THESES SIS/LIBRARY
R.G. MENZIES LIBRARY BUILDING NO:2
THE AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA ACT 0200 AUSTRALIA

TELEPHONE: +61 2 6125 4631
FACSIMILE: +61 2 6125 4063
EMAIL: library.theses@anu.edu.au

Interlanguage Phonology:
Acquisition of timing control and perceptual categorization of durational contrast in Japanese

by

Takako Toda

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

of The Australian National University

Canberra

1996

January

To my mother and father with gratitude

Except where otherwise acknowledged,
the work contained in this thesis
is based on my own research.

Takako Toda

25.1.96

Acknowledgments

The main questions addressed in this thesis arose in 1986 when I started teaching the Japanese language to students in Australia. Over a period of nine years, I have met many students who shared common problems related to their pronunciation in Japanese. Before I officially started my PhD research, my interest kept growing as I wondered why my students had similar problems and how I could help them improve their pronunciation.

It is a privilege to acknowledge that it was Dr A.E. Backhouse and Professor H.C. Quackenbush, who motivated me to start this project. I am grateful for their continuous guidance and encouragement. I would also like to express my gratitude to my fellow colleagues and students at Japan Centre, Australian National University, for giving me an opportunity to develop ideas behind this project.

Dr. T. Liddicoat of the Department of Linguistics motivated me further by the publication of my article in the Australian Review of Applied Linguistics, journal of the Applied Linguistics Association of Australia. I would also like to thank Dr. K. Rensch for his comments on the earlier version of this thesis.

This project would not be possible without informants generously contributing their time. I deeply appreciate their understanding of my aim and being so cooperative. My special thanks to the staff members of the Department of Foreign Affairs in Canberra and the Australian Embassy in Tokyo for the contribution of their valuable data, and especially to those who patiently participated in the longitudinal data collection.

Thesis writing can be a lonely process and you can easily be isolated from the rest of the world. I am grateful to Professor T. Ayusawa of the National Language Research Institute and Professor K. Nagatomo of the Ochanomizu University for kindly inviting me to join the Tokyo Association of Communication and Language and the Second Language Acquisition Research Forum. They provided me with opportunities to exchange ideas with fellow colleagues in Japan by presenting my papers. I thoroughly enjoyed the formal and informal discussions during and after our regular meetings, and appreciated these opportunities which kept me in contact with the academic community while I was overseas.

I extend heartfelt thanks to very special people, my husband, Peter, and my Australian host family for the warm support they extended to me over the years. Peter and Englert family have known me for a long time and have given me ceaseless support in numerous ways since I was an undergraduate student. Without their personal encouragement, this research project would not have been possible.

Finally, I am grateful to the Australian Government for providing me with a Commonwealth Post Graduate Research Award and an additional Priority Area Research Grant over four years.

Abstract

The timing organization of phonological durational contrast is known to be one of the most challenging areas in the acquisition of Japanese phonology (Sugito 1989; Muraki and Nakaoka 1990; Han 1992; Toda 1994). This study examines the acquisition of timing control and perceptual categorization of the durational contrast in Japanese, and aims to contribute to second language acquisition theory from the viewpoint of interlanguage phonology.

Acoustic techniques were used to investigate the mechanisms of learners' speech perception and production. In order to capture the acquisition processes and the developmental factors in the formation of interlanguage phonology, both cross-sectional and longitudinal experiments are conducted with different groups of learners at various proficiency levels, and the results were compared with those of native speakers.

The processes such as first language transfer, overexaggeration and phonetic approximation are observed in the learners' speech production. The acoustic observation revealed what has been claimed as genetically innate Universal Grammar (Chomsky and Halle 1968) pertaining to voice onset time, which is clearly observed in Japanese native speakers speech, is not operative in beginning learners' speech production. On the other hand, it is operative in advanced learners' speech which shares similarities with that of native speakers. The experimental results reported in this dissertation reveal crucial theoretical insufficiency concerning the "universal principle". This supposedly universal theory concerning voice onset time, should not be restricted to a specific language, runs into difficulties in explaining the reasons why this property does not show up with learners' speech production during the initial stages, but does show up later at more advanced stages (i.e. it is developmental, but not universal).

Variations in learner strategy were also discussed to propose a typology of learner strategies. Particularly, syllable modification strategy provided us some insight to the interlanguage syllable structure.

The findings of this study lead us to question the common assumption that the acquisition of geminate stops is a major problem for learners of Japanese (Han 1992). The experimental results obtained in this study reveal various aspects which demonstrate that this assumption does not give a full understanding in respect of learners' problems. For example, at the level of speech production, the problem of underdifferentiation is often caused by the longer duration of a single intervocalic stop, rather than the shorter duration of a geminate intervocalic stop. This problem would not have been discovered if the observations were only based on the ratio analyses of the durational contrast.

The longitudinal study provides various findings which are unique only to the longitudinal observations. For example, this study finds that the perceptual boundaries are not modified throughout the year, although the degree of perceptual categorization improved over one year (thus the degree of uncertainty involved in their perceptual judgment decreased). The crucial problem is that the learners established an inaccurate perceptual target, and their perceptual judgment gradually became more accurate in relation to their inaccurate perceptual target. This finding suggests that, unless learners' perceptual boundaries become similar to those of native speakers, their improvement in categorical perception alone cannot be regarded as acquiring native speakers' norms.

Additionally, experiments were conducted to apply the implications of the research findings to teaching situations, and to investigate the extent to which instruction can facilitate the natural processes of second language acquisition. A pronunciation lesson was given to learners, and the results were compared before and after the lessons concerning both speech perception and production. Some improvements were observed in certain, but not all, areas. The details of the improvement, and the reasons why this did not occur in certain areas are discussed. The results demonstrate that it is important for instructors to emphasise the appropriate timing control for learners, rather than just focussing on making a clear differentiation of the durational contrast in the classroom.

Finally, three areas of second language acquisition research, which are worth investigating for their likely contribution, are suggested for future study.

Chapter 1 Introduction

1.1 Background to interlanguage research

In the last two decades, interlanguage (IL) research has developed as a new conceptual framework in second language acquisition (SLA) research. Language acquisition research has gone through different stages. From the 1940s to the 1960s, research was dominated by the contrastive analysis hypothesis (CAH). Researchers were motivated by statements such as: 'The most efficient materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner' (Fries 1945:9). Within the research framework based on CAH, it is assumed that comparison between the learners' first language (L1) and the second language (L2) or the target language is a crucial key to predicting the areas and degrees of difficulties that learners will find in their learning, and also the errors they are likely to make.

Lado (1957) established the view that ease of learning was largely due to positive transfer and difficulty was attributed to negative transfer:

> ...in the comparison between native and foreign language lies the key to ease or difficulty in foreign language learning....(1957:1)
>
> We assume that the student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be easy for him and those elements that are different will be difficult. (1957:2)

This view shares a similarity with a comment by Weinreich (1953):

> The greater the difference between two systems, i.e. the more numerous the mutually exclusive forms and patterns in each, the greater is the learning problem and the potential area of interference. (1953:1)

It was found, however, that L2 learners also made errors which could not be attributed to L1 interference. For example, based on data obtained from 145 Spanish-speaking children learning English, Dulay and Burt (1973) found that only 3% of the syntactic errors were attributed to L1 interference while 85% of the errors were developmental. Such findings showed the limitations of a strong version of CAH. CAH certainly predicted some errors, but not all.

Consequently, a distinction between a strong version and weak version of the CAH was suggested by Wardhaugh (1970). In the weak version, researchers first observed errors made by learners and tried to explain those errors referring to the differences and similarities between the learners' L1 and L2 or the target language. This approach is known as error analysis and is different from the strong version of CAH which assumed that learners' errors in L2 or the target language could be predicted based on an *apriori* contrastive analysis of the L1 and the target language.

Error analysis; however, was criticised and fell into disfavour for a number of reasons (Schachter and Celce-Murcia 1977). Research findings demonstrated aspects of language acquisition which were not captured by error analyses. Larsen-Freeman and Long (1991:61) state: 'By focusing only on errors, researchers were denied access to the whole picture. They studied what learners were doing wrong, but not what made them successful. Furthermore, it was often difficult, if not impossible, to identify the unitary source of an error.' For example, Schachter (1974) reported that Japanese and Chinese learners of English made fewer errors in relative clauses than Spanish and Persian learners. This result goes against predictions based on an *apriori* contrastive analysis of the L1 and the target language. The reason behind this result was that Japanese and Chinese learners avoided using English relative clauses because they were aware that English relative clauses were difficult for them. Error analysis alone does not account for the avoidance of English relative clauses by L2 learners, and thus fails to capture areas of difficulty which L2 learners come across.

Instead of just focussing on errors L2 learners make, researchers started to analyse L2 learners' performance. Some longitudinal studies were conducted, and these studies made their contribution to the finding of developmental sequences. Another area of inquiry was an investigation of learner strategies. This type of analysis is called a performance analysis. Through the observation of developmental sequences, researchers found that the developmental sequences of L2 learners shared similarities with those of child learners acquiring the same language as an L1. For example, Milon (1974) found that a Japanese child followed a similar pattern of English negation produced by children acquiring English as an L1.

It was also found that many language learners who had different L1s showed similarities in their learning behaviours. Their errors were not predicted in terms of language transfer from their L1, but were often similar to errors made by child learners of their mother tongues. These findings provided support for the assumptions of universalist linguistic theory. For example, Tarone (1972) claimed a universal tendency towards open syllables: in her view, CV is a basic unit in all languages (according to Tarone, for both production and perception) and learners tend to break complex syllable structures into simple CV units regardless of their L1.

In order to investigate the nature of learners' language acquisition, the concept of IL has been introduced and developed in the area of SLA research. The underlying assumption of IL research is that learners' language is an entirely functional linguistic system, which is independent of both learners' L1 and the target language. The errors made by learners are systematic, and are useful indicators of their levels of competence. This conceptual framework has allowed us to observe various aspects of the unique characteristics of L2 learners' language. This is not to claim that L1 interference is unimportant; it clearly is, and L1 interference upon IL is much more prominent at the level of phonetics and phonology compared with that of morphology or syntax (Felix 1980; Odlin 1989).

Corder (1971) considers learners' language as a set of idiosyncratic dialects and calls them 'transitional dialects'. Nemser (1971a) calls it an 'approximative system'.

The term 'interlanguage', which is currently most widely used, was first introduced by Selinker (1972). Despite this difference in terminology, the basic concept referred to is more or less the same. The basic concept which is common to all of these terms is that the learners' language is not the same as their L1 or the target language, but is a dynamic continuum which should be treated as an independent linguistic system. They all regard learners' activity as a creative process, rather than a set of processes initiated by L1 transfer.

Corder's (1971) notion of transitional dialects is illustrated as a system which is distinct from both L1 and target language (TL) in Fig.1.

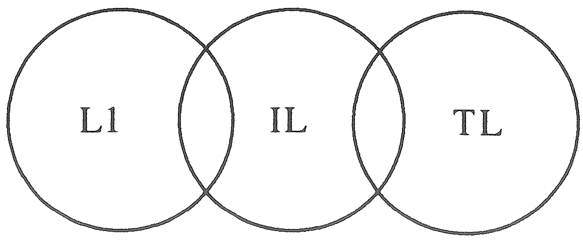


Fig.1 IL as a transitional system between L1 and TL (Corder 1971)

IL is also seen as a self-restructuring or self-recreating continuum (Corder 1978). Restructuring refers to the process whereby learners gradually substitute the target language for their L1 rules, assuming that the starting point of SLA is the L1. Recreating refers to the stage where the learner starts learning the target language by using a simple system which is independent of the L1. According to Ellis (1990:51): 'Most interlanguage theorists see interlanguage as a recreating continuum, although they do not eliminate the possibility of transfer.'

While a great deal of work has been conducted on syntactic and morphological aspects of SLA, investigations of IL phonology offer relatively limited data compared to those of other areas of IL research. Even now the number of empirical works which investigate aspects of IL phonology through acoustic data using instrumental techniques remains limited:

> The phonology of interlanguage is an area which was largely neglected by second language acquisition research until very recently. There seemed to be very little interest in the pronunciation patterns of the speech of second language learners. When Schumann summarized existing second language acquisition research in 1976, he found absolutely no studies on the phonology of interlanguage (Tarone 1978).

> In the history of SLA research, a great deal of attention has been paid to grammatical domains of language acquisition (morphology/syntax), yet until the last decade, the acquisition of L2 phonology remained relatively neglected (Enomoto 1992).

One reason why research on IL phonology has not attracted much attention could be the assumption that pronunciation is largely influenced by the speaker's native language, thus the results obtained from the data are uninteresting. This is not at all true. As Tarone (1978:70) points out: 'the research which has been done in this area quite clearly shows that transfer is only a part- and often a small part- of the influence on interlanguage phonology.' Another reason may have been the amount of difficulty involved in the research methodology and the accessibility of acoustic instruments. Recent developments in acoustic techniques, however, have allowed many researchers to investigate acoustic-phonetic characteristics of the processes involved in language acquisition, adding a new dimension to IL research. For example, identifying acoustic differences in French /t/ and /u/ between native and English speakers (Flege 1981, 1984) is only made possible through acoustic quantification of voice onset time (VOT) and formant frequencies. More acoustically-based research needs to be conducted to examine the phonetic details in order to obtain a deeper understanding of the nature of IL phonology.

Studies of IL phonology seem to have taken two major directions, the first perspective emphasising L1 transfer, and the second universal tendencies. The former originated in CAH (Lado 1957). Though not framed in a strict version of CAH, some recent works demonstrate that learners' linguistic experience of L1 in some way influences the acquisition of phonology in L2 (Flege 1979b, 1980, 1981, 1984, 1987bd; Flege and Hillenbrand 1984; Odlin 1989, Kuhl 1993; Toda 1994, 1995ab). Whereas the strong version of CAH claims that differences between L1 and L2 can predict areas of difficulties learners encounter, the weak version of CAH holds merely that differences can account for some errors they are likely to make.

The second approach originated in the structuralist tradition associated with Jakobson's (1968[1941]) theory of the acquisition of phonological contrast. Within the framework of generative phonology, the universalist/nativist approach proposed by Chomsky and Halle (1968) and the theory of Natural Phonology (Stampe 1969) presented sets of rules which are based on a proposed universal hierarchy of phonological acquisition. The markedness differential hypothesis (MDH) proposed by Eckman (1977) provides an alternative model to CAH, relating the notion of typological markedness[1] to relative degrees of difficulty (which is independent of L1 and thus universal). Furthermore, Fellbaum (1986) proposed a revised version of MDH to account for the acquisition of allophones. Phonetic universals which are based on human articulatory constraints have also been investigated in the domain of experimental phonetics (Lisker and Abramson 1964; Ohala 1974, 1980), although there is no set of empirical outcomes which accurately predicts various aspects of phonological output.

Language transfer and universal theory seem to be in many ways conflicting theoretical approaches, both in terms of their perspectives and also from the viewpoint of their historical backgrounds. Presently it is known that neither language transfer nor universal theory alone provides a complete account of the processes of language acquisition, and in fact, various processes from both of these domains have been claimed as operative in the formation of IL phonology (Tarone 1978).

[1] Typological markedness is defined in Eckman (1977:320) as "[a] phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B, but the presence of B does *not* imply the presence of A".

One of the main goals of the present study is to examine what processes are involved in the formation of English-Japanese IL phonology. Tarone (1978) summarises the processes which were reported in earlier studies as follows:

1. negative transfer from L1 (all studies)
2. L1 acquisition processes (Wode 1976; Tarone 1978)
3. overgeneralization (Johansson 1973)
4. approximation (Johansson 1973; Nemser 1971ab)
5. avoidance (Celce-Murcia 1977)

The extent to which phonological transfer and universal phonetics operate in shaping IL has not yet been defined. Further investigation in both dimensions will lead to a more thorough understanding of the nature of the processes which take place in shaping IL.

1.2 Problems

In previous studies which examined the acquisition of L2 phonology from the viewpoint of L1 transfer or universal linguistics, the interrelation between perception and production in the acquisition processes has not been thoroughly investigated. The majority of earlier works addressed questions on aspects of IL phonology based on either production, such as informative discussions of English-Japanese IL syllable structure by Tarone (1980) and Sato (1984), or perception, such as the degree of categorical perception of durational contrast in Korean-Japanese IL (Min 1987), English-Japanese IL (Hirata 1988, 1990b; Enomoto 1992) and Chinese-Japanese IL (Uchida 1993a).

Flege (1981, 1984) hypothesises that a foreign accent is not entirely due to difficulties involved at the level of speech production, but also to the cognitive mechanism of categorizing the sound in the target language as a part of an already existing category of the speakers' L1. For example, if English learners of French identify French /t/ as belonging to the same category as English /t/, then they produce French /t/ with a relatively long VOT. This hypothesis also applies to the production of French vowels. Despite the differences in the acoustic signals between the English and French /u/, mental representations of French /u/ by English learners tend to be related to English /u/. As a result, it is difficult for French listeners to distinguish /y/ and /u/ spoken by English-speaking learners of French. This is the so-called 'equivalence classification' (Flege 1987d) which causes the development of inaccurate perceptual targets. Equivalence classification is based on the assumption that learners' cognitive processes are influenced by previous experience when acquiring new information (Schults 1960; Ausubel et al. 1978). Although 'equivalence classification' alone does not account for the entire picture of foreign accent, the contribution of these studies is that they emphasise the role of perception in L2 acquisition research.

The recent model of perceptual assimilation (Best 1992; Best and Strange 1992) introduced the degree of category goodness that L2 learners can perceive and judge how the given L2 stimuli fit into their L1 category. It assumes the following four patterns of perceptual assimilation (Best and Strange 1992:306) :

1) the two members of the non-native contrast may be assimilated into *two categories* in the native phonology

2) both non-native phones may be assimilated equally well (or poorly) into a *single category*

3) both may be assimilated into a single category, but unequally, thus showing a *category goodness difference* in their fit to the native phoneme

4) the non-native phones may differ so much from the phonetic properties of native phonemes that they are *non-assimilable*.

This model differs from Flege's model of equivalence classification from the respect that it does not assume a given L2 stimuli to be either fully classified as an equivalent of a L1 phoneme category, or not to be equated at all. Consequently, the focus of perception research has shifted towards addressing questions concerning the ways in which L2 contrast are perceived, rather than whether or not the perceptual pattern is categorical or continuous.

The key to understanding the L2 acquisition process is to observe the mechanisms of both production and perception. Otherwise, it is not possible to determine whether the learners' problems are due to articulatory difficulty (at the level of speech production), or inaccurate perceptual categorization (at the cognitive level). This also relates to the issue of whether perceptual acquisition precedes production. Generally speaking, it is accepted that perceptual ability precedes acquisition of speech production, although there are some studies which argue that production may precede perception in the differentiation of /l/ and /r/ in Japanese (Goto 1971; Sheldon and Strange 1982). To the best of my knowledge, however, no previous studies appear to provide any empirical evidence which accounts for the interrelations between production and perception in the process of acquisition of L2 phonology.

Additionally, there are very few detailed acoustic observations concerning IL phonology which have been gathered on a longitudinal basis. Cross-sectional data collection has the advantage of requiring less time, and is seen as simulating observations of development at different stages of the progress of the same informants. Strictly speaking, however, true 'developmental changes' are not captured through cross-sectional studies since the data do not represent the progression of the same informants over time, and generalization from cross-sectional studies has been questioned (Reichardt and Cook 1979; Cook 1986; Larsen-Freeman and Long 1991).

This study uses both cross-sectional and longitudinal data. Through longitudinal data, developmental changes in the learners' speech production and perception are examined as their levels of proficiency increase. Cross-sectional studies are also undertaken with larger groups of subjects, supplementing the relatively smaller group of subjects who participated in the longitudinal study.

1.3 Aims of the study

The ultimate aim of the present study is to contribute to the theory of SLA from the viewpoint of IL phonology. Acoustic techniques were used in the research to achieve the following goals:

1) to examine what processes are involved in the formation of IL phonology

Current theories of IL phonology have been influenced by two major theoretical frameworks: one emphasises L1 transfer (in which negative transfer is referred to as L1 interference); the other claims universal tendencies.[2] In this dissertation, experiments were conducted to investigate further the domain of phonetic universals and of language transfer in the formation of English-Japanese IL.

2) to investigate perceptual mechanisms

In order to investigate the mechanisms by which listeners perceive durational contrast in Japanese, a set of computer-edited stimuli was created. The differences between native speakers (NS) and learners at different levels are examined. An underlying assumption is that differences in the results imply that perceptual categories amongst these groups are not identical.

3) to analyse speech production

With regard to the domain of speech production, this study focuses on the acoustic investigation of timing organization in IL. The ways in which learners manipulate segment duration in the phonetic realization of phonological long/short contrasts are observed, and syllable modification strategies observed in learners' IL syllable structure are discussed. Moreover, observations are made in order to examine how characteristics of learners' perceptual mechanisms interrelate with those of their speech production.

4) to examine developmental changes in IL phonology

---

[2] 'Universal' used in this dissertation follows the definition in Ioup and Weinberger (1987:420) "A cross-linguistic phonological pattern which has a high probability of occurrence, though it may admit exception. Many of these universals derive from the specific properties of the human articulatory and perceptual systems". (e.g. VOT increases as the place of articulation moves toward the back.) This definition of universalness, however, does not necessarily apply in the learners' speech production in the same way as native speakers of their target language.

Cross-sectional experiments were conducted with groups of beginners and advanced learners in order to compare them with a group of NS. A further group of learners participated in a longitudinal study from beginning to intermediate stages. Observations are made on developmental factors in the formation of IL phonology.

5) to observe variations in learner strategies[3]

The variety of learner strategies used to make long/short contrasts in Japanese is observed in order to draw some generalizations concerning variations in learner strategies, although the individual learner's acquisition process of L2 phonology may vary in a number of ways. This study aims to propose a typology of learner strategies in speech production.

6) to offer some insights for language teachers

A further important reason for conducting IL research is to apply the research findings to improve teaching methodology. According to Han (1992:102), 'Students of Japanese as a foreign language are puzzled and frustrated as they are repeatedly corrected on the pronunciation of geminate consonants. Experienced instructors know that the timing control of geminate consonants in Japanese is one of the most difficult phonetic skills for the students to master.'

Experiments were conducted to examine the extent to which efficiently taught SLA facilitates the natural processes of L2 acquisition. A pronunciation lesson is given to a group of students, and the results of perception and production tests are examined before and after the lesson.

1.4 Organization of the study

This thesis consists of eight chapters. This first introductory chapter has specified problems in the research area and identified the goals of the study.

Chapter 2 discusses previous works related to Japanese timing organization. The chapter first introduces the mora, an important concept in dealing with Japanese

[3] The term, variation, is often used to refer to sociolinguistic variations; for example, formal and informal context (Schmidt 1977; Beebe 1980). In this study, however, variation does not refer to any sociolinguistic context.

phonology. Discussion then follows on the question of the acoustic reality of the mora, and on its perception.

In Chapter 3, perception of long/short contrasts in obstruents, vowels and nasals is investigated by using computer-edited sound stimuli. The experiments are first conducted with a group of NS, and then with beginners and advanced learners of Japanese.

Chapter 4 examines the timing organization of Japanese from the viewpoint of speech production. This section investigates the processes which are involved in the shaping of IL phonology of Japanese. This is then followed by a discussion of the interrelation between characteristics of English-Japanese IL phonology and those of speech perception described in Chapter 3.

In Chapter 5, a longitudinal study is conducted with a group of learners undergoing intensive language training. This chapter presents findings concerning perception obtained during one year's observation. Chapter 6 deals with findings concerning speech production for the same group of subjects.

Chapter 7 deals with implications of the research for teaching. A model text is created to test the extent to which a set of instructions which focus on pronunciation can improve learners' perception and production. This is an additional chapter which relates the issues discussed in the previous chapters to teaching situations, but determining the effects of different teaching methodologies upon learners' performance is beyond the scope of this dissertation.

The final chapter presents a summary of the major results of the study, followed by an overall discussion.

Chapter 2 Linguistic theories and acoustic research on the mora

This chapter is a review of previous work concerning the mora. The concept of mora is first introduced from the viewpoint of phonology. This is followed by a consideration of literature concerning the acoustic-phonetic reality of the mora.

The aim of this chapter is to discuss some characteristic features in the timing organization of Japanese. It will consider issues raised in the previous literature, including studies on the acquisition of timing organization and perception of durational contrast by non-native speakers (NNS) of Japanese. Against the background of this material, the chapter is then followed by the data obtained in the present study.

2.1 Phonological concept of the mora

2.1.1 Syllable and mora

Many of the early works pertaining to the prosodic characteristics of various languages are typologically motivated. The studies by Pike (1945) and Abercrombie (1967) classified languages into two broad categories: 'stress-timed' and 'syllable-timed' depending on whether a given language is with or without lexical stress. Japanese was first classified into the latter category (Hockett 1955). Later, a third category, 'mora-timed', was distinguished from 'syllable-timed' languages (Ladefoged 1982[1975]) and Japanese was reclassified as a mora-timed language.[4]

The concept of mora is found in the work of Trubetzkoy (1969[1939]), a linguist of the Prague School, who drew distinctions between 'syllable-counting' and 'mora-counting' languages. Trubetzkoy considered the way in which long syllable nuclei were treated in the phonological analysis as the key to this distinction:

> The interpretation of long syllable nuclei as geminated, or in terms of multimember constituency in general, may be regarded as an 'arithmetic conception of quantity'. Languages

[4] According to Yoshiba (1983), J.W.Donaldson was the first person to use 'mora' in his book Greek Grammar (1848).

in which this conception finds expression are 'mora-counting' languages since in these languages the smallest prosodic unit does not always coincide with the syllable. (1969:177)

On the other hand, languages in which the smallest prosodic units coincide with the syllable were regarded as 'syllable-counting' languages. According to his definition, languages such as Finnish, Japanese and Lithuanian are 'mora-counting', and English, German and Dutch are 'syllable-counting' languages.

In Trubetzkoy's terms, the smallest prosodic unit of a language is called a 'prosodeme'; the prosodeme of a syllable-counting languages is a syllable and that of a mora-counting language is a mora. The prosodeme is differentiated in a syllable-counting language by intensity, and in a mora-counting language by pitch. Therefore, according to this system, accent of syllable-counting languages is defined in terms of intensity, while that of mora-counting languages is defined in terms of pitch.

Isochrony is often referred to in the definition of the mora. Ladefoged (1982[1975]), for example, emphasises the isochronous aspect of timing organization in Japanese and explains this characteristic feature as follows:

> Probably one of the most interesting languages in the way that it uses length is Japanese. Japanese may be analysed in terms of the classical Greek and Latin unit called a mora. A mora is a unit of timing. Each mora takes about the same length of time to say. (1982[1975]:226)

In this view, the phonological concept of mora is directly related to duration, which is a phonetic property. The isochronous aspect of the mora can also be extended to word-level duration (Ladefoged 1982[1975]):

> The most common type of Japanese mora is formed by a consonant followed by a vowel. Japanese words such as [kakemono] (scroll) and [su̥kiyaki] (beef stew) each consist of four morae of this type. Note that in the latter word the high vowel /u/ is voiceless because it occurs between two voiceless consonants; but it still takes about the same length of time to

say as do the vowels in the other syllables. Another type of mora is a vowel by itself, as in the word [iki] (breath). This word has two morae, each of which takes about the same length of time to say. A consonant cannot occur after a vowel within a mora, but it too can form a mora by itself. The word [nippoŋ] (Japan) must be divided into four morae [ni p po ŋ]. Although it has only two vowels, it takes approximately the same length of time to say [nippoŋ] as it does to say [kakemono] or [su̥kiyaki]. (1982[1975]:226)

The importance of duration in the definition of mora has commonly been emphasised, as seen in Hockett's (1955:59) earlier claim that mora is 'defined fundamentally in terms of duration and nothing else'.[5]

In traditional Japanese grammar, the temporal system has long been claimed as being governed by mora timing (Bloch 1942; Hockett 1955; Hattori 1960; Kindaichi 1967). A basic unit which constitutes one mora is CV or V. For example, *kokoro*[6] 'heart' is simply segmented as *ko-ko-ro* and thus has three morae.[7] When the number of syllables agrees with the number of morae as in this example, there do not seem to be any problems. Scholars have claimed slightly different views, however, with regard to the treatment of a phonological category called *tokushuhaku* 'special morae', which are represented as Q (first half of geminate obstruents), N (moraic nasals), and (V)V (the last half of long vowels). The phonetic values of the special consonantal morae are predictable in terms of assimilation with the consonants which follow. Phonetically speaking, /Q/ is realized as [p], [t], [k], [s], [ʃ] or [tʃ]; /N/ as [n], [m], or [N] depending on the environment.[8] /N/ can appear word-finally, and it can be realized as [N] such as [paN] for 'bread'. It may also be realized as a nasalised version of the preceding

[5] Later Lehiste (1977) reconsidered the definition of isochrony, and the distinction was drawn between phonetic isochrony and perceptual (psychological) isochrony.

[6] The system of romanization used in the text is the Hepburn system.

[7] Both *morae* and *moras* are widely used as a plural form of *mora*.

[8] Despite the phonetic differences, all moraic obstruents are written as a small *'tsu'* and all moraic nasals as an *'n'* in *kana* orthography. Japanese NS seem to regard all sounds written in the same *kana* orthography as one unit. According to J. Tsuchiya (personal communication), some Japanese NS who do not have a linguistics background consider that moraic [s] in *issho* [iʃʃo] 'together' in Japanese is a silence, which is the same as stop consonant [k] in *ikko* [ikko] 'one (small miscellaneous object)'. They do not recognize the fact that the [s] is a continuum of a fricative noise, and tend to feel that there is a silent closure period in [iʃʃo] just the same way as in [ikko].

vowel, [paã], or a nasalisation of the preceding vowel, [pã]. The former ([paN]) tends to occur in careful speech while the latter ([paã] or [pã]) in less careful speech. /N/ also occurs before a vowel or semi-vowel such as /teNin/ 'shopkeeper' or /koNwaku/ 'confusion'. In this situation, /N/ is a nasal sound which is assimilated to the following sound such as [teĩin] and [kow̃waku].

Arisaka (1957) was one of the first Japanese linguists to suggest a distinctive treatment between the 'phonetic syllable' and the 'phonological syllable'. In order to analyse Japanese phonology, he claimed that it was necessary to have a unit 'phonological syllable', which was distinct from the so-called 'syllable' in Western literature (Arisaka defined this as 'phonetic syllable'). Hattori (1960) then suggested a distinction between the syllable and the mora.

Another important claim that Hattori (1960) made was that the phonetic duration of utterances was determined by the number of morae, which is the same as the description by Ladefoged (1982[1975]) mentioned previously. According to his claim, *Nippon* 'Japan' can be divided as niQ-poN or ni-Q-po-N and therefore has two syllables or four morae. The duration of all the two-mora items shown below is supposed to be identical:

1. [koto]	/koto/	CVCV	'Japanese musical instrument'
2. [ko:]	/koo/	CVV	'turtle shell'
3. [koŋ]	/koN/	CVN	'navy blue'

In terms of the number of syllables, however, 1 is the only disyllabic item, while 2 and 3 are monosyllabic. In other words, Hattori (1960) allows a mismatch between the number of morae and syllables. The basic structure of Japanese syllables in his view is thus described as /(C)(y)V/, /(C)(y)VV/, /(C)(y)VN/, /(C)(y)VQ/. The mora, on the other hand, can be represented as /(C)(y)V/, /N/ and /Q/.

Hattori (1960) argues that a mora is an independent unit which is distinct from the syllable, and thus the number of morae does not necessarily correspond with the number of syllables. Word duration is mainly determined by the number of morae when spoken in a consistent speech tempo. In this claim, he argues that the phonological concept of mora is directly related to phonetic duration. Additionally, Hattori (1960) accepts discrepancies between the number of syllables in slow and fast speech; he treats a word which contains devoiced vowels such as *hashi* [haʃi̥] 'bridge' as basically a disyllabic expression with two morae, but in fast speech as a monosyllabic expression (yet still with two morae).

Hattori's (1960) suggestion, however, brought confusion to Japanese linguists, amongst whom the tradition of treating his 'mora' as *onsetsu* 'syllable' was already established practice (i.e. they had seen no need to distinguish the 'syllable' from the mora). In order to avoid this confusion, other scholars adopted different approaches to this issue. Shibata (1962) suggested the use of 'syllabeme' for the phonological syllable in Hattori's (1960) definition, and 'mora' for the analysis of pitch. Kindaichi (1967), on the other hand, suggested calling Hattori's (1960) mora *haku*, adopting a musical term (translated as 'beat' in Vance 1987).[9] In Kindaichi's (1967) view, each basic phonological unit, represented by a single symbol in the *kana* orthography consists of one *haku*, which is also a basic metrical constituent of traditional Japanese poetry.[10] Using this concept, Kindaichi (1967) also analysed the system of pitch accent in standard spoken Japanese and its various dialects. The basis of his analysis is slow, careful speech, in which he claims that each mora has equal duration. He treats /ni-Q-po-N/ as an expression with four *haku*, and all of Hattori's (1960) examples given previously, [koto] /koto/, [ko:] /koo/ and [koŋ] /koN/, are treated as words with two *haku*.

The above section discussed the syllable and mora. The question is, whether we need syllable, mora or both in the analysis of Japanese phonology. In the analysis of Japanese accentuation, it is a widely accepted view that both mora and syllable

---

[9] The term *haku* is first used by Kamei (1956). Kindaichi (1967) recognized the usefulness of Kamei's terminology, and suggested using *haku* in the phonological description and *onsetsu* in the phonetic description.

[10] Palatalised consonants such as *kya*, *nya* are written as a combination of 'ki' or 'ni' with a small 'ya' in *kana* orthography and they are also treated as one mora. Thus both *ka-shi-ya* 'rented house' and *ka-sha* 'cargo train' are written with three *kana* symbols (the latter, however, has a small '*ya*'), but the former has three *haku* while the latter only has two.

are necessary units. Traditionally, it has been claimed that one of the main justifications for recognizing the mora as a necessary element in analysing the Japanese accentual system is that it is a tone bearing unit. In modern standard Tokyo speech, the initial tone is restricted to one-mora length. Thus, the tone assigned to the first mora is always different from that of the second mora. There are minimal pairs of disyllabic expressions with contrastive pitch patterns, such as *hashi* (HL) 'chopsticks' and *hashi* (LH) 'bridge', or *kaki* (HL) 'oysters' and *kaki* (LH) 'persimmon' (H: high tone; L: low tone). However, for most Tokyo speakers, there is no pitch rise from L to H across two sonorant mora boundaries within a syllable. For example, *shimbun* 'newspaper' which is phonemicized as /siNbuN/ would be pronounced as HHHH rather than LHHH (i.e. L tone is assigned to /si/ and an H tone to /N/ within one syllable). Recent instrumental studies also show some acoustical data which support the above claim (Pierrehumbert and Beckman 1988).

McCawley (1968) analyses accentuation in Japanese in terms of pitch while incorporating the concept of mora as a unit of phonological distance. In McCawley's analysis, the number of morae determines the placement of accent (i.e. the mark of pitch fall). The third last mora is the default location of Japanese accent and this tendency is shown even in meaningless words or English loan words as shown below (1968:133-134):

[kakikúkeko] LHHLL (a line of the *kana* syllabary)

[pádʒama] HLL pyjamas

The syllable is also an inevitable notion in Japanese accentuation because the accent is only assigned to syllabic units. If the third last mora is non-syllabic, the accent is shifted to the previous mora which is syllabic. In other words, two-mora syllables always have the accent on the first mora (Shibatani 1990):

[kóorogi] HLLL cricket

* [koórogi] LHLL

This example demonstrates that the mora in isolation cannot account for the system of Japanese accentuation without the notion of syllable. Therefore, it can be said that both mora and syllable are important in the analyses of Japanese accentuation.

2.1.2 Mora as a metrical unit

Another characteristic feature of the mora which has been attracting a great deal of attention in recent literature is that it plays an important role as a metrical unit in traditional Japanese poetry and folk songs.[11] Rhythm in Japanese poetry has frequently been discussed in the literature on phonology (Bekku 1977; Nakamichi 1980; Homma 1982; Jooo 1988: Beckman 1994). In recent literature, the bimoraic hypothesis proposed by Poser (1990), which claims that a basic Japanese rhythmic unit consists of two morae, encouraged researchers to examine the extent to which NS treat two mora as one unit. For example, Kurisu (1994) examines the perceptual segmentation of words which are extended to six morae, and claims that even-numbered segmentation (or multiples of a bimoraic foot) in Japanese is supporting evidence for the bimoraic hypothesis from the viewpoint of speech perception.

One of the first linguists to claim an early version of bimoraic rhythm in Japanese poems was Bekku (1973; 1977). Then Nakamichi (1980) claimed that the basic Japanese rhythmic unit was a *takt*, a unit consisting of two morae. In his view, four takts make an ideal Japanese rhythm. There is either a one- or a three-beat pause at the end of each line as shown below in a short poem (*tanka*), which expresses the admiration of Matsushima's (place name) beauty:

ma.tsu | shi.ma | ya.# | # # |

a.a | ma.tsu | shi.ma | ya.# |

ma.tsu | shi.ma | ya.# | # # |

[11] Lehiste (1989, 1990, 1994) conducted extensive research on poetic metrics in Estonian, Swedish, Latvian, Lithuanian, Faroese, Icelandic, Hungarian and Serbocroatian, and concluded that 'speakers of different languages behave in very different ways when they produce poetic metres that bear the same label and thus may be considered similar (if not identical), and that the differences in their productions are related to the prosodic structure of the languages.' (1994:2243)

(. moraic boundary, | rhythmic boundary, # pause)

A pause can also occur within an initial bimoraic unit as described in the following example in Jooo (1988:30). According to Jooo, this initial pause is deliberately used to break the repetition of a simple rhythm, as shown in the famous poem by Basho Matsuo:

fu.ru | i.ke | ya.# | # # |

#.ka | wa.zu | to.bi | ko.mu | Initial pause

mi.zu | no.o | to.# | # # |

Without an initial pause, the rhythm of this poem could be reinterpreted as follows:

fu.ru | i.ke | ya.# | # # |

ka.wa | zu.# | to.bi | ko.mu |

mi.zu | no.o | to.# | # # |

Some NS may feel that one is more acceptable than the others. It seems that a great number of idiosyncrasies are also involved in the treatment of metrical boundaries between two consecutive bimoraic units. This area requires further investigation.

The basic concept of the bimoraic rhythm is applied in the pronunciation lesson conducted in Chapter 7.

2.2 Acoustic investigations on mora

2.2.1 Phonetic isochrony of mora

In recent years, research interest in the mora hypothesis has increased, and a number of phoneticians have examined the phonetic reality of the mora. Through acoustic observations, it has been found that the strict version of the mora hypothesis, which claims that each mora is strictly isochronous, does not capture the phonetic reality (Beckman 1982b; Hoequist 1983ab; Port et al. 1987; Toda 1991). It is also known that syllable duration in Japanese speech is not strictly proportional to the number of morae in the syllable (see 2.2.2 for further details).

In the strict version of the mora hypothesis, the duration of the first half of geminate consonants needs to be one mora length. An alternative to the strict version of the mora hypothesis is a weak version. In the case of a single [k] (e.g. *iken* 'opinion') and a geminate [kk] (e.g. *ikken* 'one (house)'), if the duration of the [kk] is more than twice that of the [k], it means that the [kk] is not merely the equivalent of a geminated single [k] (i.e. [kk]>[k]x2). This means that there must be a larger unit, namely a 'moraic unit'. Beckman (1982b:118) claims that moraic consonants need 'only to be longer than the consonants in the CV syllable'. Also, Vance (1987:71) suggests that 'as long as the average duration of a phonetic long stop is significantly longer than twice that of a short stop, we can maintain the weaker claim'.

2.2.2 Acoustic evidence for the mora hypothesis

While the phonetic reality of a strict version of the mora hypothesis has been questioned, it was found that some phenomena were difficult to account for if the concept of mora was not accepted as an abstract timing unit in the phonology of Japanese. The supporting phonetic evidence for the mora hypothesis presented in previous works (Han 1962, 1992, 1994; Port et al. 1980, 1987; Homma 1981; Bradlow et al. 1995) consists of three main categories: segmental duration, the domain of temporal compensation, and the relationship between word duration and the number of morae. The following section deals with these three categories.

Firstly, the early work of Han (1962) provided some phonetic evidence to support the mora hypothesis in terms of segmental duration. She claims that 'the duration of short and long consonants is, on average, in the ratio of 1.0 to 2.6 and often 1.0 to 3.0.' It appears that these ratios successfully support the mora hypothesis as the duration of long consonants is more than twice that of short consonants. Han also demonstrates that, when vowel deletion occurs, the duration of voiceless obstruents is longer (due to temporal compensation) than the duration of the prevocalic obstruents in CV when the vowel is maintained.

Beckman (1984), on the other hand, examined voiceless obstruents where following high vowels (/i/ and /u/) were devoiced, and argued against Han's claim that the duration of a deleted vowel was taken up by the preceding syllabic consonant. In Beckman's data, a CV mora with a devoiced vowel and a moraic voiceless obstruent are not as long as CV mora. She consequently argued against the phonetic reality of mora pertaining to syllabic consonants as follows: 'the fact that they are not consistently longer is enough to disprove even the weaker prediction of the mora hypothesis in respect to compensatory lengthening with vowel deletion' (1982:120). Later, Port et al. (1987) argued against Beckman's view, on the basis that the mora hypothesis was valid at the word level (see below).

Han's (1992) reexamination of closure duration and VOT of geminate and single stops shows that on average, the closure duration of geminate stops is 2.8 times as long as that of single stops: 'Individual differences and contextual factors cause the geminate/single ratios to vary between 2.5:1.0 and 3.2:1.0.' Thus Han (1992) confirms her 1962 results.

The results obtained in Homma (1981) also support Han (1962) by referring to the ratio of long and short consonants being as large as 3:1. Homma's (1981:279) results concerning consonant closure duration are as follows:

(1) Closure duration was longer for voiceless stops than voiced ones.

(2) Closure duration for labials was longer than apicals and velars.

(3) The ratio of closure duration between single stops and geminated stops was about 1:3. This ratio implies that the duration of geminated stops is not only doubling the stop segment but also including the length which corresponds to a larger unit, namely a mora, as Han pointed out.

Thus, Homma's (1981) results are similar to those of Han's (1962), as far as consonant closure duration is concerned, as seen in (3).

Of the three major categories examined in acoustic investigation of the mora hypothesis, perhaps the most controversial issue on which researchers have taken different views is that of temporal compensation. There are two main groups: those who consider the domain of temporal compensation as the mora (Han 1962), or those who consider it to be a larger unit than the mora, such as a sequence of morae or a word (Port et al. 1980; Homma 1981).

Han (1962) considers that the mora maintains consistent duration despite the intrinsic difference of adjacent segments. Thus, in her view, the domain of temporal compensation is the mora. In Homma's view, however, the domain of temporal compensation is a word, which is larger than the CV unit.

> Although Han wrote that a unit of duration in Japanese is associated with a syllable, it may be more appropriate to say that domain of durational pattern is not a syllable, but a word. (1981:279)

In Homma's data, the total duration of the words /papa/ (260ms) and /gaga/ (267ms) is almost the same despite the large difference in the duration of the first syllable /pa/ (85ms) and /ga/ (122ms) as illustrated below:

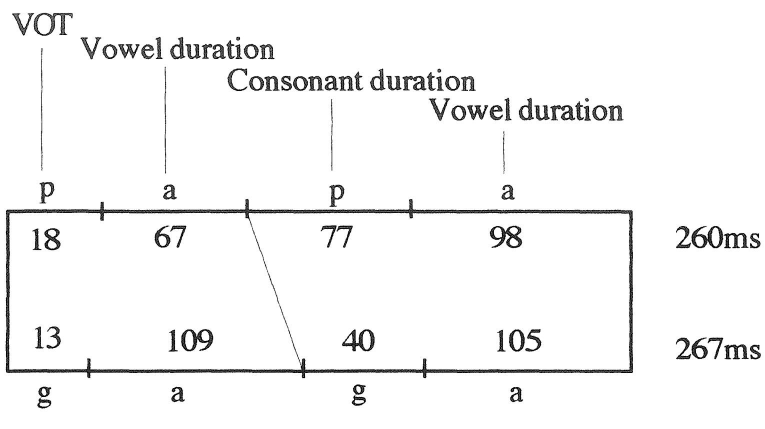


Fig.2.1 Word as the domain of temporal compensation (Homma 1981)

With regard to word duration, Homma claims that disyllabic expressions with long stops (thus three-mora expressions such as *tatta*) are one and half times as long as those with short stops (two-mora expression such as *tata*), that is, they show a

3:2 ratio which corresponds to the number of morae (see below). Closure duration, VOT, and vowel duration are closely related to fixed word duration, although the duration of each mora in a word is not necessarily the same.

A similar claim was made earlier by Port et al. (1980), namely that the domain of temporal compensation is extended over several syllables rather than just within CV as shown below:

> Our results support the traditional observation that macrostructural timing is highly regular in Japanese. They do not, however, support the specific proposal of Han (1962) that the domain of temporal compensation is strictly within the CV unit. Instead, it seems that temporal compensation extends across several syllables (if syllables are defined conventionally as CVs) so that two-syllable units, at least, have a nearly constant duration despite dramatic changes in the duration of individual consonants and vowels. (Port et al. 1980:244)

In Port et al. (1980), temporal compensation is described as bidirectional (i.e. both anticipatory and regressive): the duration of intervocalic consonants changes for both preceding and following vowels. Vowels on both sides of a consonant vary inversely with the consonant, and the consonant constriction duration varies inversely with inherent vowel duration. For example, low vowels tend to be longer than high vowels and vowels preceding voiced consonants are longer than those preceding voiceless ones (Chen 1970; Klatt 1976).

The relationship between the duration of a word and the number of morae is the third category which is referred to in support of the mora hypothesis as mentioned above. Port et al. (1987) further investigated aspects of Japanese timing organization in a larger unit such as a word. They measured the duration of test words (1-5 morae) and demonstrated that, for each additional mora, the word duration increased by equal increments. It was also found that this situation was the same in both fast and slow speech tempi. Port et al. (1987) claim that, in their results, the heavy syllables did not exhibit the expected shortening of other components of the syllable. Instead, heavy syllables showed lengthening in their results. Thus timing organization in Japanese is constrained by an abstract timing unit defined as a mora, not by a tendency to regularise syllable duration. Port et al. (1987:1576) comment that:

> This simple experiment provides clear evidence that the mora is a real unit in Japanese, since results of this sort could not be obtained from English, or German, or any other language known to us. For English, we would expect to see that syllables other than the first one would add successively less to word duration.

The most recent work of Han (1994) claims that the duration of a fricative is longer when the following vowel is deleted or devoiced (this agrees with her 1962 results and disagrees with Beckman's 1984 results). This elasticity makes the average duration of the mora consistent, keeping the durational ratio of two- and three-mora expressions 2:3 and that of three- and four-mora expressions 3:4. As seen in this study, Han (1994:81) has abandoned her previous argument, which was based on segmental duration, and agrees with the conclusions of Port et al. (1980, 1987) and Homma (1981) as shown below:

> The results of the ratio analysis definitely support the notion that the timing of segments in Japanese is planned and controlled using a word-level target. This confirms the conclusions reached earlier by Port et al. (1980, 1987) and Homma (1981) . The word in the present context is a free word pronounced without an intervening juncture. The question of boundaries at which mora timing breaks down requires further research.

As described above, many studies have been conducted on the phonetic aspects of the mora in Japanese. More recent studies have examined various aspects of Japanese timing organization in larger units such as sentences, and demonstrate that the Japanese segmental duration is influenced by the position of the segment in an utterance (Kaiki et al. 1990; Kaiki and Sagisaka 1992). Various factors seem to be operative in segmental duration control such as: lengthening the phrase-initial consonant, shortening only the final vowel in a sentence, and lengthening the final mora in a breath group (Kaiki and Sagisaka 1992:401). Other factors such as context and style of speech may affect the results concerning temporal compensation.[12] Perhaps the reason why there are differences in the results of previous studies is inconsistency in dealing with various phonetic factors. Further study using controlled data is needed in order to examine the nature of Japanese temporal compensation.

Bradlow et al. (1995), on the other hand, introduced variations in supra-moraic prosodic features, speaking rate and sentence-level focus, in order to investigate the characteristics of mora timing in Japanese. It was found that the differences in contrastive and broad focus were not reflected in the differences in word duration, instead, it is reflected in the peak F0. This is additional evidence that Japanese is governed by mora timing. In English, the duration would also be affected, therefore, these two languages demonstrate different characteristics.

Studies of Japanese alone, however, may not be able to give an insight into the nature of temporal organization governed by mora timing. The recent cross-linguistic study by Nagano-Madsen (1992) provides interesting insights into the comparison of the temporal organization of mora-based languages such as Eskimo, Yoruba and Japanese. Although she does not draw any conclusions on the domains of temporal compensation for these languages, she demonstrates that the structures of temporal compensation in those mora-based languages are dissimilar, while the relationship between word duration and the number of morae within the word bears similarities.

For future studies, we need to be aware that investigation of the segmental duration (and the durational ratios) alone fails to capture the entire picture of the reality of mora timing. More studies need to be conducted using different approaches to investigate various aspects of mora timing.

---

[12] Campbell (1992) claims that the durational variance in vowels of Japanese is 'more influenced by their context than is that in consonants' (1992:208), while that of English is influenced by the following consonants.

In this dissertation, acoustic observation is conducted on speech production and discussion on the relationship between word duration and the number of morae is presented in Chapter 4.

2.2.3 Universal phonetics in timing organization

Recent literature on the concept of the mora hypothesis has seen an animated debate between linguists who support the view that Japanese timing organization is language-specific (i.e. learnt behaviour of NS) (Port, et al. 1980; Homma 1981; Han 1992, 1994) and those who take a universal perspective (Beckman 1982b, 1984; Otake 1989bd). The latter group argues that the phonetic characteristics which have been claimed as language-specific are in fact controlled by universal tendencies.

The extent to which Japanese timing structure is influenced by universal constraints is not yet clearly understood. One way to examine this issue is to investigate whether the same phenomenon is observed in the speech of NS of various other languages. If the same kind of temporal compensation was observed across different languages, then it is possible to say that it may be controlled by universal tendencies. Discussions pertaining to Japanese temporal compensation have already been presented in 2.2.2. In this section, cross-linguistic aspects of temporal compensation are discussed in order to examine the extent to which Japanese timing organization is language-specific.

Port et al. (1980), for example, investigated Arabic and Japanese and conducted two experiments on these languages in order to explore the extent to which vowels preceded by longer and shorter apical consonants (/t, d, r/) would compensate for duration (Fig.2.2). They found that in Arabic, there was no significant difference in medial stop closure duration between /d/ and /t/, and that the duration of the following vowel also showed no effect for voicing. They found, however, that the voicing of intervocalic consonants did affect the duration of the preceding vowel (which seems to be a phonetic universal). Thus their results show that in Arabic there is little evidence for timing compensation within a CV unit. Otake's results (1989b) on Arabic, on the other hand, provide counter-evidence to Port et al. (1980). He claims that Arabic speakers also show similar temporal compensation effects and that the characteristics of Japanese timing organization are thus not language-specific. Otake emphasises universal aspects and in this respect, supports Beckman (1982b).

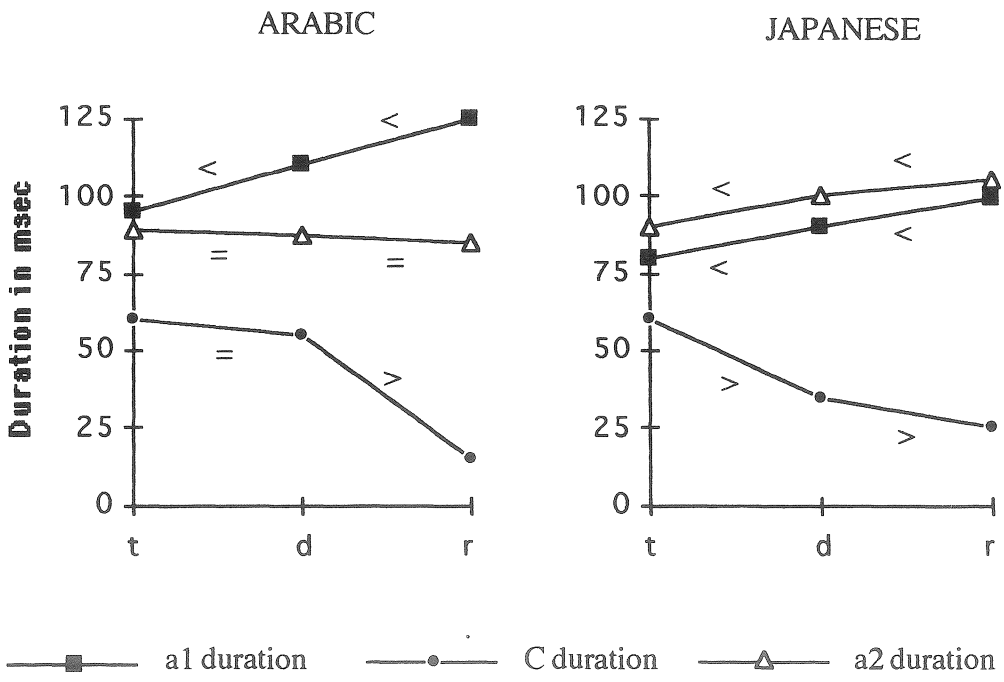


Fig.2.2 /aCa/ sequences of Arabic and Japanese (Port et al. 1980)

Fig.2.2 demonstrates that the preceding and following vowels compensate for the changes in consonant duration to about the same degree in Japanese. Vowels on both sides of a consonant varied inversely with the consonant, and consonant constriction duration varied inversely with the inherent vowel duration. Based on the different behaviour between Arabic and Japanese, Port et al. (1980) claim that the timing structure of Japanese is language-specific and thus learned behaviour on the part of speakers of Japanese. Port et al. (1980:249) give this as the reason why 'Chomsky and Halle's phonetic theory cannot take responsibility for an acoustically manifested multisegment structure like the Japanese mora.' The results obtained in their study suggest that timing structure in Japanese is not only a perceptual reality, constructed by listeners, but that it also exists as an acoustic signal, and therefore is also a physical reality.

Beckman (1982b) emphasises the effect of universal tendencies upon the formation of Japanese temporal compensation, and argues against Port et al. (1980) and Homma's (1981) claim that the domain of compensation spreads across mora boundaries, with vowel duration varying inversely with the duration of the consonant in VCV sequence. Beckman's (1982b:132) comments on this claim as follows:

> Finally, what PORT et al. [1980] and HOMMA [1981] call temporal compensation across mora boundaries may be nothing more than the apparently universal tendency for vowels to be shorter before voiceless stops than before voiced.

More recent work by Port et al. (1987) claims that the lengthening of vowels before voiced obstruents in fact operates as part of universal phonetics in Japanese timing organization, emphasising the importance of the linguistic unit of the mora in Japanese sound structure: 'The concept of the mora as an abstract isochronous unit of timing in Japanese captures many of the most salient features of timing in this language, despite the presence of universally observed "inherent duration" effects due to segment type.' They suggest further that universal models will not account for the entire picture of Japanese timing structure:

> But, altogether, attempts to model these results on this basis alone would miss some extremely fundamental regularities about the temporal system of this language. Instead, we should perhaps think of the implementation of mora timing as a system that is constructed to be compatible with these universal effects, one in which these universal tendencies collude with each other to achieve something more abstract and more clearly linguistic. This is not too different from arguing that the "universal voicing effect" is made grammatical and linguistic within English and other Germanic languages. (Port et al. 1987:1584)

Other aspects of Japanese timing structure which have been suggested as being universal are, for example, the relative duration of VOT amongst [p], [t] and [k], and the shorter duration of VOT for geminate stops than for single stops (Han 1992). Homma (1981), on the other hand, claims that gemination of stops does not affect VOT. The recent study by Nagano-Madsen (1992) also shows that VOT is not significantly different between single and geminate stops. In the present study, VOT is investigated to examine whether or not it is affected by the gemination of stops (see Chapter 4). It is beyond the scope of this thesis, however, to determine the extent to which temporal compensation in Japanese is governed by universal phonetics based on comparison with other languages.

2.2.4 Problems in the timing control of L2 learners

Han (1992) approached the issue of language-specific timing rules of Japanese by examining foreign accent:

> A study of a foreign accent may offer some insight into the controversy over timing-control rules in phonology. In general, a foreign accent is heard in phonological domains where the speaker's native language system conflicts with the target language structure. A foreign accent detected in the pronunciation of geminate stops involves the stop closure duration, or a silent period, as the primary cue. If different language speakers control the timing of stop closure differently, then language specific rules will be needed. If the temporal implementation of phonetic features is to be 'supplied by universal rules' [Chomsky and Halle, 1968, p.295] and thus the same for all languages, a foreign accent detected in the timing of segments cannot be explained by the theory. (1992:103)

Han used four male American English speakers who were fluent in Japanese and compared the results with those of ten Japanese NS. Her results showed that Japanese NS had a clear durational contrast between single and geminate stops

(2.8:1.0) in examples such as *ite kudasai* 'please stay' and *itte kudasai* 'please go', while American learners underdifferentiate the distinction (2.0:1.0).

She also observed the durational ratios of English [t] and [tt] occurring at the morpheme boundary such as 'catfish' and 'cattail' (although observation was not extended to other examples such as [k] and [kk]). English does not have the same durational contrast as Japanese at the phonemic level, as represented in minimal pairs such as *ite* and *itte*, but geminate stops appear across the morphological boundary in examples such as 'white tie' and 'rock cake'. Her subjects were asked to read out the following sentences and the results are shown below (1992:122):

(1) Look at my cattail.

(2) Look at my catfish.

(3) Look at my catnip.

(4) Please get Tom.

(5) Please get Mary.

Han's (1992) results show that in English, the duration of [tt] was twice the duration of [t] (Table 2.2). She comments, 'The results of this analysis seem to indicate that Americans generally pronounce [tt] in English by doubling the closure duration of [t].' This is similar to their corresponding durational ratios in Japanese as L2, and therefore Han (1992) claims that learners' underdifferentiation is due to negative transfer from their L1. From this evidence, Han (1992:126) concludes that 'The timing control of single and geminate stop closures was found to differ between Japanese and American English, and is thus language-specific.'

Words	Subject C [tt]	[t]	Subject P [tt]	[t]
cattail	135.0		150.7	
catfish		50.6		70.1
catnip		78.5		108.2
get Tom	162.2		131.0	
get Mary		63.9		69.0

Table 2.1 Examples of mean closure duration (in msec) of [tt] and [t] in American English (Han 1992)

Words	Subject C [tt]:[t]	Subject P [tt]:[t]
Ratio No.1	2.09:1.00	1.69:1.00
Ratio No.2	1.97:1.00	1.90:1.00
Mean ratio	2.03:1.00	1.80:1.00

Table 2.2 Examples of ratios of [tt] to [t] in English as spoken by two American subjects (Han 1992)

In order to investigate whether durational ratios in Australian English are similar to those in American English, I conducted an experiment following Han (1992). A velar stop [k] and fricative [s] were also included in the experiment since the experiments conducted in the following chapters deal with these phonemes are well as with an alveolar stop [t]. Three Australian subjects (also fluent in Japanese) were asked to read the following sentences three times each.[13]

(1) Go and get Mary.

(2) Go and get Tom.

(3) Go and get Tony.

(4) Go and kick Cathy.

(5) Go and kick Mary.

(6) Go and punish Cheryl.

(7) Go and punish Mary.

[13] In order to obtain compatible data with the speakers of American English, I used similar sentences to Han (1992).

The closure duration of the underlined consonants was measured. Table 2.3 and 2.4 shows the average figures (standard deviation in brackets) of the three speakers.

Words	Average (standard deviation)	
	[CC]	[C]
get Tom [tt]	114.90 (20.44)	
get Tony [tt]	108.84 (16.99)	
get Mary [t]		65.99 (9.50)
kick Cathy [kk]	124.72 (16.88)	
kick Mary [k]		73.96 (10.02)
punish Cheryl [ʃʃ]	156.97 (14.45)	
punish Mary [ʃ]		93.23 (15.50)

Table 2.3 Mean closure durations (in msec) of [tt] and [t] in Australian English

Words	Average ratios
	[CC]:[C]
[tt]:[t] Tom	1.74:1.00
[tt]:[t] Tony	1.65:1.00
[kk]:[k]	1.69:1.00
[ʃʃ]:[ʃ]	1.68:1.00

Table 2.4 Average ratios of [CC] to [C] in English as spoken by three Australian speakers

The ratio of [C]:[CC] obtained in my results is about 1.7:1, and it is interesting to note that this is a smaller ratio differentiation than for the American speakers obtained in Han (1992). The shorter closure duration of stops in both American and Australian English demonstrated above appears to present a problem (i.e. of L1 interference) in acquiring Japanese timing control. If learners' difficulty can be simply attributed to L1 interference, this result suggests that Australian learners will have greater difficulties in acquiring Japanese timing organization with respect to long consonants than American learners.

Han (1992:126) makes a pedagogical suggestion as follows:

> Instructors of Japanese as a foreign language bear a responsibility of calling the students' attention to the phonological contrast between geminate and single

consonants, and its phonetic timing control. A suggestion to 'make the geminate consonants 3 times as long as the single ones' may have practical, pedagogical merit.

Toda (1994) investigated the IL typology of timing organization in Japanese by beginning learners, and found that Han's (1992) pedagogical suggestion could not be applied uniformly to learners of different levels. It is because the causes of underdifferentiation can be attributed, not just to the shorter duration of geminate phonemes, but also to the following IL syllable modification strategies:

(1) longer duration of single C2, and

(2) longer duration of vowels preceding geminate C2.

Also, the timing control of CVCV in Japanese, which has not been discussed in great detail in the previous literature, was found to be a problem for learners (see Chapter 4). Based on these findings, it appears that Han's (1992) approach needs to be revised since her analyses are based only on negative language transfer. If learners' speech output was determined only by L1 interference, it might be possible to attribute the difficulties learners encounter to phonological differences between English and Japanese. It has been shown in Toda (1994), however, that learners develop their own strategies to distinguish phonological durational contrast, and that their IL is independent of both English and Japanese. Therefore, I would argue that a contrastive acoustic-phonetic analysis would not determine the extent to which Japanese timing control is governed by language-specific rules on the basis of errors involved in speech production by learners of Japanese. The absence of certain rules in the learners' speech only implies that certain principles do not operate in the formation of their IL. Thus it would appear that studies of foreign accent do not offer any answers to the controversy of timing-control rules in Japanese phonology. This issue is further investigated in Chapter 4 of this dissertation.

L1 interference has been considered the most prominent feature at the level of phonetics and phonology, and in fact a number of previous studies have investigated the typology of L1 interference in SLA. There is no doubt that it is very useful for language instructors to know the details of L1 interference so that they can account for errors which learners are likely to make. However, there is a risk of overlooking the full picture (including various processes and learner strategies which are involved in the formation of IL) if one tries to explain learners' errors only in terms of L1 interference.

There are almost no previous studies which investigate the various processes involved in the formation of English-Japanese IL phonology. By focussing on this aspect in the present study, it is hoped that this will contribute to the study of IL in SLA research.

2.3 Perception of the mora

Chapter 3 of this dissertation deals with the perception of mora. Before presenting data concerning the categorical perception of Japanese durational contrasts in the following chapter, it is useful to introduce the basic mechanisms of speech perception: by categorical and continuous perception. The perception of durational contrasts in Japanese vowels and stops by NS has been claimed as categorical (Ootsubo 1980, 1981).

2.3.1 Categorical and continuous perception

Categorical perception can be described as a phenomenon in which discontinuity is observed across categorical boundaries within a physical continuum which shows no obvious changes.

Repp (1984) defines categorical perception as follows:

> Phenomenally speaking, categorical perception refers to the experience of discontinuity as a continuously changing series of stimuli crosses a category boundary, together with the absence of clearly perceived changes within a category. It must be emphasized here that categorical perception is a very striking and readily demonstrated phenomenon. All persons who sit down and listen to

one of the standard series of stop consonants varying in VOT or formant transitions, provided they are able to hear the synthetic sounds as speech, will experience abrupt perceptual changes at certain places on the continuum. The continuing attraction of categorical perception to both the novice and the seasoned investigator lies in its permanent and replicable vividness in the listener's experience. (1984:251-252)

Categorical perception research was first conducted by Liberman and his collaborators at the Haskins Laboratories (Liberman et al. 1957; Bastian et al. 1961; Liberman et al. 1961b; Liberman et al. 1967). Liberman et al. (1957) created digitized sound stimuli, a physical continuum of /b/, /d/, /g/ (preceding a vowel approximating /e/), by increasing the onset frequency value of the second formant. They used speech synthesizer, Pattern Playback, and asked informants to identify the sounds given. It was found that the informants showed a sharp shift in their identification from one phoneme to another. In other words, the subjects did not capture the physical reality of the stimuli as a continuum of sound, but instead discriminated amongst the given sounds according to phonological categories.

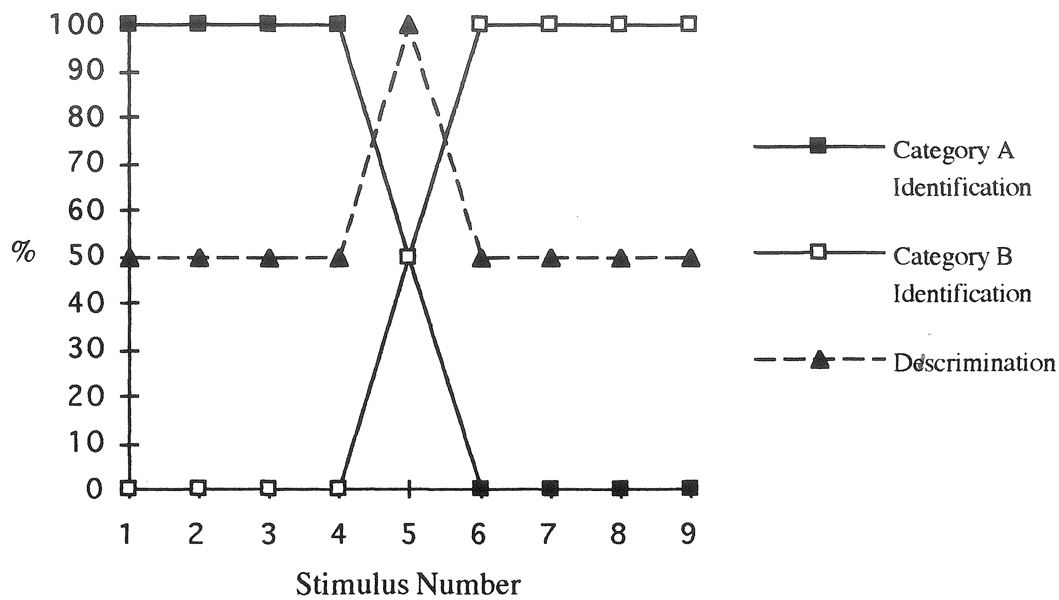


Fig.2.3 Idealized version of categorical curves (categorical perception of category A and B). See Lass et al. (1982:300)

This pioneering work encouraged a number of subsequent researchers to conduct experiments to find out more about the mechanism of categorical perception. The work by Liberman et al. (1961a) found that the distinction between intervocalic /p/ and /b/ was categorical, and was cued by closure duration. The voicing contrast between /t/ and /d/ (cued by 'first formant cutback') was also categorical (Liberman et al. 1961b).

Bastian et al. (1961) did experiments on word initial consonant clusters, such as those in /slit/-/split/ and found that the perception of the stop manner (cued by closure duration) was also categorical. The contrasts in liquid /r/ and /l/ (Miyawaki et al. 1975) and in stop /b/ and nasal /m/ (Miller and Eimas 1977) also demonstrated a similar tendency towards categorical perception.

The concept of categorical perception which was first introduced by the Haskins group was a strict model (i.e. an idealized model) which predicted a perfect shape for categorical curves as shown in Fig.2.3. In an empirical context, it is almost impossible to obtain such perfect results from most experiments and therefore the concept of 'degree of categorical perception' (Repp 1984) should also be included to account for the extent to which categorization was made.

The other type of perception is continuous perception; Fry et al. (1962) had shown that the vowel continuum /ɪ/ - /ɛ/ - /æ/ was not perceived categorically. Instead, these vowels were discriminated evenly within and between categories. Other examples of continuous perception include F0 and amplitude, acoustic correlates of pitch and loudness. However, later studies found that the results of categorical perception tests depended on the types of task and stimulus used in the experiment. For example, vowel stimuli could be perceived as categorically as consonant stimuli if the stimuli were presented with longer intervals than Fry et al. (1962). Thus, it is overly simple to claim that consonant perception is categorical and vowel perception is continuous. This is also relevant to the claim made in 1.2, that is, the current research focus is not just to find out whether or not L2 learners' perception is categorical or continuous.

Wode (1990, 1991, 1994) claims that both categorical and continuous perception are required mechanisms for speech perception. Categorical perception is necessary, because if speech processing was not conducted in some categorical manner, the human brain could not process all the information fast enough. On the other hand, continuous perception functions as 'a monitoring device that enables speakers to check whether their own production is still in tune with the norms of their speech community' (Wode 1994:146), and it is also important.

2.3.2 Perceptual development of L2 learners

The universal theory of language acquisition states that there is an infinite capacity for language learning by human beings, as stated in Wode (1994:145):

> ... the language learning capacity of humans is not restricted in any principled way as to the number of languages that can be learned, the types of languages and the kinds of linguistic structures that can be acquired, any age ranges during which this has to happen, or the learning situations as long as language input is available.

Following this view, an underlying assumption of the present study is that any L2 learners have ability to learn NS perceptual norms, although L2 learners have language-specific categorical boundaries which are influenced by their L1. Chapters 3 and 5 in this dissertation describe the perceptual development of Australian learners of Japanese. Prior to these chapters, previous studies pertaining to the L2 learners' perceptual categorization of Japanese durational contrast are described, and some methodological considerations are discussed in this section.

2.3.2.1 Problems

In accordance with the finding that the NS perception of durational contrast in Japanese vowels and stops was categorical (Ootsubo 1980, 1981), researchers have investigated whether the perception of Japanese durational contrast by learners who had different language backgrounds was also categorical (Min 1987; Hirata 1990b; Enomoto 1992; Uchida 1993a).

Previous studies have found that:

1) NS perception is more categorical than learners of Japanese, and

2) learners develop their perceptual system towards categorical as their proficiency level increases.

For example, Min (1987) investigated the perception of the durational contrast in Japanese stops by Korean learners of Japanese, and found that learners did not exhibit a clear categorical boundary as NS did. He claimed that the main perceptual cue for NS for judging the durational contrast was based on the physical duration of the silent closure period for voiceless stops. Korean learners seemed to rely on the spectral characteristics of voiceless stops rather than duration; however, Min did not provide any phonetic evidence to justify this assumption.

In the present study, I claim that just comparing the degree of categorization between different groups do not fully capture the differences in their perceptual mechanisms (Chapter 3). It is suggested that the observation be extended to other dimensions of perception such as actual threshold values and the interrelation between threshold values and the duration of other segments within the word. It seems inappropriate to associate learners' perceptual mechanisms with something which is similar to those of NS, unless learners' perceptual boundaries become similar to those of NS. It is not possible to say that learners' perceptual mechanisms have become similar to that of NS merely because learners' results demonstrate clearer categorical curves (also in Toda 1995ab).

Additionally, it is suggested that longitudinal data observation be conducted in order to investigate the nature of the mechanisms of L2 learners' speech perception. To the best of my knowledge, however, all previous studies concerning this issue employ cross-sectional data observation. There are no previous studies which investigate both speech perception and production based on longitudinal observations, which incorporate various aspects of categorical perception mentioned above.

2.3.2.2 Methodological considerations

Categorical perception research often involves the following two methods: the method of constant and the method of limit (or method of minimal changes). With regard to the previous studies pertaining to the acquisition of Japanese phonological categories, the former was used by Min (1987), Hirata (1990b) and Enomoto (1992) and the latter by Uchida (1993a).

The typical curves of categorical perception shown in the idealized model (Fig.2.3) may be obtained by using the method of constant (Repp 1984). In this method, stimuli are presented to informants in randomised order and the informants are asked to identify them. This is considered to be one of the most reliable statistical methods used in the area of psychophysics (Guilford 1959[1954]).

One disadvantage of this method is, as mentioned in Min's (1987) study, the threshold values are difficult to define within the categorical curves obtained by the method of constant. This is mainly due to the fact that it is almost impossible to obtain a perfect shape for categorical curves as shown in Fig.2.3. The other disadvantages which were pointed out in Guilford (1959[1954]:182) are: it is difficult to obtain the value of standard error, and also the data collection tends to be time consuming in comparison with other methods. In order to randomise the stimuli, it is suggested that each stimulus be presented to the subjects 50-200 times. Tanaka (1977) points out that subjects tend to lose concentration if the required time for experiments is lengthy, and therefore the obtained results may not be reliable. Researchers may be tempted not to present sound stimuli repeatedly, in order to save time for data collection. However, if the sound stimuli were not sufficiently repeated, the results could not be depended on to achieve statistical significance.

The other method which is also widely used in the area of categorical perception research is the method of limit (or method of minimal changes). In this method, the stimuli are presented in such a way that neighbouring stimuli differ only slightly so that subjects can indicate the just noticeable difference (JND) when they perceive a change. The main criticism of this method is that because stimuli are presented in successive order, the results tend to involve error of habituation and expectation (Guilford 1959[1954]:127). However, these can be minimised by using ascending and descending series of stimuli, which is applied in this study (see Chapter 3 for further details).

The method of limit was used in the present study since it was intended to investigate the ways in which NS and learners process the same acoustic information, not just the degree of categorization but also other aspects of perceptual mechanisms as well. Thus one of the observation objectives was the interrelation between the ascending and descending series (see Chapter 3). Also, the average values between the ascending and descending series are defined as threshold values, and by using these values, the investigation of the differences in the threshold values between NS and learners is conducted.

In order to produce additional randomising effects to minimise the disadvantages of this method (i.e. error of expectation), the following effects are added in this study when creating the sound stimuli: (1) the starting point of successive sound stimuli was randomised, and (2) the number of sound stimuli in each series was varied, as well as using ascending and descending series of sound stimuli.

2.4 Summary

This chapter reviewed previous studies. Firstly, works on Japanese prosodic characteristics concerning mora were reviewed from the viewpoint of phonological analyses (phonology) and phonetic reality (acoustic phonetics). Secondly, perception of mora was discussed. The previous studies on both production and perception of Japanese timing by L2 learners were discussed in this chapter.

Prosodic characteristics of languages are different, and there is no doubt that learners' linguistic experience of L1 influences the acquisition of L2 phonology concerning both speech perception and production. It was emphasised, however, that one cannot assume that the differences between L1 and L2 would account for all the errors made by L2 learners.

In the following chapter, the investigation is conducted on perception of mora and the analyses are presented on that of NS and learners of different proficiency levels.

Chapter 3 Perception

This chapter examines the perceptual judgment of the phonological durational contrasts in voiceless obstruents, vowels and nasals in Japanese. The experiment was first conducted with ten NS, and was followed by further tests with ten beginning and eight advanced learners of Japanese.

3.1 Experiments

This section presents the details of the sound stimuli and the procedures for the experiments.

3.1.1 Sound stimuli

The recording of the stimuli was conducted in a sound-proof recording room at the Department of Language and Culture, Nagoya University on 4 December 1992.[14] The word-tokens were pronounced by an adult female speaker in her early 30's, who had been engaged in Japanese language teaching. She was born in Tokyo and was teaching Japanese language while undertaking Ph.D research in Japanese linguistics at Nagoya University. She spoke the Tokyo variety of Japanese in daily life, and was married to a Japanese NS who was also from Tokyo.

A digital audio tape recorder Sony DTC-300ES was used for the recording. The sound was recorded onto the hard disk of an NEC PC-9801VX21 using a digital-to-analogue converter (with a sampling frequency of 10 kHz and an accuracy of 16 bits). The sound stimuli used in this study were created by PICOLA Plus.[15]

The following six words (natural word-tokens pronounced by the speaker) were used to produce the stimuli; *kate* 'food, provisions', *rika* 'science', *iso* 'beach', *honne* 'true feeling', *komma* 'comma' and *sanga* 'mountains and rivers' (Table 3.1).[16] The

---

[14] I would like to thank Mr Uchida of the Department of Educational Psychology, Nagoya University, for his assistance in creating the sound stimuli.

[15] See Uchida (1993b) for details of this system.

[16] The use of meaningless words was first considered in order to avoid semantic interference on the perceptual judgment, but later it was decided that words with meanings should be used in this study. This decision is supported by the results obtained in Uchida (1993) and Minagawa (1995), which

CVCVV stimuli (e.g. *katee*) were produced by lengthening the duration of the second vowel in the CVCV stimuli (e.g. *kate*), and the CVCCV stimuli (e.g. *katte*) by lengthening the duration of the intermediate consonant. In the case of the words with nasal consonants, the CVNV stimuli (e.g. *hone*) were created by shortening the duration of the intervocalic moraic nasals in the CVNNV series (e.g. *honne*). This is because the appropriate pitch pattern is not produced by the reverse operation (i.e. lengthening the duration of a non-moraic nasal in *hone* to produce *honne*).

1. kate 'source'	2. katte 'selfish'	3. katee 'process'
4. rika 'science'	5. rikka 'first day of summer'	6. rikaa 'liquor'
7. iso 'beach'	8. isso 'rather'	9. isoo 'removal'
10. hone 'bone'	11. honne 'true feeling'	
12. koma 'spinning top'	13. komma 'comma'	
14. saga 'nature'	15. sanga 'mountains and rivers'	

Table 3.1 Words in the response set

SERIES	WORDS	TONE
KATE	kate/katte/katee	LH
RIKA	rika/rikka/rikaa	HL
ISO	iso/isso/isoo	LH
KOMA	koma/komma	HL
HONE	hone/honne	LH
SAGA	saga/sanga	HL

Table 3.2 Tone pattern of each series

The words used in the experiments were carefully selected: (1-9) provide minimal pairs with phonological contrasts in the duration of vowels and consonants, and (10-15) provide minimal pairs with durational contrasts in intervocalic nasals (there are no minimal pairs involving CVNVV for words with nasals). All the minimal pairs have comparable pitch patterns (Table 3.2).[17]

showed no statistical differences in the learners' perception between meaningless and meaningful words in Japanese. Also, from the viewpoint of Japanese language learning, the use of real words seemed preferable.

[17] The acquisition of Japanese pitch patterns is beyond the scope of this study.

The duration of the segments below was manipulated in order to investigate the perceptual categorization of the phonological durational contrast (see Table 3.3 for original duration of manipulated segments):

(a) V1: The duration of the vowel in the first syllable (V1) was reduced and expanded - 2 stages were produced (70 and 140%) (See Table 3.4).

(b) V2: The duration of the vowel in the second syllable (V2) was expanded by 20% intervals - 16 stages were produced (60-360%) (Table 3.5).

(c) C2 (obstruents): The closure duration of the intervocalic obstruents was expanded by 20% intervals - 16 stages were produced (60-360%) (Table 3.5).

(d) C2 (nasals): The duration of the intervocalic nasal consonants was reduced by 10% intervals - 14 stages were produced (20-150%) (Table 3.5).

In (a), V1 duration was reduced and expanded in order to observe the influence of V1 duration upon the perceptual judgment of C2 and V2.[18] Watanabe and Hiratoo (1985) found that NS perceptual judgment of C2 was influenced by the duration of V1.[19] In the present study, the influence of V1 duration upon the perceptual judgment of C2 and V2 was investigated in the data of NS and learners at different proficiency levels, and the differences between these groups were compared.[20]

A total of 276 stimuli were created for this experiment, with 96 stimuli for long vowels (3 word-tokens x 2 stages of V1 duration x 16 stages of V2 vowel duration), 96 stimuli for long obstruents (3 word-tokens x 2 stages of V1 duration x 16 stages of C2 duration), and 84 stimuli for long nasals (3 word-tokens x 2 stages of V1 duration x 14 stages of nasal C2 duration) (see Table 3.5 for the absolute duration of C2 and V2).

The sound stimuli were presented in continuous sequences organized as ascending and descending series. In the ascending series, the sequence of stimuli was presented

---

[18] Sugito (quoted in Nakaoka 1990) claims that Japanese native speakers perceive long C2 when the duration of C2 is more than 1.41 - 1.69 times as long as V1.

[19] Fukui (1978) compared NS perception of reduced geminate C2 and expanded single C2. He found that the C2 duration was a primary cue and the V1 duration could be a secondary cue for the geminate perception of C2. Longer V1 before geminate C2 has been reported in the literature (Fukui 1978; Beckman 1992; Han 1992; Smith 1995).

[20] V1 duration can be related to speech speed in this context, although it is not a direct equivalent of speech tempo.

from shorter duration to longer duration. Thus, phonologically speaking, informants were expected to perceive two-mora expressions gradually changing to three-mora expressions. In the descending series, the stimuli were presented from longer duration to shorter duration, so that informants would hear three-mora expressions gradually changing to two-mora expressions.

The series of words were presented in randomised order and the starting and ending points of successive stimuli were changed. The number of stimuli in each series was also changed (ranging between 13 and 16) to cause additional randomising effects.

Thus the 276 stimuli were repeatedly presented to the subjects and the total number of stimuli used in this experiment (recorded on a 90 minute cassette tape) was 1044 (Appendix 3.1).

NO	C1V1C2V2	V1	C2	V2
1	kate	78.1	114.7	153.4
2	rika	58.3	125.7	131.5
3	iso	111.9	106.1	154.8
4	honne	75.3	181.8	N/A
5	komma	67.4	212.6	N/A
6	sanga	128.7	163.3	N/A

Table 3.3 Original duration of manipulated segments (msec)

NO	C1V1C2V2	V1	70%	140%
1	kate	78.10	54.67	109.34
2	rika	58.30	40.81	81.62
3	iso	111.90	78.33	156.66
4	honne	75.30	52.71	105.42
5	komma	67.40	47.18	94.36
6	sanga	128.7	90.09	180.2

Table 3.4 Absolute duration of V1 (msec)

NO	words		60%	80%	100%	120%	140%	160%	180%	200%	220%	240%	260%	280%	300%	320%	340%	360%
1	KAT_E	C2	68.82	91.76	114.70	137.64	160.58	183.52	206.46	229.40	252.34	275.28	298.22	321.16	344.10	367.04	389.98	412.92
	KATE_	V2	92.04	122.72	153.40	184.08	214.76	245.44	276.12	306.80	337.48	368.16	398.84	429.52	460.20	490.88	521.56	552.24
2	RIK_A	C2	75.42	100.56	125.70	150.84	175.98	201.12	226.26	251.40	276.54	301.68	326.82	351.96	377.10	402.24	427.38	452.52
	RIKA_	V2	78.90	105.20	131.50	157.80	184.10	210.40	236.70	263.00	289.30	315.60	341.90	368.20	394.50	420.80	447.10	473.40
3	IS_O	C2	63.66	84.88	106.10	127.32	148.54	169.76	190.98	212.20	233.42	254.64	275.86	297.08	318.30	339.52	360.74	381.96
	ISO_	V2	92.88	123.84	154.80	185.76	216.72	247.68	278.64	309.60	340.56	371.52	402.48	433.44	464.40	495.36	526.32	557.28

NO	words		20%	30%	40%	50%	60%	70%	80%	90%	100%	110%	120%	130%	140%	150%
4	HON_E	C2	36.36	54.54	72.72	90.90	109.08	127.26	145.44	163.62	181.8	199.98	218.16	236.34	254.52	272.7
5	KOM_A	C2	42.52	63.78	85.04	106.30	127.56	148.82	170.08	191.34	212.6	233.86	255.12	276.38	297.64	318.9
6	SAN_A	C2	32.66	48.99	65.32	81.65	97.98	114.31	130.64	146.97	163.3	179.63	195.96	212.29	228.62	244.95

Table 3.5 Absolute duration of C2 and V2

The method used for this experiment was the method of limit (as discussed in Chapter 2). The number of stimuli in each series and also their starting and ending points are also shown in Appendix 3.1.

On the answer sheets provided, word pairs were listed, indicating that the original stimuli were edited and gradually changed (e.g. *kate* -> *katte*, *iso* -> *isoo*). Words were written both in *kana* and *romaji*. Subjects were asked to mark with an X (i.e. the

X equals to JND) when they thought that the original stimuli changed as shown below (e.g. See the tenth stimulus in Table 3.6 and the seventh stimulus in Table 3.7):

NO	words	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	KATE -> KATTE	O	O	O	O	O	O	O	O	O	X	X	X	X	X	X	X

Table 3.6 Example of JND - ascending series

NO	words	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	KATTE -> KATE	O	O	O	O	O	O	X	X	X	X	X	X	X	X	X	X

Table 3.7 Example of JND - descending series

Subjects listened carefully and marked each stimulus before the JND with an O. Once they used an X, they could not revert to O after that point. Subjects were allowed to use a question mark if they were not sure, but it was suggested that they use as few question marks as possible because their frequent usage could make the results unreliable. Table 3.8 shows an example of JND with a question mark. In this case, the eleventh stimulus was interpreted as the point of JND and the tenth stimulus was marked on the perception graphs, indicating the point where the question mark was used (see an example of Fig.3.2).

NO	words	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	KATE -> KATTE	O	O	O	O	O	O	O	O	O	?	X	X	X	X	X	X

Table 3.8 Example of JND with a question mark

This study was designed to investigate the mechanisms of perceptual categorization, and how they differ between NS and learners of Japanese for the same given stimuli. Threshold values were obtained by averaging out all the samples of JND figures for each word observed in the experiment, and were used as the basis for the analyses of the perceptual mechanisms amongst different groups.

3.1.2 Procedure

The informants were asked to listen to the sound stimuli in a quiet room using a professional quality tape recorder. They indicated where a JND occurred (by marking with an X) on the answer sheet provided as shown in Table 3.6 and Table 3.7. Prior to this experiment, the informants were given some instructions and a practice session (kake/kakke/kakee) was conducted to familiarise subjects with the test procedures.

The experiment required a total of one and half hours, divided into two sections, side A (45 minutes) and side B (45 minutes) on the tape. It was usually conducted over two days, each session taking 45 minutes. Informants were asked to rest (about 15 minutes) when they started to lose concentration, or felt tired.

3.2 Japanese NS

This experiment examines the mechanisms of perception by Japanese NS.

3.2.1 Subjects

Ten Japanese NS with normal hearing were the subjects of this experiment. All Japanese subjects were from the Tokyo area and spoke standard Tokyo Japanese. There were four males and six females in this group, aged between twenty and thirty.

3.2.2 Results

Tables 3.9 presents the average values and the standard deviation for the ascending and descending series obtained from the ten subjects. Average threshold values are the average percentage increase for JND between the ascending and descending series.

Vowels

CV07CV_	KA07TE_	RI07KA_	I07SO_
A.S	164%	170%	156%
S.D	15%	17%	17%
D.S	155%	181%	150%
S.D	24%	21%	24%
CV14CV_	KA14TE_	RI14KA_	I14SO_
A.S	178%	188%	177%
S.D	14%	18%	16%
D.S	170%	183%	165%
S.D	26%	27%	16%

Obstruents

CV07C_V	KA07T_E	RI07K_A	I07S_O
A.S	179%	158%	206%
S.D	27%	26%	34%
D.S	165%	152%	200%
S.D	34%	32%	22%
CV14C_V	KA14T_E	RI14K_A	I14S_O
A.S	191%	170%	216%
S.D	28%	34%	41%
D.S	177%	168%	203%
S.D	33%	23%	25%

Nasals

CV07N_V	KO07M_A	HO07N_E	SA07N_A
A.S	62%	74%	76%
S.D	12%	15%	19%
D.S	57%	60%	66%
S.D	12%	16%	16%
CV14N_V	KO14M_A	HO14N_E	SA14N_A
A.S	72%	74%	92%
S.D	15%	16%	19%
D.S	58%	63%	76%
S.D	11%	17%	13%

Tables 3.9 Average Values - Japanese NS

A.S: ascending series

D.S: descending series

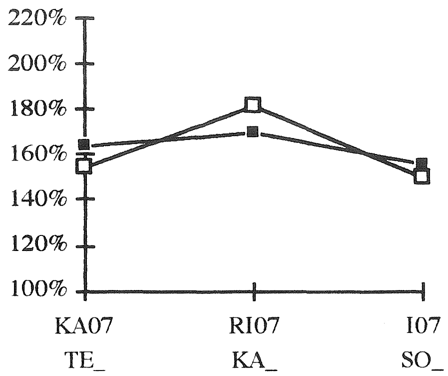
S.D: standard deviation

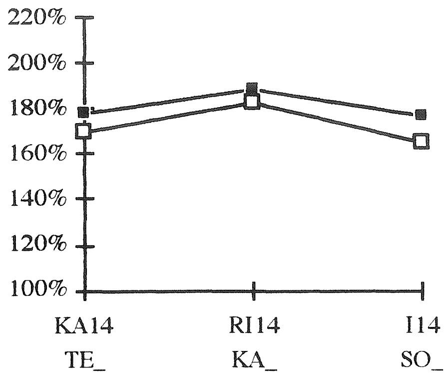
CV07CV_	KA07TE_	RI07KA_	I07SO_
T.V	160%	176%	153%
A.D	244.67	230.78	236.84
CV14CV_	KA14TE_	RI14KA_	I14SO_
T.V	174%	186%	171%
A.D	266.92	243.93	264.71
CV07C_V	KA07T_E	RI07K_A	I07S_O
T.V	172%	155%	203%
A.D	197.28	194.84	215.38
CV14C_V	KA14T_E	RI14K_A	I14S_O
T.V	184%	169%	210%
A.D	211.05	212.43	222.28
CV07N_V	KO07M_A	HO07N_E	SA07N_A
T.V	60%	67%	71%
A.D	126.50	121.35	115.13
CV14N_V	KO14M_A	HO14N_E	SA14N_A
T.V	65%	68%	84%
A.D	138.19	124.08	136.76

Table 3.10 Average Threshold Values - Japanese NS*

T.V: Average Threshold Values of ascending and descending series (%)

A.D: Absolute Duration (msec)





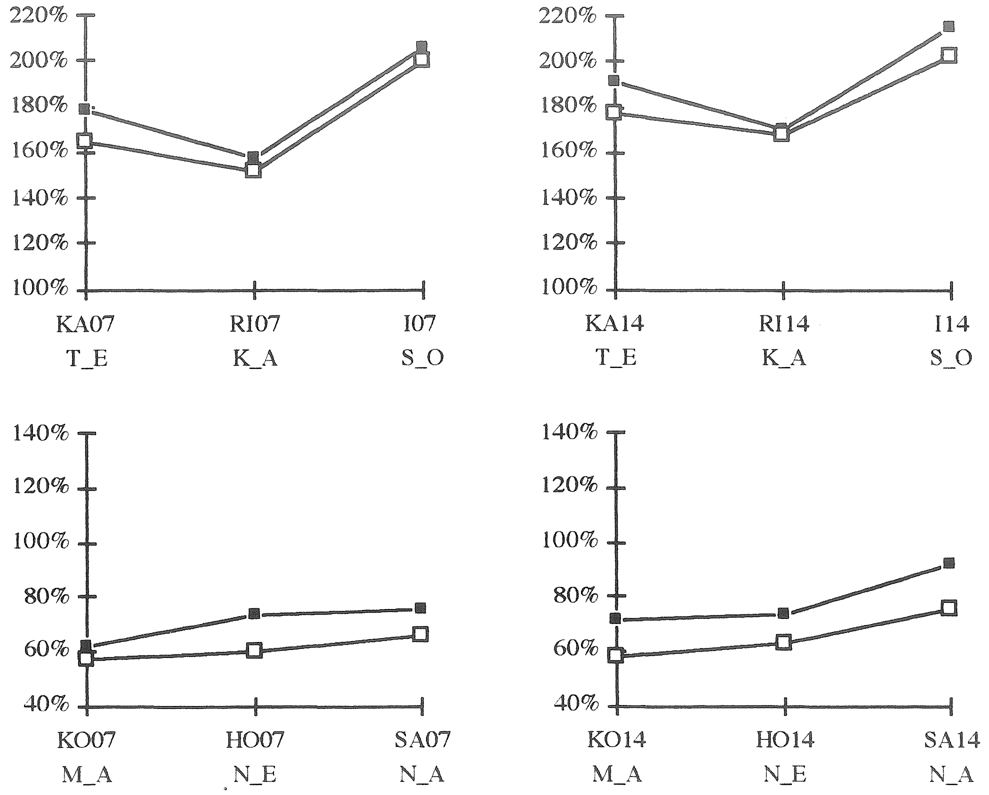


Fig.3.1 Graphs for perception by Japanese NS

Ascending (black), Descending (white)

Threshold values

It was found that average threshold values were always greater when V1 was expanded (140%) in comparison with when it was shortened (70%). This situation is observed in both C2 and V2. Statistical significance was observed in all V2. Interestingly, as far as C2 was concerned, it was observed only in KATE series (see *V1 duration*).

Note that the original duration is different for each stimulus: the original duration of V2 in RIKA_, for example, is the shortest (V2 in RIKA series is 131.5 msec, while it is 153.4 and 154.8 msec in KATE and ISO series respectively). The threshold values for RIKA_ are the greatest of those shown in the top two figures in Fig.3.1. This is largely due to the fact that the duration of the original stimuli is shorter than the other stimuli, and therefore requires a greater percentage increase for JND. On the other hand, the middle two figures in Fig.3.1 show a dipped curve with RIK_A at the lowest point, which is due to the longest duration of C2 in the original stimuli of RIK_A (C2 in RIKA is 125.7 msec, while it is 114.7 and 106.1 msec in KATE and ISO respectively).

Ascending and descending series

It was found that the average values for the ascending series were generally greater than those for the corresponding descending series. For example, the average values of KA07TE_ is 164% (ascending) and 155% (descending) (see Tables 3.9). This situation is also shown in the graphs (Fig.3.1). Since the stimuli are presented in ascending and descending order, this situation may be regarded as a 'cross over' effect of the ascending and descending series. Thus it can be said that, as far as this group of Japanese NS is concerned, the results do not seem to be distorted by errors of expectation arising from the presentation of stimuli in a successive order. If the results were so affected, the average values obtained from the descending series would be greater than those obtained from the ascending series: if subjects indicated JND earlier due to errors of expectation, the threshold values would be greater when the stimuli were presented in descending order, and they would be smaller when the stimuli were presented in ascending order.

In the present data, the order is reversed and thus it appears that the subjects indicated JND at the point of time when a given stimulus went just above (in the case of ascending series) or just below (in the case of descending series) the expected threshold values.

In order to check statistical significance for the differences between the ascending and descending series, t-tests were undertaken for all examples. It was found that in most examples, the difference was not statistically significant (Appendix 3.2). This means that NS perceptual judgment is consistent and involves little uncertainty because their perceptual boundaries are not affected by the way the stimuli are presented.

V1 duration

As mentioned previously, it was found that the average threshold values were usually affected by the duration of V1.

2-factor Anova was undertaken for the analysis of variance. The dependent variable was the percentage increase in duration required for a JND (i.e. the threshold value); the independent variables were V1 duration (70% and 140%). All of the statistical results for word-final V2 show that the threshold value is significantly greater when V1 duration is 140% than when it is 70% (Appendix 3.3). This means that NS perceptual judgment of V2 is influenced by the duration of V1. Amongst examples with obstruent C2, RIK_A showed the same tendency ($F=4.643$; $p=.0343$). KAT_E and IS_O, on the other hand, did not show any statistical significance ($F=3.071$; $p=.0837$ and $F=.854$; $p=.3583$ respectively). With regard to the examples with nasal C2, KOM_A and SAN_A demonstrated statistical significance ($F=3.971$; $p=.0499$ and $F=12.238$; $p=.0008$ respectively) but not HON_E ($F=.177$; $p=.6752$).

The above results may be related to the pitch pattern of word tokens. As far as the results obtained in this study are concerned, it can be said that the perceptual judgment of C2 is influenced by the duration of V1 in a word with an HL tone (i.e. when V1 is assigned an H tone), while it is not influenced by the V1 duration in a word with an LH tone (i.e. when V1 is assigned an L tone).

Although Watanabe and Hiratoo (1985) claim that the perceptual boundary of C2 is directly influenced by the duration of V1, their observation is based only on three sets of meaningless words ([apa/appa], [ata/atta] and [aka/akka]), all of which had an HL tone, and observations of words with an LH tone were not conducted. Consequently, they did not investigate the interrelation between the perceptual boundary and pitch pattern. The results obtained in the present study seem to suggest that NS perceptual boundary of C2 is influenced by the duration of V1 only when an H tone is assigned to the first CV. Further study is required of the details of interrelation between pitch patterns and NS perceptual judgment.

The results of this experiment suggest that a longer duration is required for NS to perceive a long V2, in accordance with the increase of V1 duration.[21] This situation seems to be the same in some C2. The subjects' perceptual judgment is, in other words, affected by the duration of V1. There are no 'fixed' threshold values between short and long phonemes for NS, but listeners' perceptual judgments are made in relation to the duration of other segment within the word.

If we consider V1 duration as a correlate of speech tempo, as mentioned previously, we can say that NS perceptual mechanism allows their perceptual judgment to be adjusted in accordance with the speech tempo.

Stability of judgment

The standard deviations generally range between 15-25% for V2, 20-40% for obstruent C2 and 10-20% for nasal C2.[22] The standard deviation is generally greater with C2 and thus it can be said that NS judgment is more stable with V2 than obstruent C2.

3.2.3 Summary

In this section, we have claimed that a longer duration was required for NS to perceive a long V2, in accordance with the increase of V1 duration. This situation is the same with regard to C2 in a word with an HL pitch pattern, but not with an LH pitch pattern. This is an interesting finding because Watanabe and Hiratoo (1985) did not point out the relationship between the influence of V1 duration upon the perceptual judgment of C2 and the pitch pattern of the word. As far as the results obtained in this study are concerned, it is not possible to say that NS perceptual boundary is uniformly influenced by the duration of V1.

---

[21] This does not mean, however, that the threshold values become two times greater when V1 is expanded from 70% to 140%. The percentage increase for threshold values is usually only about 10% more when V1 is expanded than when V1 is shortened.

[22] The smaller standard deviation for nasal C2 may be due to the fact that the stimuli intervals were 10% for nasal C2 while those were 20% for obstruent C2 and the word-final V2.

This section has also claimed that differences between the ascending and descending series were generally not statistically significant.

These findings seem to imply that (1) NS perceptual judgment of V2 and some C2 operates as a function of the V1 duration (2) NS perceptual judgment does not involve a large amount of uncertainty (thus it is categorical).

In summary, the NS results imply that there is no fixed duration which can be considered as appropriate for a long phoneme for NS, but that the perceptual judgment seems to be relative to the duration of other segments within the word. This is also an indication of the interrelation between speech tempo of the utterances and the listeners' perceptual judgment. Without reference to such a relative parameter, it is impossible to state the appropriate duration for a long phoneme purely in relation to the duration of its single counterpart.

The standard deviation is generally greater in geminate C2 than long V2 and thus it is claimed that NS judgment is more stable with long V2 than geminate C2.

3.3 Australian beginning learners

In this section, the learners' results are presented and compared with those of the Japanese NS described in the previous section.

3.3.1 Subjects

Ten Australian beginning learners with normal hearing were the subjects of the experiment.

The subjects were enrolled in the first year Japanese course at the Australian National University (ANU). All subjects had an Australian background with little previous knowledge of the Japanese language before they started their ANU course. All students spoke Australian English as their L1.[23] The recording was conducted two

[23] See Oasa (1980) for geographically correlated sound varieties in Australian English. He concludes that certain acoustic characteristics appear consistently among speakers of Australian English from different regions, but the general identification rate to detect speaker's origin is low. His conclusion is that 'the regional differences in Australian English are felt by many speakers, who express them as stereotypes, and the differences do exist, but they are not 'marked' enough for general identification' (1980:90).

months into the course in the first semester of 1993, and at that stage all the subjects were aware of phonological durational contrast in Japanese. The age group of the subjects was between eighteen and twenty except for BS who was in his mid-forties. Three male and seven female subjects participated in the experiments.

3.3.2 Results

This section presents beginners' results. They are compared with the NS results which were discussed above.

Vowels

CV07CV_	KA07TE_	RI07KA_	I07SO_
A.S	165%	178%	156%
S.D	32%	27%	29%
D.S	204%	209%	195%
S.D	33%	39%	27%
CV14CV_	KA14TE_	RI14KA_	I14SO_
A.S	174%	192%	163%
S.D	32%	29%	22%
D.S	208%	220%	204%
S.D	36%	28%	21%

Obstruents

CV07C_V	KA07T_E	RI07K_A	I07S_O
A.S	173%	163 %	172%
S.D	40%	39%	42%
D.S	212%	200%	205%
S.D	31%	30%	32%
CV14C_V	KA14T_E	RI14K_A	I14S_O
A.S	165%	169%	165%
S.D	40%	45%	37%
D.S	191%	202%	176%
S.D	31%	39%	33%

Nasals

CV07N_V	KO07M_A	HO07N_E	SA07N_A
A.S	68%	74%	74%
S.D	14%	13%	18%
D.S	88%	87%	88%
S.D	15%	13%	15%
CV14N_V	**KO14M_A**	**HO14N_E**	**SA14N_A**
A.S	72%	69%	73%
S.D	16%	16%	19%
D.S	80%	84%	71%
S.D	16%	14%	14%

Tables 3.11 Average Values - Beginners

A.S: ascending series

D.S: descending series

S.D: standard deviation

CV07CV_	KA07TE_	RI07KA_	I07SO_
T.V	185%	194%	176%
A.D	283.02	254.45	271.67
CV14CV_	**KA14TE_**	**RI14KA_**	**I14SO_**
T.V	191%	206%	184%
A.D	292.99	270.89	284.06
CV07C_V	**KA07T_E**	**RI07K_A**	**I07S_O**
T.V	193%	182%	189%
A.D	220.80	228.15	200.00
CV14C_V	**KA14T_E**	**RI14K_A**	**I14S_O**
T.V	178%	186%	171%
A.D	204.17	233.17	180.90
CV07N_V	**KO07M_A**	**HO07N_E**	**SA07N_A**
T.V	78%	80%	81%
A.D	165.30	145.89	131.86
CV14N_V	**KO14M_A**	**HO14N_E**	**SA14N_A**
T.V	76%	77%	72%
A.D	161.58	139.08	117.58

Table 3.12 Average Threshold Values - Beginners

T.V: Average Threshold Values of ascending and descending series (%)

A.D: Absolute Duration (msec)

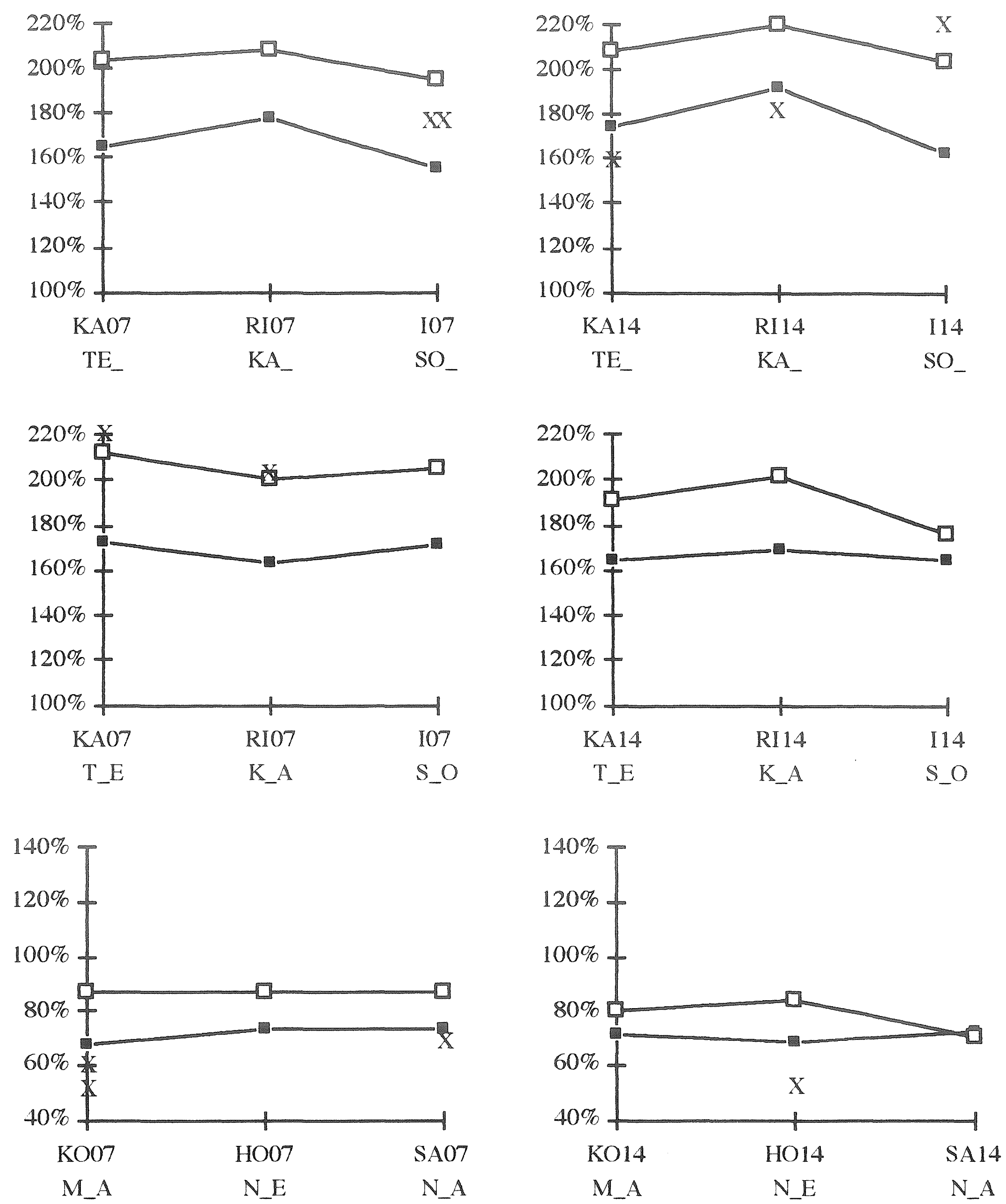


Fig.3.2 Graphs for perception by beginning learners[24]

Ascending (black), Descending (white)

Threshold Values

Firstly, t-test results found that the V2 threshold values obtained from this group of beginners were always greater than those obtained from NS (see Appendix 3.5). This means that the duration of V2 has to be longer to be perceived as long phonemes by

[24] X in the graphs shows where learners indicated a question mark before JND. Appendix 3.4 shows the values for which question marks were indicated.

the learners. For example, the threshold values in KA07TE_ are 159.5% (NS) and 184.5% (beginners) and the difference is significant (t=-3.68; p=.0004). In terms of duration, these figures are 245 msec (NS) and 283 msec (beginners). In KA14TE_ , the values are 174% (NS) and 191% (beginners) and the difference is also significant (t=-2.486; p=.015). These figures are 267 msec (NS) and 293 msec (beginners). In terms of duration, the perceptual boundary for V2 in the KATE series differs from 30 to 40 msec between beginners and NS subjects.

Within this perceptual domain, beginners would perceive the stimuli as short V2 while NS would perceive the same stimuli as long V2. Examples of differences in the perceptual boundaries (i.e. the differences between ascending and descending series are averaged) for CVCV_ between NS (normal line) and beginners (dotted line) are shown below:

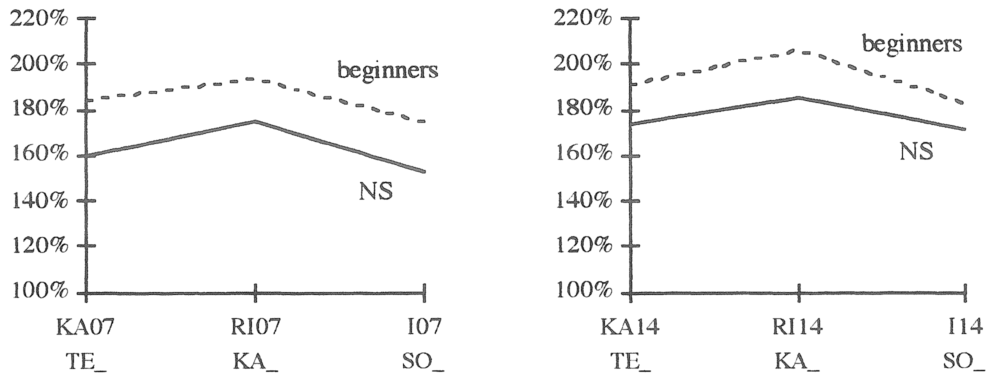


Fig.3.3 Differences in the perception of NS and beginners (V2)

The beginners show greater values in both graphs compared to NS with regard to long V2. If we define the perceptual boundary of NS (normal line) as a target, the distance between the NS and learners' values can be regarded as an inaccurate perceptual target on the part of beginners. This could also be the domain in which a perceptual discrepancy between NS and beginners may occur when they listen to the same stimuli.

The following graphs show the differences in C2 threshold values between NS and beginners.

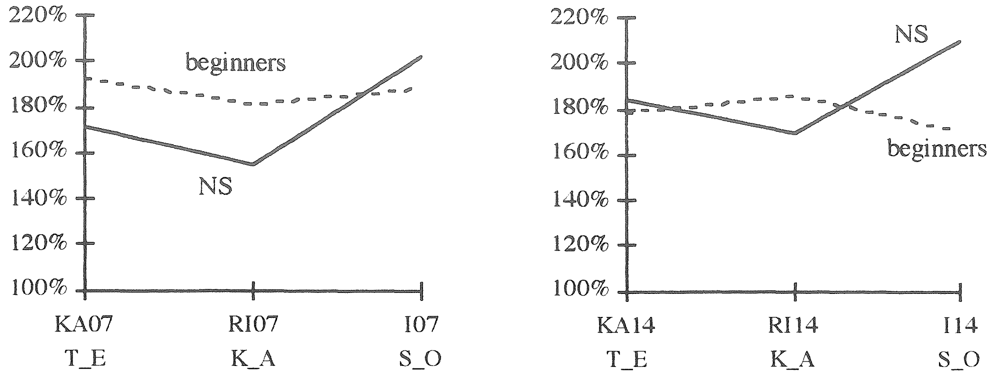


Fig.3.4 Differences in the perception of NS and beginners (C2)

It was found that the differences in the average threshold values of beginners in KA07T_E and RI07K_A are significantly greater than those of NS (t=-2.535; p=.0132 and t=-3.449; p=.0009 respectively). In I14S_O, on the other hand, NS had a greater threshold value than beginners. The threshold values in I14S_O are 209.5% (NS) and 170.5% (beginners) and the difference is significant (t=5.057; p=.0001). No significant difference was observed in I07S_O, KA14T_E and RI14K_A.

As far as KA07T_E and RI07K_A are concerned, it can be said that the learners have inaccurate perceptual targets which require a longer duration for the perception of long phonemes than it does for NS. This situation is similar to the perception of V2 described above.

On the other hand, it first appears as if the beginners' perceptual boundary is similar to that of NS in KA14T_E and RI14K_A. This situation is different from the V2 data, where the learners consistently required longer duration to perceive long phonemes. The present data, however, cannot simply be interpreted as demonstrating similarities in the actual perceptual mechanism between the two groups. The similarity seems to have been caused by the fact that beginners' average threshold values did not increase in accordance with V1 duration (see below *V1 duration*). Thus it is not possible to say that the nature of the learners' perceptual boundary is the same as that of NS.

Interestingly, the threshold values for fricative C2 are significantly smaller for beginners than for NS. This means that the duration of fricative C2 does not need to be as long to be perceived as a long phoneme by learners. For example, beginners may perceive an IS_O stimuli as containing a long fricative C2 while NS would regard the same stimuli as containing a short fricative C2. Smaller threshold values for beginners' data only appear in fricative C2. It is interesting to note that this is the reverse situation from the general tendency of stop C2 and V2.

The following graphs demonstrate the perceptual boundaries for nasal C2.

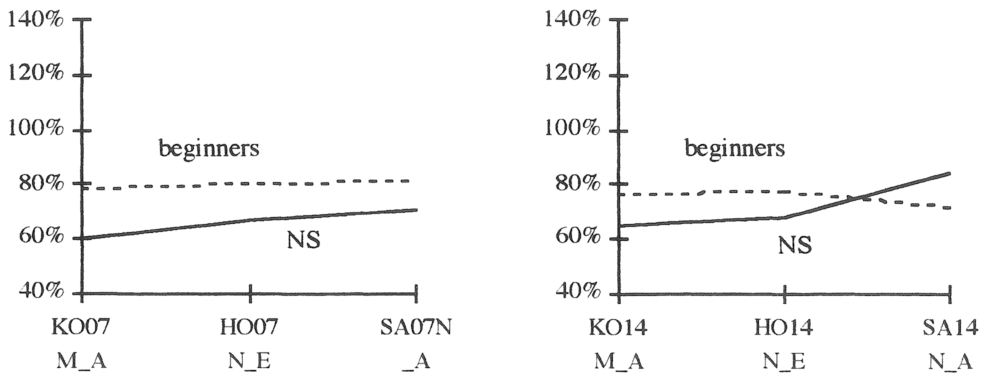


Fig.3.5 Differences in the perception of NS and beginners (nasal C2)

The differences in the threshold values between NS and beginners were statistically significant in all examples. The examples show greater values for beginners than for NS, except for SA14N_A (see Appendix 3.5 for statistics). Beginners' threshold values tended to be greater than that of NS, which is a similar tendency as observed in V2.

In Chapter 4, we investigate whether the incorrect perceptual targets found in this chapter account for any production errors by beginning learners.

Ascending and descending series

In the case of NS, it was generally the case that the differences between the ascending and descending series were not statistically significant, although the average threshold values for the ascending series were greater than the values for the corresponding descending series. Conversely, the results obtained for beginning learners of Japanese showed that absolute values for the descending series were almost always larger than those for the ascending series. The differences between the ascending and descending series were statistically significant in most cases (Appendix 3.6).

These results may be interpreted as evidence that the learners have different perceptual mechanisms from those of NS. In the case of beginners, the results may be affected by the error of expectation (Guilford 1959 [1954]) because unlike NS, they have not yet established certain threshold values to which they relate their perceptual boundaries.

The average distance between the ascending and descending series is obviously greater in learners' data than in the NS data. Thus it can be said that the degree of uncertainty (i.e. 'grey area') involved in learners' perceptual judgments is greater than for NS, and thus it is less categorical than NS perception.

V1 duration

In the NS data, it was claimed that perceptual judgment of V2 and some C2 was affected by the V1 duration, since the average threshold values were generally greater when the V1 duration was expanded (i.e. 140%) than when it was shortened (i.e. 70%).

In order to observe the interrelation between the duration of V1 and the threshold values for beginners, 2-factor Anova was undertaken in the same way as for the NS data shown above. It was found that no example showed significantly greater threshold values for longer V1 duration (Appendix 3.7). Instead, IS_O and SAN_A showed that the average threshold values was significantly greater when the duration of V1 was short than when it was long, which is opposite from the expectation. This tendency is not seen in NS results since NS data generally show an increase in threshold values accompanying expanded V1 duration. These results imply a difference in the nature of perceptual judgment between the two groups. If we can relate V1 duration to speech tempo, it could be said that beginners have not yet acquired the skill to adjust their perceptual judgment in accordance with the increase in speech tempo.

From these results, it can be said that the percentage increase in perceptual boundaries does not operate as a function of the duration of V1 with regard to beginners. Therefore, it is not possible to say that the perceptual judgment of beginners is systematically affected by the duration of V1 in the same way as that of NS.

Stability of judgment

The standard deviation of threshold values for long obstruents is generally greater than those shown in the NS data. This means that the learners' perceptual judgment is more unstable than that of NS. It is interesting to note that the standard deviation of the obstruent C2 is greater than that of the word-final V2 for both beginning learners and NS. It seems that the perceptual judgment of C2 involves greater instability regardless of whether the subject is NS or NNS.

3.3.3 Summary

Here the results of ANU beginning learners are summarised, and compared with those of NS. We have claimed that the NS and beginners have different perceptual mechanisms. In other words, learners had inaccurate perceptual targets when the NS norms were regarded as targets.

1. Threshold values were generally greater than NS for V2 and some C2, which means that the duration of these sounds needed to be longer for beginners to judge them as long phonemes. (I14S_O and SA14N_A were exceptions to this.)

2. The results obtained from beginners demonstrated that the average values of the descending series were almost always greater than those of the ascending series, and the differences were generally statistically significant. Therefore, it was claimed that the beginners' perception was not as categorical as that of NS.

3. It was found that beginners' perceptual judgment was not affected by the duration of V1 in the same way as NS.

In summary, the results obtained in this section seem to imply that different processes are in operation in the perceptual mechanisms of NS and beginners.

3.4 Australian advanced learners

The advanced learners' results are presented in this section, and compared with those of NS and beginners discussed above.

3.4.1 Subjects

Eight advanced speakers participated in this study; five from the Department of Foreign Affairs and Trade (DFAT) in Canberra (AA, BL, RA, RS and SF), one from the ANU (SB), and two from the Australian Embassy in Tokyo (CI and TN). None of the subjects suffered from hearing or speaking disorders. All subjects have an Australian background and speak Australian English as their L1. There were five male and three female subjects. The subjects from DFAT had higher than S3/R3 rating of the Japanese language proficiency level as assessed by the Language Training Unit, DFAT.[25] The speaker from the ANU (SB) had studied Japanese for eight years and was working for the Australia-Japan Research Centre at the ANU while continuing her studies in Japanese. This subject has a Japanese mother and was brought up in a Japanese-English speaking environment. The two subjects from the Australian Embassy were also fluent in Japanese. CI is engaged in English-Japanese translation and interpreting at the Translating section at the Australian Embassy, and also speaks Japanese at home with her Japanese husband. She studied Japanese at high school and university, then studied for one year at an interpreting school in Japan. TN had taken two year's intensive Japanese language training and had received a result higher than S3/R3 before he started working at the Embassy. It was towards the end of his three-year posting in Tokyo when TN participated in my project. All subjects have lived in Japan and have used Japanese at work.

---

25 S=Speaking; R=Reading/Writing; 5=native, 4=near native; 3=advanced; 2=intermediate; 1=beginner.

3.4.2 Results

The following are the results obtained from the advanced learners.

Vowels

CV07CV_	KA07TE_	RI07KA_	I07SO_
A.S	156%	175%	153%
S.D	21%	27%	31%
D.S	161%	178%	153%
S.D	31%	46%	40%
CV14CV_	KA14TE_	RI14KA_	I14SO_
A.S	173%	179%	168%
S.D	27%	22%	25%
D.S	169%	195%	159%
S.D	41%	42%	37%

Obstruents

CV07C_V	KA07T_E	RI07K_A	I07S_O
A.S	170%	143%	176%
S.D	27%	22%	28%
D.S	175%	143%	190%
S.D	40%	27%	31%
CV14C_V	KA14T_E	RI14K_A	I14S_O
A.S	198%	155%	186%
S.D	54%	26%	33%
D.S	184%	173%	159%
S.D	44%	40%	40%

Nasals

CV07N_V	KO07M_A	HO07N_E	SA07N_A
A.S	69%	81%	74%
S.D	11%	13%	13%
D.S	76%	78%	71%
S.D	13%	13%	16%
CV14N_V	KO14M_A	HO14N_E	SA14N_A
A.S	71%	84%	83%
S.D	12%	13%	14%
D.S	74%	80%	71%
S.D	14%	18%	17%

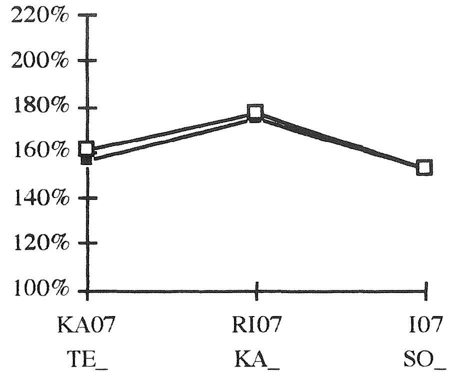
Tables 3.14 Average Values - Advanced learners

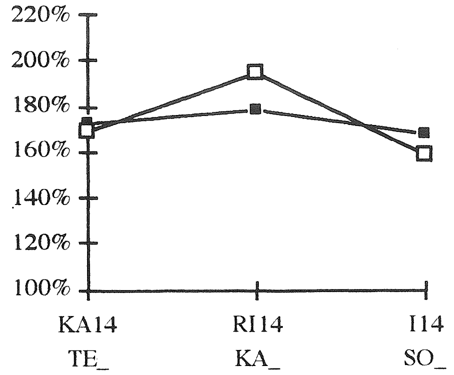
CV07CV_	KA07TE_	RI07KA_	I07SO_
T.V	159%	176%	153%
A.D	243.52	231.77	236.07
CV14CV_	KA14TE_	RI14KA_	I14SO_
T.V	171%	187%	163%
A.D	261.74	245.74	252.52
CV07C_V	KA07T_E	RI07K_A	I07S_O
T.V	173%	143%	183%
A.D	197.86	179.12	194.30
CV14C_V	KA14T_E	RI14K_A	I14S_O
T.V	191%	164%	173%
A.D	218.65	205.83	183.02
CV07N_V	KO07M_A	HO07N_E	SA07N_A
T.V	73%	79%	72%
A.D	154.14	144.30	117.88
CV14N_V	KO14M_A	HO14N_E	SA14N_A
T.V	73%	82%	77%
A.D	154.14	149.42	126.05

Table 3.15 Average Threshold Values - Advanced learners

T.V: Average Threshold Values of ascending and descending series (%)

A.D: Absolute Duration (msec)





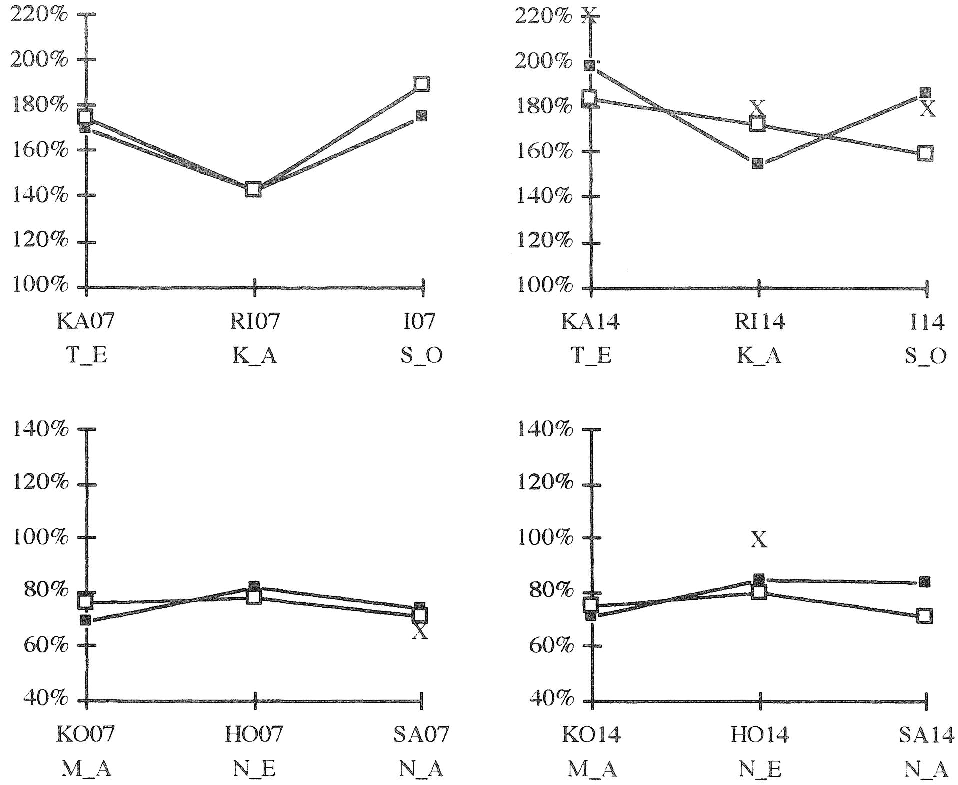


Fig.3.6 Graphs for perception by advanced learners[26]
Ascending (black), Descending (white)

Threshold values

The threshold values of advanced learners are generally smaller than beginners. For example, the threshold values for KA07TE_ and KA14TE_ are 283 msec and 293 msec for ANU beginners, while the corresponding values for advanced learners are 244 msec and 262 msec. In fact, these values are very similar to NS (245 msec and 267 msec respectively). All CVCV_ examples share this close similarity with NS data (Appendix 3.9). CVC_V examples also showed the same tendency in the KATE and RIKA series.

With regard to the ISO series, however, the threshold values were significantly smaller than the corresponding data obtained from NS (I07S_O ($t=2.869$; $p=.0054$) and I14S_O ($t=4.327$; $p=.0001$) respectively). This is an interesting finding because fricative C2 also showed smaller values in the beginners' data than the NS data. On

[26] Appendix 3.8 shows the values for which question marks were indicated by advanced learners.

the other hand, the values of KOMA and HONE series were found to be significantly greater than those of NS (i.e. KO07M_A (t=-4.543; p=.0001), KO14M_A (t=-2.303; p=.0243), HO07N_E (t=-3.558; p=.0007) and HO14N_E (t=-3.555; p=.0007)). This finding is also interesting because the same tendency was found in the beginning learners' data.

T-tests were conducted in order to observe if the differences in threshold values between beginning and advanced learners were significant. It was found that all the differences involved in the above examples were not statistically significant between the two groups (i.e. I07S_O (t=.628; p=.5321), I14S_O (t=-.23; p=.8188), KO07M_A (t=1.46; p=.1488), KO14M_A (t=1.009; p=.3162) and HO07N_E (t=.264; p=.7927) and HO14N_E (t=-1.485; p=.1421)). Thus from this respect, we can assume that perception of CVC_V by advanced learners shares more similarity with beginning learners than NS when C2 is fricative and nasal C2.

The above findings imply that the perception of advanced learners is intermediate between NS and beginners (see 3.5 Discussion). The results obtained in this experiment could also imply that the perceptual judgment of fricative and nasal C2 may be more influenced by language specific rules, and perhaps more challenging for advanced learners (as well as for beginners) to acquire than that of stop C2.

Ascending and descending series

With regard to the advanced learners' data observed in this study, the distance between the ascending and descending series is very small. The results of the t-test showed that the differences between the ascending and descending series were not significant except for I14S_O (t=2.126; p=.0418) and SA14N_A (t=2.122; p=.0422) (see Appendix 3.10). This was similar to the data obtained from NS, although the 'cross-over' effect described in NS data was not consistently observed.

This situation is different from the results for beginners, which showed that the percentage duration of the descending series tends to be greater than that of the ascending series.

The advanced learners had little 'grey area' in their perceptual judgment, unlike beginning learners, which implied that the advanced learners' perceptual judgment was similar to that of NS (i.e. more categorical than that of beginners).

V1 duration

The perceptual boundary does not seem to be affected by V1 duration with either C2 or V2 for advanced learners (Appendix 3.11). The only example which demonstrates a greater threshold value when the duration of V1 is expanded to 140% in comparison with when the V1 duration is shortened to 70%, is RIK_A (F=8.321; p=.0054).

It is interesting to note that statistical significance was obtained only in RIK_A concerning the influence of V1 duration upon NS perceptual boundary as well. This may be further supporting evidence for the assumption that the influence of V1 duration depends on the pitch pattern of the words (i.e. HL tone).

No other C2 nor V2 show a significant difference. Thus as far as the influence of V1 duration upon perceptual boundary is concerned, advanced learners' data share more similarity with those for beginners than those for NS.

Stability of judgment

The standard deviation of threshold values is usually greater than that of Japanese NS.

3.4.3 Summary

1. It was found that the results obtained from advanced learners were similar to those of NS with regard to CVCV_ and CVC_V when C2 was a stop C2. The threshold values of these examples are smaller than those of beginning learners.

2. When C2 was a fricative or nasal C2, on the other hand, advanced learners' results shared more similarity with those of beginning learners than those of NS. The threshold values were smaller in fricative C2 and greater in nasal C2.

3. The differences between the ascending and descending series were generally not significant. This means that the advanced learners' perception is more categorical than that of beginners, and thus similar to the results obtained from NS. This finding implies the developmental nature of learners' perceptual categorization.

4. V1 duration did not affect their perceptual boundary. This was similar to beginners and different from NS.

3.5 Discussion

The above results are summarised in the following table. ANU beginning learners and advanced speakers show the following tendencies (O= similar to NS; X= dissimilar to NS) when the NS norm is regarded as the perceptual target:

	BEGINNER	ADVANCED
ascending / descending	X	O
V1 duration	X	X
threshold values	X	O (stop C2 and V2) X (fricative and nasal C2)

Table 3.16 Similarities shared with NS results (perception)

The advanced learners' data shown above demonstrate a good example of perceptual approximation (instead of phonetic approximation) where the characteristics of their speech perception are at an intermediate stage between NS and beginners. In previous studies, examples of phonetic approximation have been presented concerning learners' speech production (e.g. VOT in advanced learners speech in Toda (1994)). To the best of my knowledge, however, no other previous studies have investigated the details of Japanese learners' speech perception as an intermediate stage between NS and beginners. The present study found that L2 approximation is not only present at the level of speech production, but also at the level of speech perception.

The following is the finding of this chapter:

1) In terms of perceptual boundaries both beginners and advanced learners demonstrated some differences from NS. In other words, all L2 learners (even near-native advanced learners) had inaccurate perceptual targets at the level of speech perception.

2) Advanced learners' data, however, shared more similarities with NS data than did beginners' data. This finding implies the developmental nature of learners' perceptual judgment. The advanced learners' data shows an example of L2 approximation where the characteristics of their perception is in between those of beginners and NS.

3) Neither beginners nor advanced learners were able to perform like NS with respect to V1 duration. This phenomenon may be rephrased as learners not being able to adjust their perceptual judgment in accordance with speech tempo. This may be a limitation pertaining to the acquisition of NS perceptual norms by learners of Japanese.

4) The characteristics of the learners' perceptual judgment seem to account for some of the characteristic features of their pronunciation reported in Toda (1994).

The present study found that longer duration was generally required for the perception of V2 and C2 (except for a fricative C2) by beginners, and this seems to correspond with the finding that longer duration of V2 and C2 is involved in their speech production (Toda 1994). Also, Toda (1994) reported that the durational ratios of C2/V1 and V2/V1 were inconsistent in beginners' speech production. This seems to be related to the fact that the learners' perceptual judgment does not operate as a function of V1 duration. In other words, although learners have the concept of long/short contrast, they have not acquired the concept of the segmental duration in relation to the preceding vowel as an important part of NS timing organization. The interrelation between their speech perception and production seems to imply that learners' inaccurate perceptual categorization affects their speech production (see Chapter 4 for further discussion).

One area worthy of further investigation for future study is the interrelation between learners' perceptual mechanisms and different IL strategies they apply in their speech production. Another area of interest is the further investigation of intermediate-advanced level learners in order to examine when learners' perceptual boundaries become similar to the NS norm. This chapter found that the advanced group shared similar perceptual boundaries with NS, while beginners did not show any similarity with NS in this regard. (Additionally, Chapter 6 demonstrates how beginners' perceptual judgments gradually become more accurate in relation to perceptual targets over one year.) Thus the learners' perceptual target had been modified somewhere between intermediate level to near-native advanced level in the process of language acquisition. Further study needs to be conducted on the acquisition of NS perceptual boundaries by intermediate-advanced level learners.

3.6 Conclusion

This section has examined the development of inaccurate perceptual categories by learners of Japanese. Three different groups, namely, NS, beginning and near-native advanced speakers were observed in order to examine the differences involved in their perceptual boundaries. The results obtained from the three groups of informants were summarised in 3.5, and the similarities and differences were discussed.

The acquisition of the perceptual norms of NS is a challenging task for learners which requires as much time and practice as (if not more than) that of speech production.

Chapter 4 Production

In this chapter, aspects of timing organization are discussed from the viewpoint of speech production.

4.1 Experiments

4.1.1 Test words

The subjects were asked to read fifteen test words which included durational contrasts in intervocalic consonants and word-final vowels (see Table 4.1).

1. kate 'source'	2. katte 'selfish'	3. katee 'process'
4. rika 'science'	5. rikka 'first day of summer'	6. rikaa 'liquor'
7. iso 'beach'	8. isso 'rather'	9. isoo 'removal'
10. hone 'bone'	11. honne 'true feeling'	
12. koma 'spinning top'	13. komma 'comma'	
14. saga 'nature'	15. sanga 'mountains and rivers'	

Table 4.1 Test words

The words were presented in Japanese orthography; pitch patterns and romanized equivalents of the Japanese words were also presented for the benefit of the non-native subjects.

4.1.2 Procedure

The recordings were made in a quiet room using a professional quality recorder and microphone. Each token was repeated three times at normal speed. Auditory impressions of the recorded data were transcribed before acoustic observations were made by the author. The recorded data were then analysed using Signalyze 2.08 (Keller 1992) (see 4.1.3).

The average values and ratios of the following items were then obtained using Microsoft Excel:

1) The duration of the first consonant (C1),[27] the first vowel (V1), the intervocalic consonants (C2), VOT[28] and the word-final vowel (V2).

2) Total word duration,

3) Long/short contrast (the ratio of long to short C2 or V2, normalizing long C2 or V2 as 100).

4) C2/V1 (the ratio of the duration of the intervocalic consonant to that of the preceding vowel).

The duration of VOT was measured and presented as mentioned above, but whether VOT should be included in C2 or V2 is a controversial issue (Weismer 1979; Beckman 1982b). In this paper, ratios for short/long phonemes are presented without including VOT in either C2 or V2. The observation of VOT is conducted separately because of the concern that if VOT is included in the preceding or following segment at this stage, the observations for the durational contrast may be affected by the different nature of VOT between NS and learners.

4.1.3 Acoustic measurement

This section reports the details of acoustic measurement and of spectrum segmentation.

The recorded data were analysed using Signalyze 2.08. The speech signal was fed into a Macintosh computer through MacRecorder with a sampling frequency of 22 Hz. This device allows the time scale (X axis) of the spectrogram to be expanded, which is useful for detailed observation.

Wide band spectrograms (8 ms/125 Hz) were obtained together with waveforms for segmentation. The range of the spectrogram can be set at one of three levels (full-range 0-11128 Hz; half-range 0-5564 Hz; quarter-range 0-2782 Hz), which also proved useful. The setting was changed depending on what observations were being

[27] The word [iso] has a word-initial vowel and thus does not have C1. Also it was not possible to measure the duration of the closure period for the word-initial velar stop [k], and therefore the measurement of the first consonant (C1) only refers to VOT. In intervocalic position (C2), duration refers to the hold and VOT.

[28] The observation of VOT is given only for voiceless stops.

sought: for example, it was set at half-range in order to observe the details of fundamental frequencies (F0), first formant (F1) and second formant (F2) patterns of vowels, and at full-range for the observation of fricative consonants since they were associated with high frequency noise.

Narrow band spectrograms (25 ms/40 Hz) also provided additional information about pitch patterns (see below). The criteria for segmentation are given in the following section.

4.1.3.1 C1

Stops

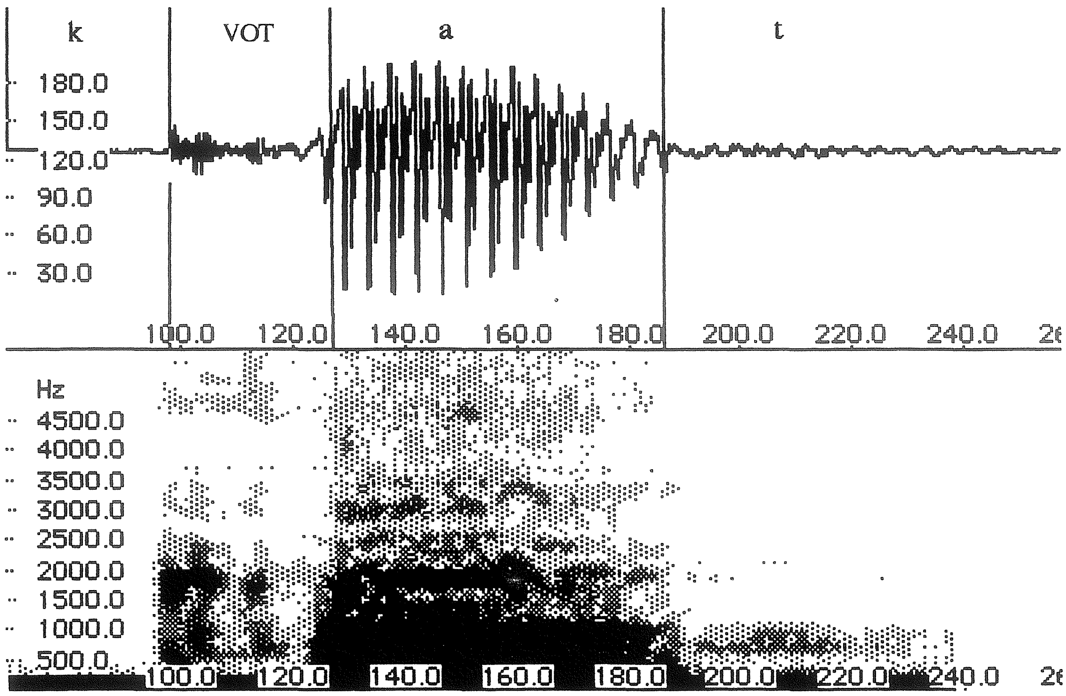


Fig.4.1 [k] in [katɛ] (example from a Japanese NS).

The closure period of voiceless stops cannot be measured in a word-initial position. Therefore the acoustic measurement of [k] in this study refers to the VOT which is measured from the point of release of oral closure until just before the onset of phonation of the following vowel which is indicated by clear periodicity.

Vertical lines on speech waves show the starting and ending point of the measurement for each segment.

Fricatives

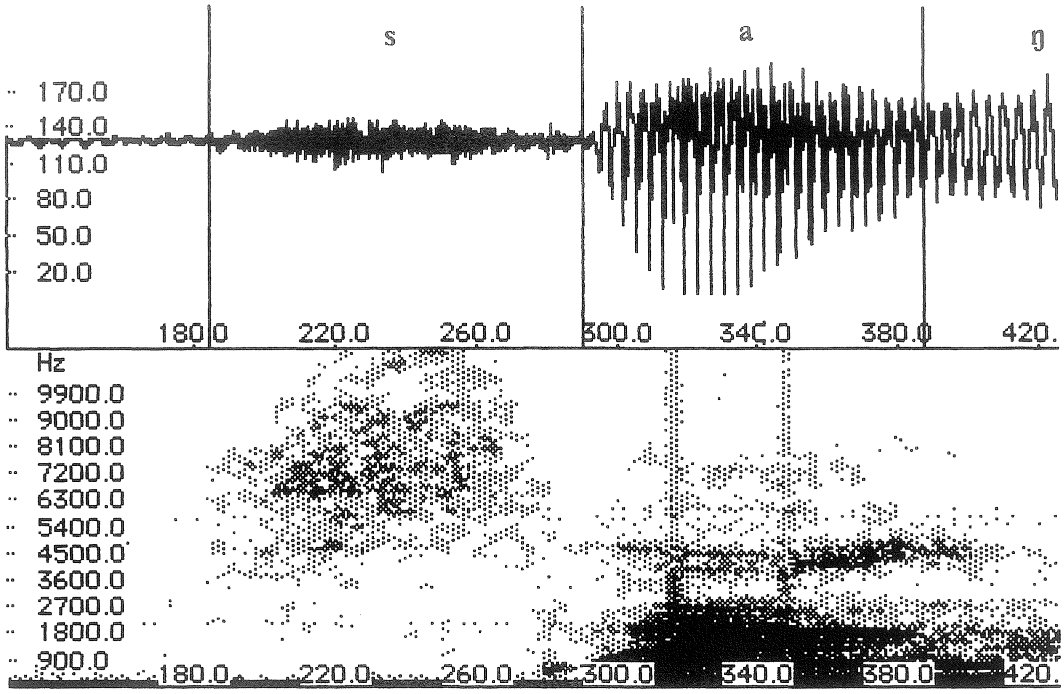


Fig.4.2 [s] in [saŋga] (example from a Japanese NS).

The onset of word-initial fricatives such as [s] is measured as the point where high frequency energy is present (around 5000-6000 Hz and above) before the onset of phonation for the following vowel. Therefore, wide band spectrograms are set at full-range for the analysis of fricatives as illustrated in Fig.4.2.

Liquid

Word-initial /r/[29] is measured from where discontinued formant frequencies, which indicate a flap, are observed (Fig.4.3).

[29] Japanese /r/ is regarded as a flap [ɾ] and does not occur in word-initial position in etymologically native words (Shibatani 1990:163). The allophones of this phoneme are as follows: a flap [ɾ], an approximant [ɹ], an alveolar lateral [l] and a retroflex plosive [ɖ] (International Phonetic Association 1984[1949]:44).

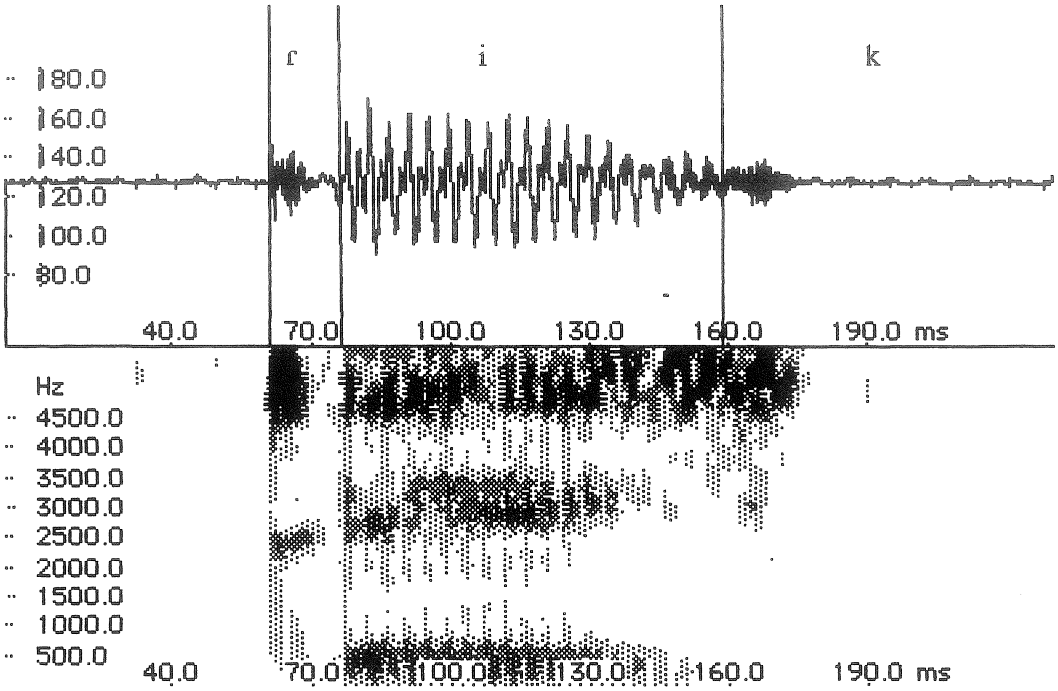


Fig.4.3 [ɾ] in [ɾika] (example from a Japanese NS).

Some learners produced a multiple-rolled sound [ʀ][30] instead of a single flap [ɾ] as shown below.

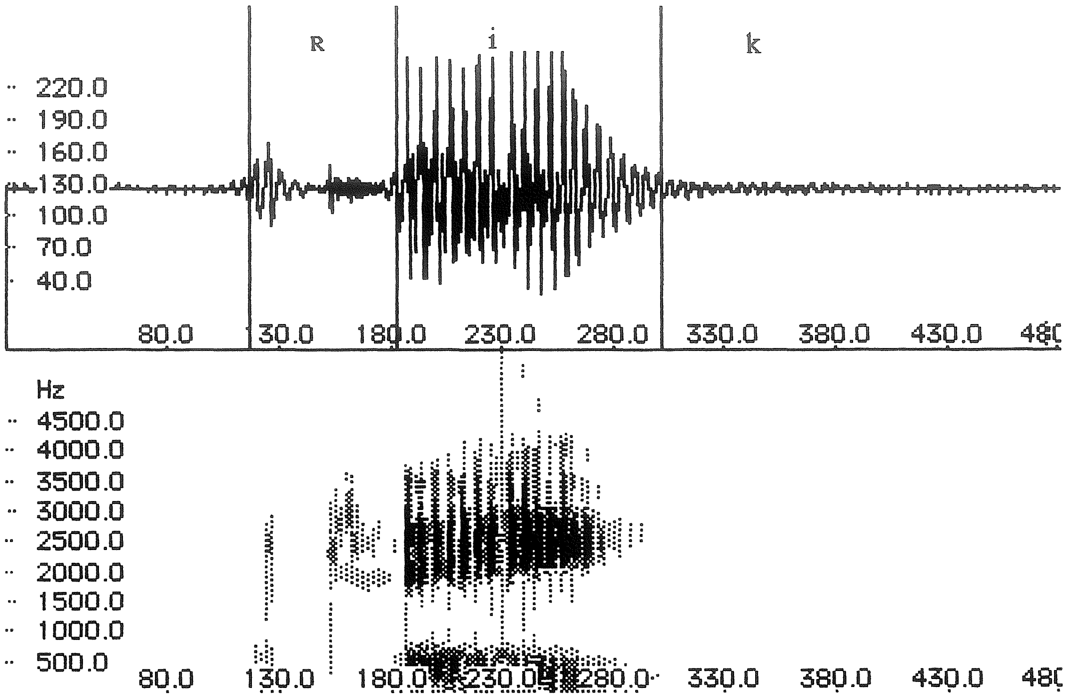


Fig.4.4 [ʀ] in [ʀika] (example from a beginner)

---

[30] In the IPA system, [ʀ] is used to transcribe a uvular flap/trill. In this dissertation, [ʀ] is used to distinguish a fully rolled alveolar trill from a single flap [ɾ].

The duration of a fully rolled trill is usually longer than a single flap. More than two discontinued shadings are observed in the spectrogram (Fig.4.4).

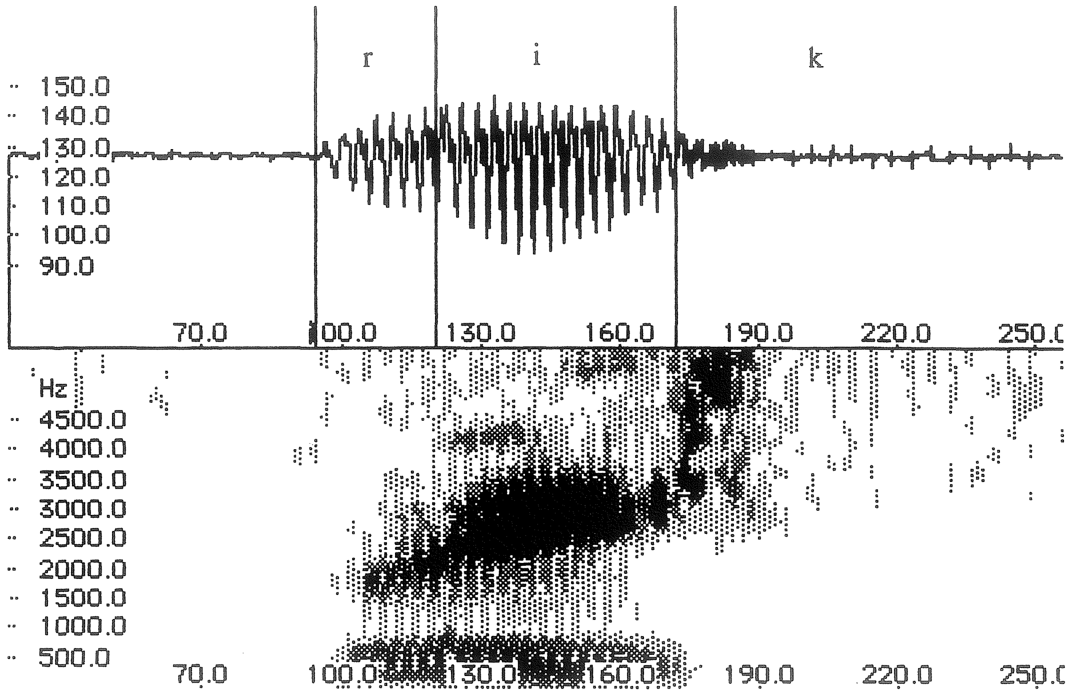


Fig.4.5 [r] in [rika] (example from a beginner)

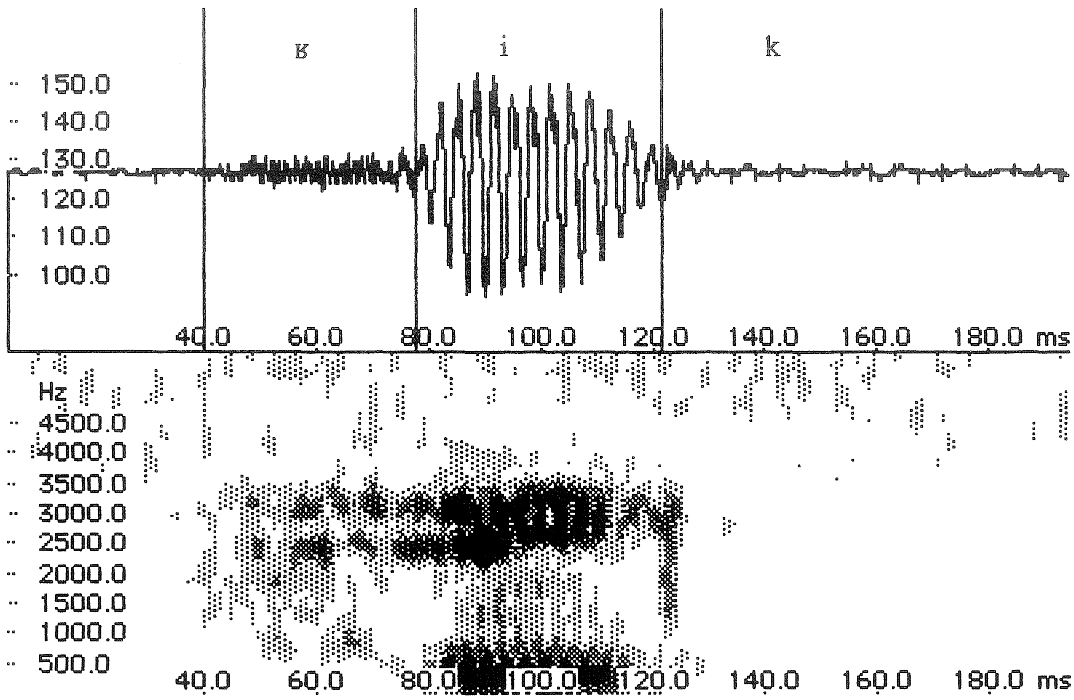


Fig.4.6 [ʁ] in [ʁika] (example from a beginner)

English-like [r], which may be due to L1 interference, was also observed in the learners' speech (Fig.4.5). The second formant frequency (F2) is low at the initial position, which is followed by a rise of F2 in the high vowel [i].

One informant pronounced a fricative [ʁ], which may be due to interference from her knowledge of French (Fig.4.6). Thus there is a four-way contrast in learners' /r/: English-like [r], a flap [ɾ], a trill [ʀ] and a fricative [ʁ].

4.1.3.2 V1

Word-internal position

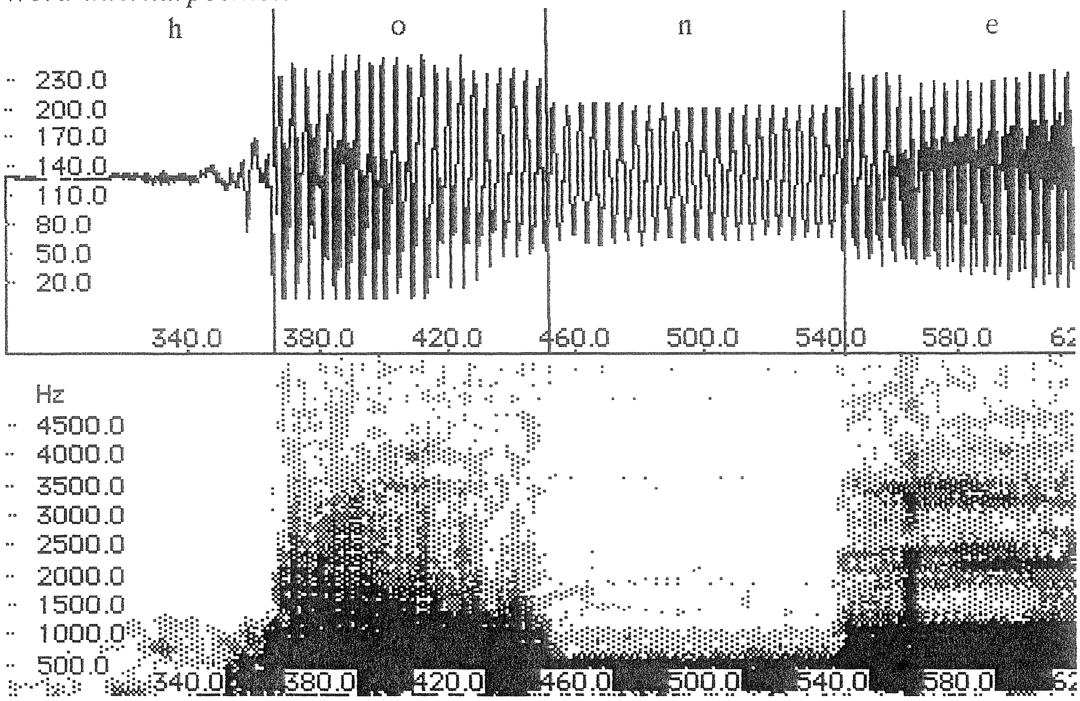


Fig.4.7 [o] in [hone] (example from a Japanese NS).

As in word-initial position, the onset point of a word-internal vowel is usually clear from both waveform and spectrogram (Fig.4.7). The offset point of the word-internal vowel is also usually obvious in the waveform when the following consonant is [t] or [k] because of the silent closure period of the voiceless stop. When C2 is a nasal consonant, however, the offset point of the preceding vowel is less clear, since the nasal is a sonorant and thus the waveform is continuous. In this case, additional information from the spectrogram is useful since the vowel offset point can be identified with reference to the ending of the F1 (first formant) and F2 (second formant) of the vowel.

4.1.3.3 Single C2

Stops

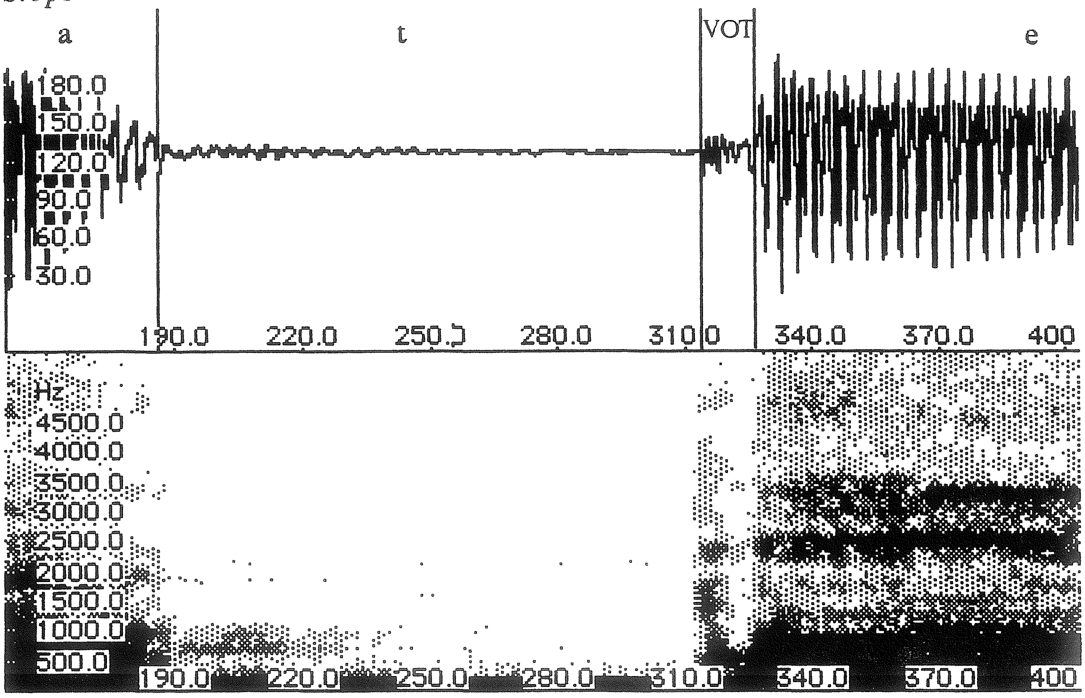


Fig.4.8 [t] in [kate] (example from a Japanese NS).

Voiceless stops in word-internal position are relatively easy to handle in terms of segmentation (Fig4.8). The closure duration and VOT are measured independently as shown above.

Fricatives

The onset of word-internal fricatives is measured using a full-range wide band spectrogram (Fig.4.9). The criteria for segmentation are similar to those for fricative consonants in word-initial position (see Fig.4.2).

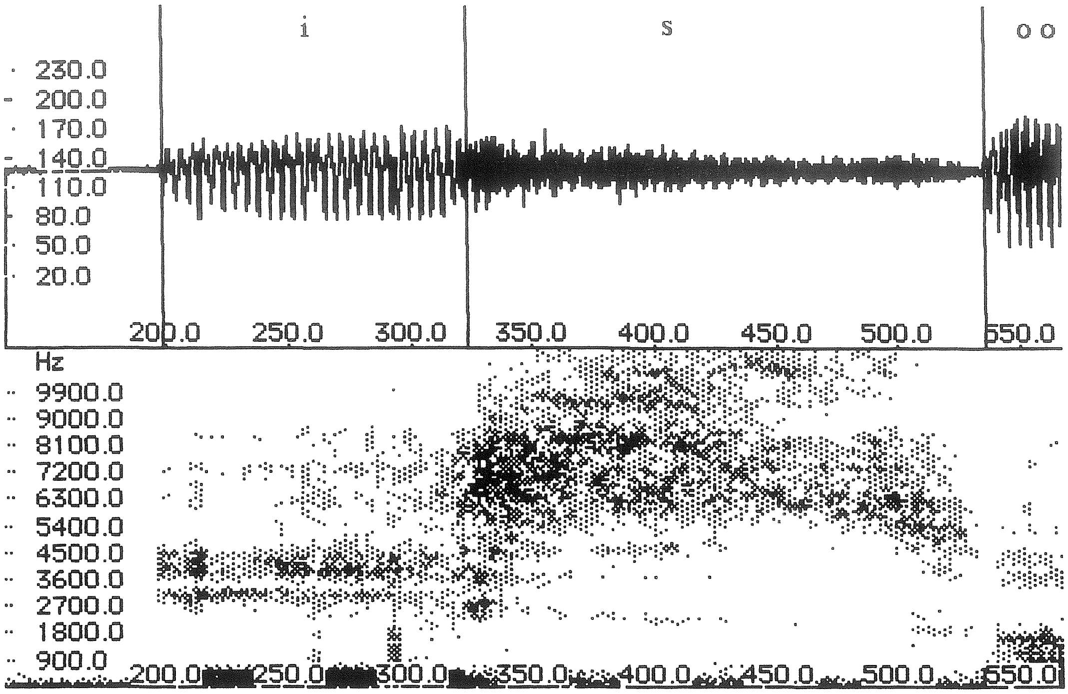


Fig.4.9 [s] in [isoo] (example from a Japanese NS).

Nasals

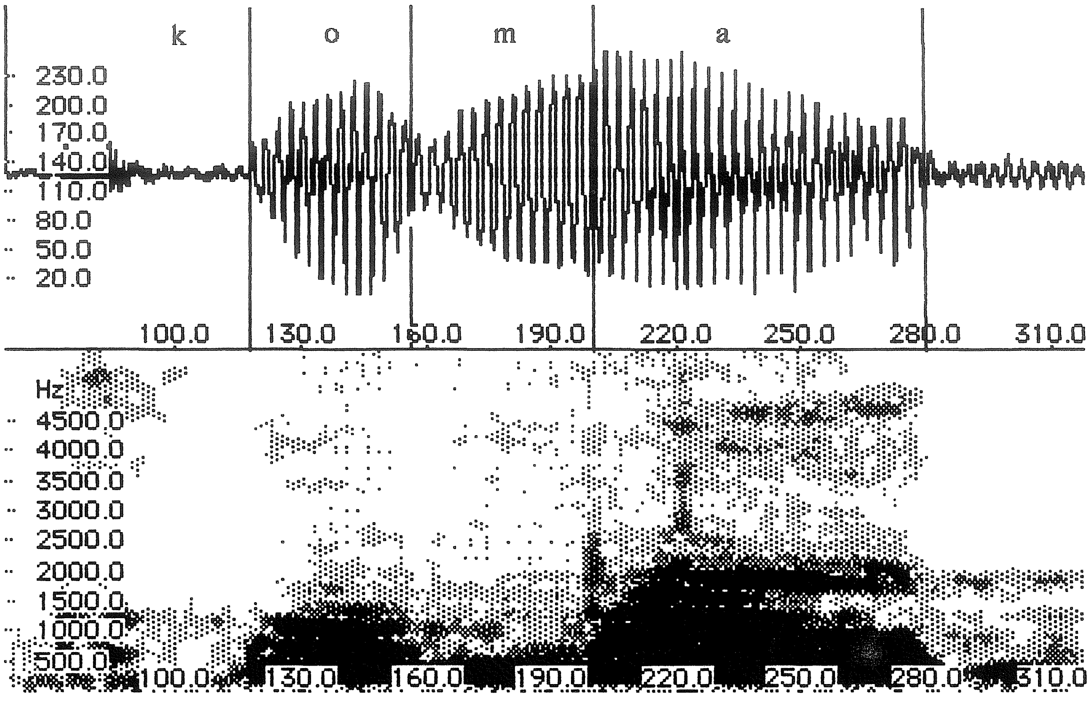


Fig.4.10 [m] in [koma] (example from a Japanese NS).

The duration of nasal consonants such as [m], [n] and [ŋ] in word-internal position is defined with reference to the offset and onset of the preceding and following vowels respectively (Fig.4.10).

Saga can be pronounced as either [saŋa] or [saga], and in some cases learners may alternate between these two.[31] An example of [saga] is shown in Fig.4.11:

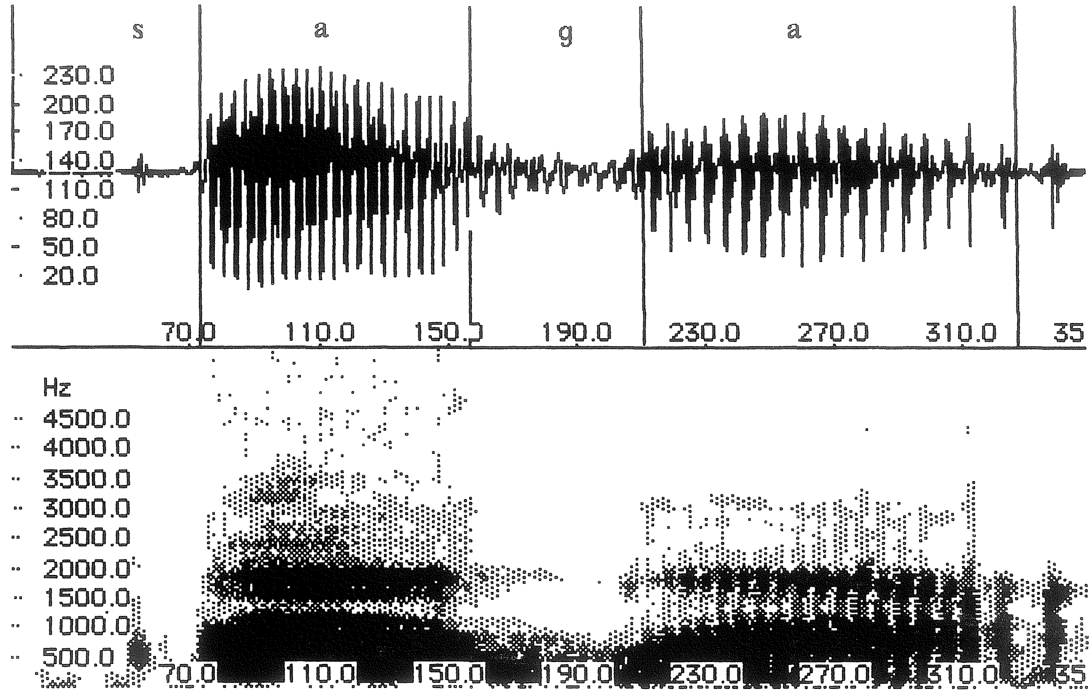


Fig.4.11 [g] in [saga] (example from a Japanese NS) (See Fig.4.14 for [ŋ]).

4.1.3.4 Geminate C2[32]

Stops

The duration of long stops is measured in a similar way to single stops. The offset of the V1 is often abrupt in the speech production of NS as shown in Fig.4.12.

---

[31] In the present study, there is a tendency for Australian learners to use the form [saŋŋa] more often than [saŋga], because they are taught that it is the correct pronunciation. On the other hand, all Japanese NS observed in this study pronounced *saga* as [saga]. The age group of the Japanese NS was between twenty and thirty. It is known that young speakers of Tokyo standard Japanese use a velar stop [g] while older people use a velar nasal [ŋ] word internally (Kindaichi 1942). According to E. Iwamoto (personal communication), there is a new tendency for young Tokyo speakers to pronounce a velar fricative, for example, [eːɣa] instead of [eːga] 'movie'.

[32] In this thesis, single/geminate distinction is used rather than short/long distinction (eg. [k] vs [kk]) in the description of the phonological durational contrast of C2. This is not because of my preference for the concept of gemination, but simply to avoid the following confusion: during the observation of learners' speech production, I realised that the duration of geminate segments (eg. long intervocalic consonant in [rikka]) was often shorter than that of single segments (eg. short intervocalic consonant in [rika]). I call this a reverse realisation of phonological contrast (see Chapter 4). When this kind of phenomenon is observed in the speech production of L2 learners, it is unclear whether 'short' or 'long' refers to the contrast at the phonetic or phonological levels.

The short/long distinction, however, may be used in the description of V2 since it is not a common practice to refer to a long V2 as a geminate V2.

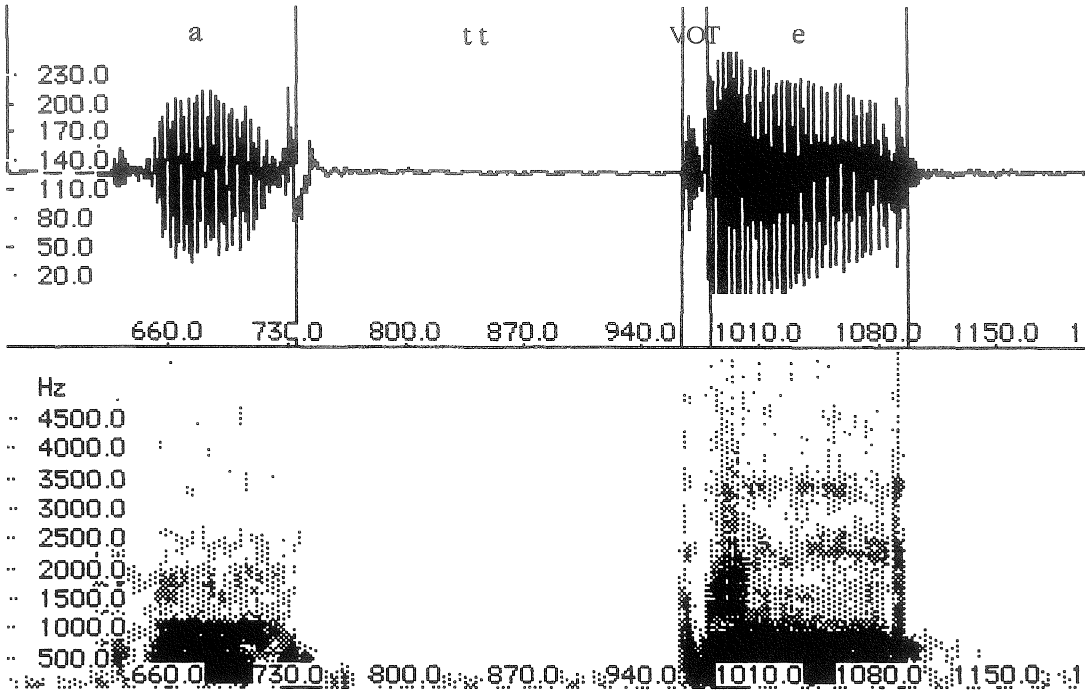


Fig.4.12 [tt] in [katte] (example from a Japanese NS).

Fricatives

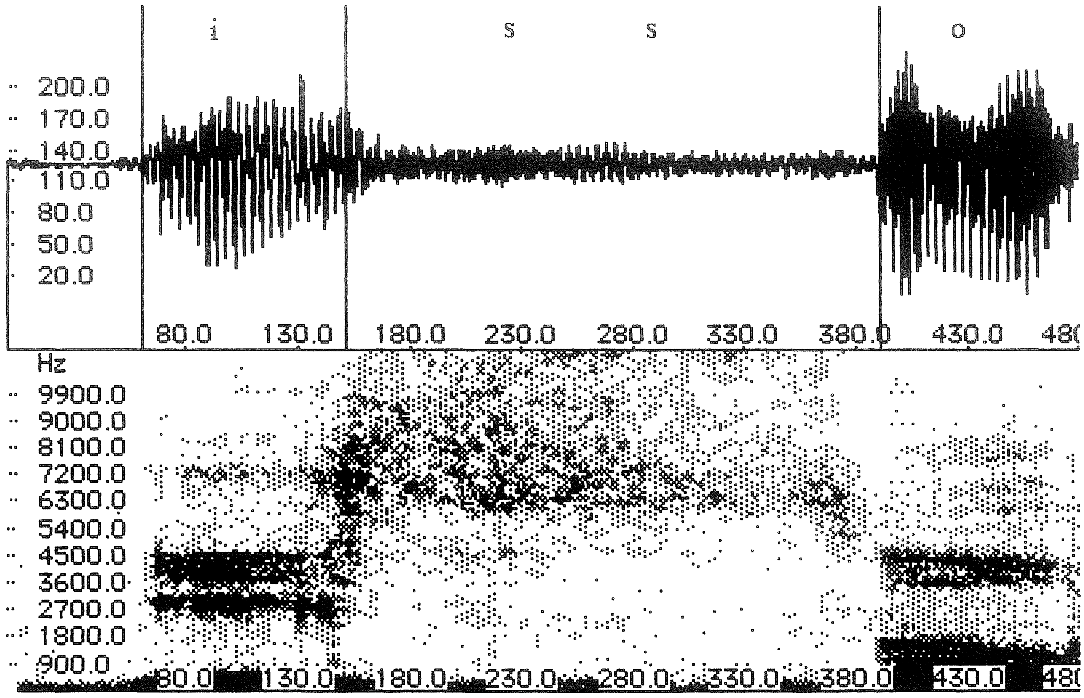


Fig.4.13 [ss] in [isso] (example from a Japanese NS).

The onset of word-internal long fricatives is measured using a full-range wide band spectrogram (Fig.4.13). The criteria for segmentation are similar to those for word-internal short fricatives (Fig.4.9).

Nasals

Moraic nasals in *komma* and *honne* are measured as one unit in the same way as geminate stops in order to compare their duration with their single counterparts.

When *sanga* is pronounced as [saŋŋa], [ŋ] is measured as one unit in the same way as the moraic nasal in *komma* and *honne*. When it is pronounced as [saŋga], on the other hand, [ŋ] and [g] are segmented and their absolute values are obtained independently.

The offset of [ŋ] is recognized by the end point of nasal resonance, which occurs when the soft palate rises for the glottal closure for the following [g], and it is observed between 1000-2000 Hz.

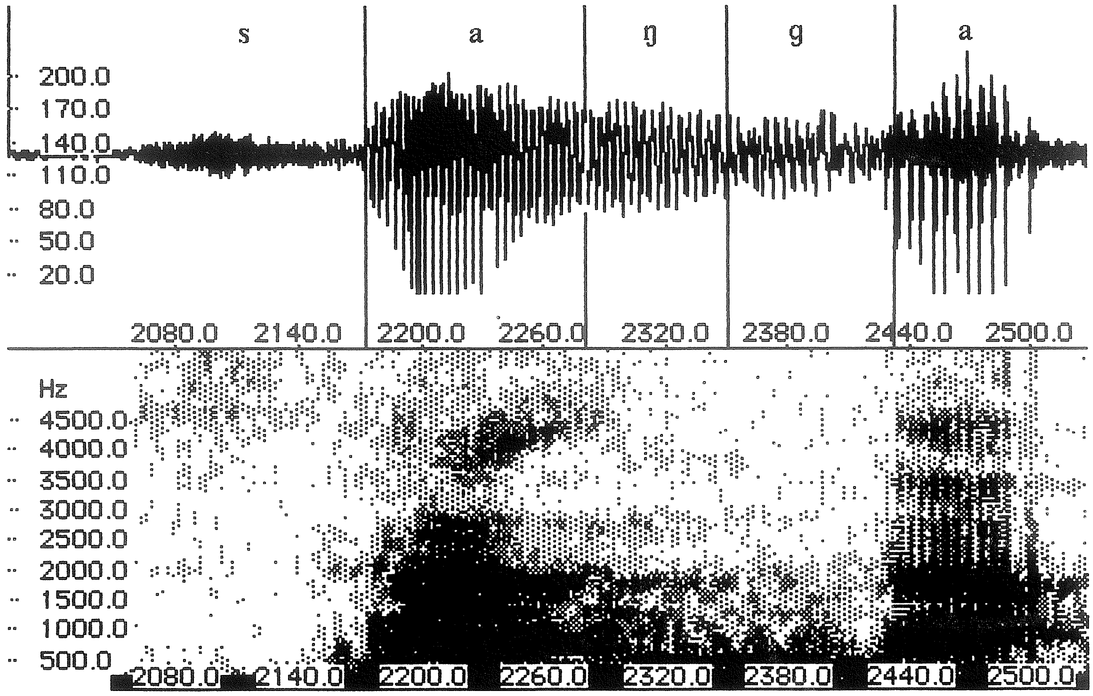


Fig.4.14 [ŋ] in [saŋga] (example from a Japanese NS).

4.1.3.5 V2

Word-final vowels

Vowels in word-final position are usually accompanied by a glottal stop for NS, and often terminate with irregular periodicity (Fig.4.15). The offset position is determined in terms of the end point of periodicity in the waveform and the energy present indicating phonation in the spectrogram.

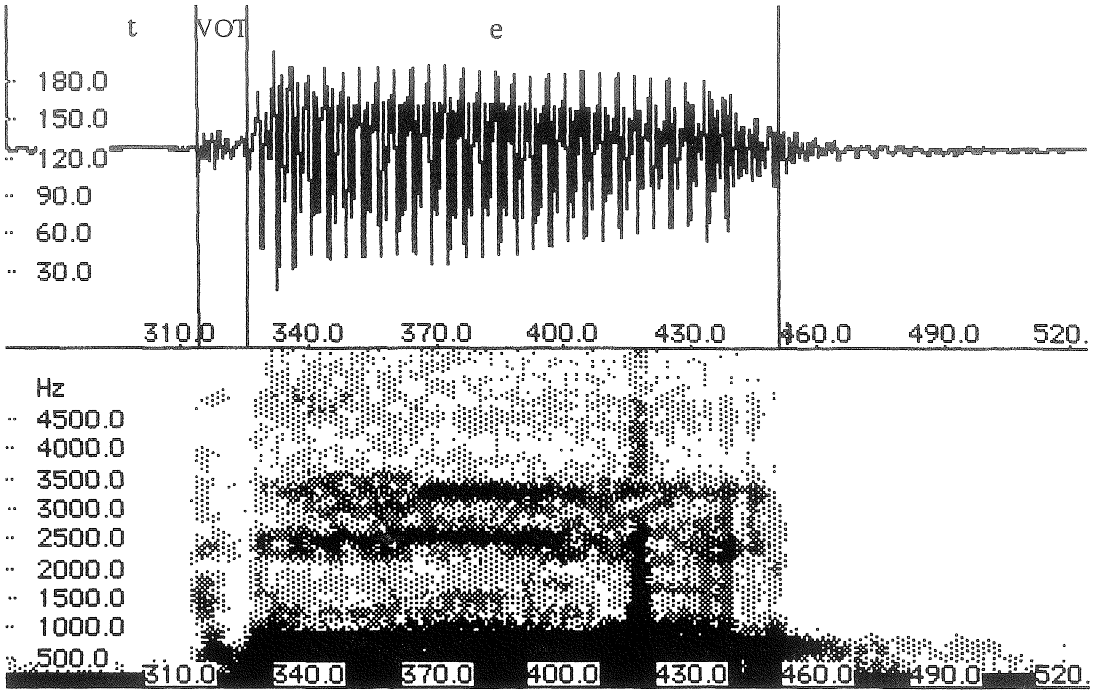


Fig.4.15 [e] in [kate] (example from a Japanese NS).

4.2 Japanese NS

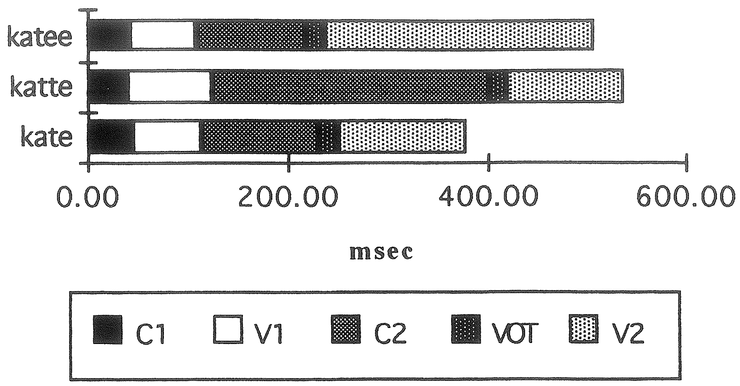
4.2.1 Subjects

The ten Japanese NS who participated in the perception test (Chapter 3) were the subjects of this experiment.

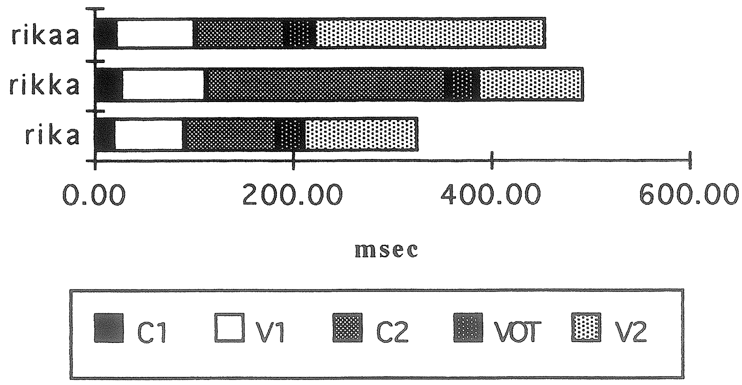
4.2.2 Results

The following graphs and charts show the average figures for the ten Japanese NS (see Appendix 4.1 for average values).

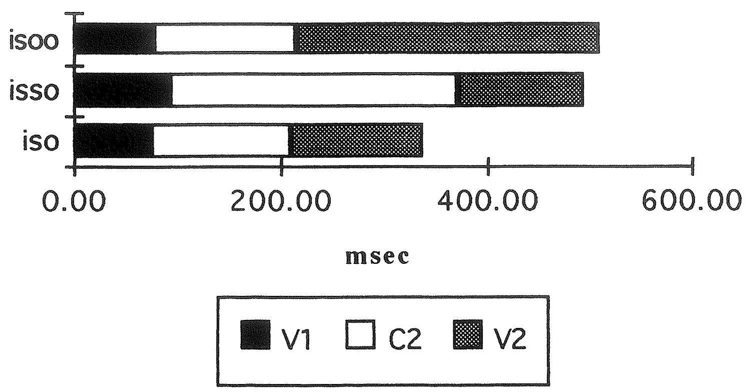
NS Average (1)

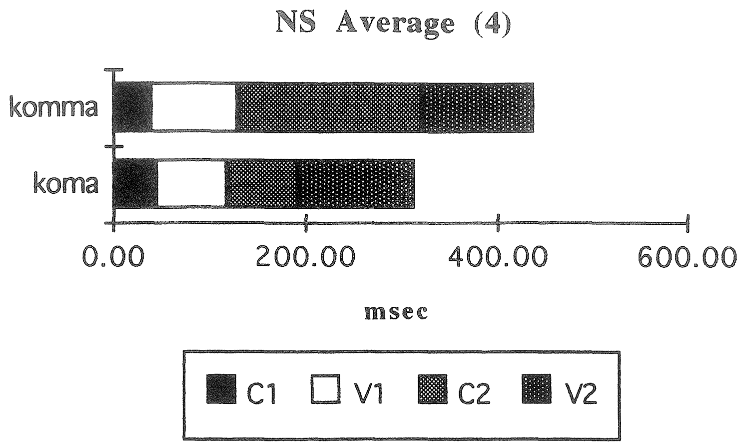


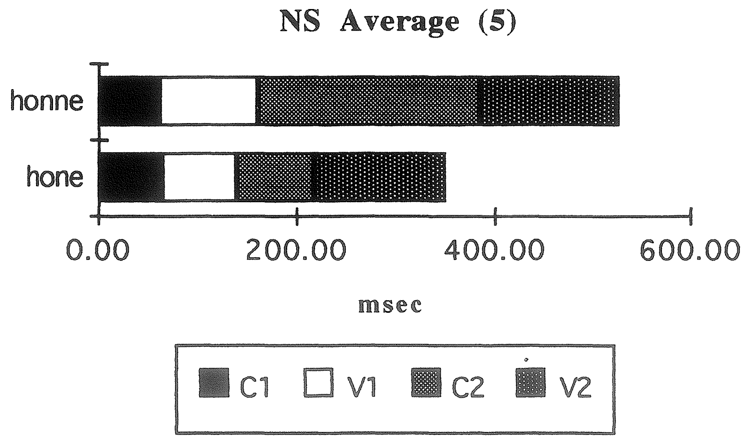
NS Average (2)



NS Average (3)







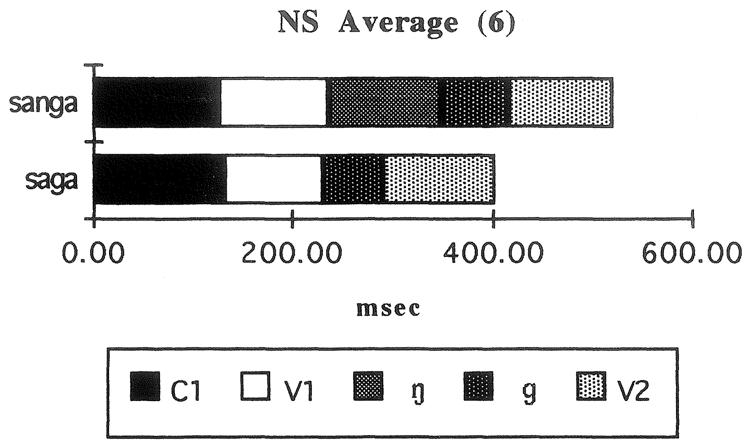


Fig.4.16 NS average (1-6)

4.2.2.1 Word duration

With regard to word duration, the difference between two- and three-mora expressions is always significant (Fig.4.16).

When the word duration of two-mora CVCV expressions is treated as 100, the ratios of the corresponding three-mora CVCCV and CVCVV expressions are as shown in Table 4.2:

SERIES	RATIO
kate/katte/katee	100:142:134
rika/rikka/rikaa	100:151:140
iso/isso/isoo	100:146:151
koma/komma	100:140
hone/honne	100:150
saga/sanga	100:129

Table 4.2 Word duration (ratio)

The average ratio of the three-mora expressions is 143% of the corresponding two-mora expressions. Amongst all the examples, the SAGA series showed the smallest differentiation ratio with respect to word duration (100:129). Therefore, the phonetic reality of the mora is not strictly isochronous. As far as the data presented in this study are concerned, it can be said simply that the duration of a three-mora expression is always longer than that of a two-mora expression.

4.2.2.2 Long/short contrast

There is a clear phonetic contrast at the level of speech production corresponding to the phonological long/short opposition. Long C2 and V2 are at least twice as long as their short counterparts (see Table 4.3 below). For example, the average duration of the long C2 in [rikka] is 2.6 and 2.7 times as long as the short C2 in [rika] and [rikaa] respectively. (The ratio of the RIKA series is 39/100/37 when the duration of the long C2 is treated as 100.) This is similar in the case of nasal C2, where the duration of the long [m] is 2.6 times as long as its short counterpart.

	C2	V2
1 kate	42	47
2 katte	100	44
3 katee	39	100
4 rika	39	49
5 rikka	100	46
6 rikaa	37	100
7 iso	48	44
8 isso	100	42
9 isoo	49	100
10 koma	39	
11 komma	100	
12 hone	34	
13 honne	100	
14 saga	34	
15 sanga	100	

Table 4.3 Long/short contrast (ratio)

The question is whether any phonetic evidence for the mora hypothesis, such as 'A geminate stop is composed of a moraic stop plus a single stop, and its total duration is expected to be longer than twice the duration of a single stop' (Han 1992:103) can be found in the present data. Vance (1987:71) also suggests that we can maintain the claim that morae are isochronous as long as the duration of a geminate C2 is significantly longer than twice the duration of a single C2. In the present study, it was found that the data supported this hypothesis as far as stop C2 was concerned: the duration of a geminate stop was longer than twice the duration of a single stop.

The data with fricatives, on the other hand, demonstrated that the duration of a geminate C2 was only just twice the duration of a single C2. For example, in the case of the fricative intervocalic C2, the duration of the long C2 in [isso] was only 2.1 and 2.0 times as long as the short C2 in [iso] and [isoo] respectively. (The ratio of C2 in the ISO series is 48:100:49 when the duration of long C2 is treated as

100.) In other words, [ss] is merely an equivalent of a geminated single [s] (i.e. [ss]=[s]x2), and there is not any larger unit, namely a 'moraic unit'.

This situation is similar for the long V2 in [rikaa], which is 2.02 and 2.19 times as long as its short counterparts. (The ratio of V2 in the RIKA series is 49:46:100 when the long V2 is regarded as 100.)

The findings imply that the concept of the mora, which treats geminate stops and fricatives as one unit at the phonological level, actually demonstrates significant phonetic differences in their timing organizations. This was not reported in the study by Han (1992), since she only observed words with voiceless stops [p], [t] and [k]. I argue that theories concerning the phonetic reality of the phonological concept of 'mora' need to be powerful enough to account for words with fricatives as well as stops. It is also necessary to account for the long/short contrast in vowels since their durational contrast is also related to the concept of 'mora'. Unless the theory applies to all the elements which are regarded as 'mora', it is only partially valid.

The results obtained from this study show that geminate stops are more than twice as long as their short counterparts, thus providing phonetic justification for the assumption that geminate stops are a combination of a moraic stop and a single stop. It was found, however, that a geminate fricative C2, which is also phonologically composed of a moraic fricative and a fricative stop, did not have a total duration that was significantly longer than twice that of a single fricative C2. Similar examples were also found for vowel duration. Based on these results, it can be said that Han's (1992) hypothesis only accounts for stops, and not the entire picture of the phonological concept of mora in Japanese which should include long fricatives, nasals and vowels.

Simple generalizations cannot be made about the relationship between the long/short contrast and the place of articulation since the segmental structures and pitch patterns of the examples are not identical. As far as the present data are concerned, it seems that the degree of durational contrast is greater in stops than in fricatives (both the

RIKA and KATE series have a greater durational contrast than the ISO series), and that the velar stops in the RIKA series show the greatest ratio of differentiation.

4.2.2.3 C2/V1 and V2/V1 ratios

A clear short/long contrast can also be found for the C2/V1 ratio. This is also about twice as large for words with a long obstruent C2 as for those with a short C2. V2/V1 ratios also show similar tendency.

This result demonstrates that the V1 duration of two- and three-mora words in the same series is consistent in NS speech production. C2/V1 and V2/V1 ratios are given in Table 4.4.

	C2/V1	V2/V1
1 kate	1.68	1.82
2 katte	3.21	1.37
3 katee	1.62	3.99
4 rika	1.31	1.62
5 rikka	2.82	1.24
6 rikaa	1.13	2.93
7 iso	1.79	1.72
8 isso	3.01	1.33
9 isoo	1.75	3.75
10 koma	0.97	1.54
11 komma	2.13	1.29
12 hone	1.00	1.71
13 honne	2.28	1.40
14 saga	0.63	1.13
15 sanga	1.66	0.94

Table 4.4 C2/V1 and V2/V1 ratios

The ratio of long C2 and V2 in relation to V1 is around 2.8-3.2 for obstruent C2, 2.9-4.0 for V2 and 1.7-2.3 for nasal C2. The ratio of a long nasal C2 is the smallest in relation to the duration of the preceding vowel. This may be related to the fact that the vowel preceding a nasal C2 is usually nasalised.

4.2.2.4 VOT

VOT was measured in order to test the following two hypotheses (Chapter 2):

1. VOT is significantly smaller for [t] than [k].

2. VOT is significantly smaller for geminate C2 than single C2.

	t	k	t-test
C	18.27 msec (5.90)	25.92 msec (5.24)	t=-7.51, p=.0001
CC	15.27 msec (4.43)	28.26 msec (7.27)	t=-8.36, p=.0001
t-test	t=2.46, p=.0159	t=-1.752 p=.0832	

Table 4.5 VOT

T-tests were undertaken in order to observe if the difference in VOT between [t] and [k] were statistically significant. It was found that VOT for single [t] (18.27 msec) was significantly smaller than single [k] (25.92 msec) (t=-7.51, p=.0001). This was the same in the corresponding geminate C2 (t=-8.36, p=.0001). This result is consistent with Lisker and Abramson's (1964) phonetic universal; that is, VOT increases as the place of articulation moves toward the back and supports the claim that the phonetic universals influence speech production by Japanese NS.

On the other hand, the difference between VOT for single and geminate C2 was significant in [t/tt] (t=2.46, p=.0159) and not significant in [k/kk] (t=-1.752, p=.0832). Therefore, as far as the present data are concerned, it is difficult to say that the Japanese geminate C2 is consistently associated with a shorter duration of VOT, and it may be universal (Han 1962a).

4.2.2.5 Pitch assignment

Some NS speakers pronounced [honne] (LHH) with a high-level tone (HHH). This may be due to the anticipatory (progressive) assimilation of the L tone, which is assigned to the initial syllable [ho], with the H tone which is assigned to the following moraic /N/, and seems to occur with natural or fast speech. When the word is pronounced slowly and clearly, the L tone seems to be assigned to the first mora of the word and the pitch pattern of the word is the underlying assignment of the tone, LHH.

4.2.3 Summary

The foregoing section showed the average values for NS speech production. It was found that the word duration of two- and three-mora expressions was clearly differentiated in NS speech. Also there was a clear indication of the long/short contrast in single/geminate stop C2, although it was found that the geminate fricative C2 did not show a duration which was significantly longer than twice the duration of a single fricative C2. The results obtained in this study provide phonetic evidence that a geminate stop is a combination of a single stop and a moraic stop which is a larger unit than a single stop because the duration of a geminate stop was longer than twice the duration of a single stop. This does not apply for the fricative C2 because the duration of a geminate fricative C2 was just twice the duration of a single C2.

With regard to VOT, it was found that longer VOT was associated with velar stops than with alveolar stops, which is consistent with universal phonetic principles. The present data, however, do not support the claim that the gemination of C2 is universally associated with a shorter VOT (Han 1962a).

Perception and production

Generally speaking, the C2/V1 ratios with geminate C2 obtained for production fall in between the two ratios for perception when the V1 duration is 70 and 140%. For example, the C2/V1 ratio of [katte] in the data of production (3.21) is between those of KA07T_E (3.61) and KA14T_E (1.93). This situation also applies to V2/V1 ratios. For example, the C2/V1 ratio of [katee] (with long V2) in the data of production (3.99) is between those of KA07TE_ (4.48) and KA14TE_ (2.44).

These observations seem to characterise the relationship between perception and production by NS. The ratios for NS production are similar to those for perception if the V1 duration is around 90-100%, i.e. similar to the original utterance before V1 duration was adjusted. In this respect, NS production of long phonemes shows consistency with their perception of long phonemes.

With regard to words with single C2 and V2, the C2(V2)/V1 ratios in speech production are generally smaller than those obtained from perception data. For example, the C2/V1 ratios of [kate/katee] (1.68/1.62) are smaller than KA07_TE (3.61) and KA14_TE (1.93). Similarly, the V2/V1 ratios of [kate/katte] (1.83/1.37) are smaller than KA07TE_ (4.48) and KA14TE_ (2.44).

This means that the C2(V2)/V1 ratios for words with short C2(V2) in NS speech production do not override those of the sound stimuli in which they perceived long C2. In other words, NS generally produce words with short phonemes within the domain of their perceptual category of short phonemes.

4.3 Australian beginning learners

The results obtained from the learners of Japanese show a number of dissimilarities from those of NS. In the following section, the average figures of the ten beginning learners are described first, and then the typological categories which characterise individual learners' speech production are discussed.

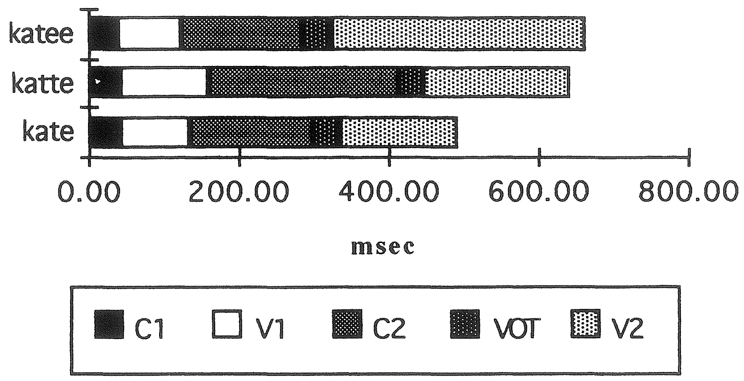
4.3.1 Subjects

The ten ANU beginning learners who participated in this experiment are the same beginners who participated in the perception test (Chapter 3). My auditory impression of their speech production is presented in Appendix 4.2. See Appendix 4.3 for the average values of acoustic measurements.

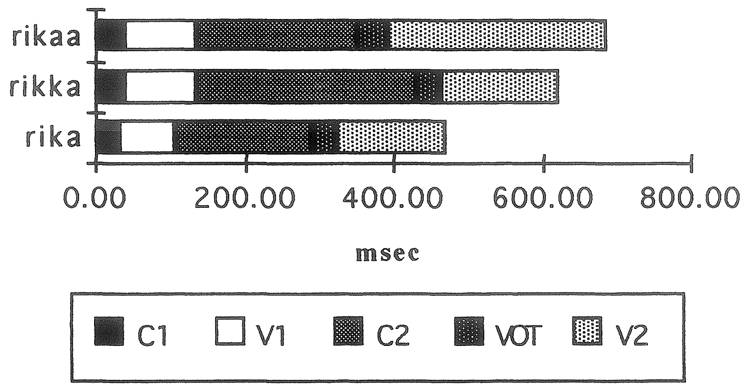
4.3.2 Results

In this section, the average results of the ten beginners are presented. The observation is also extended to the individual speakers' data in order to discuss individual learner strategies.

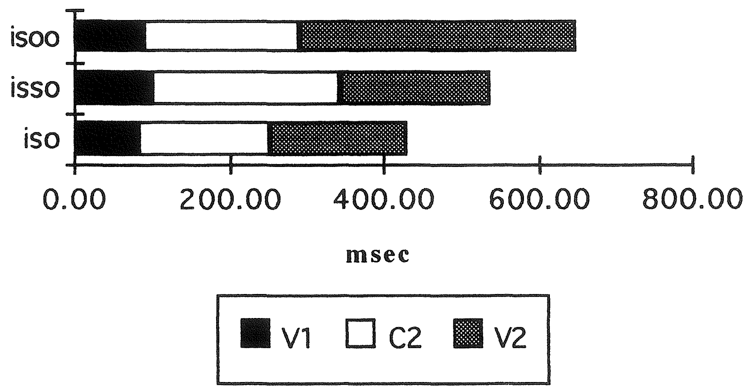
ANU Average (1)



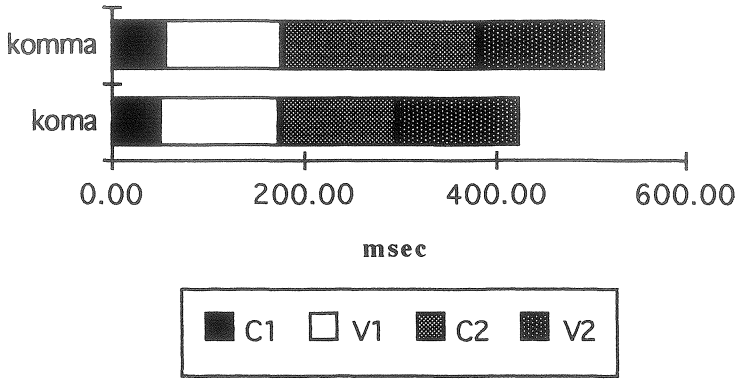
ANU Average (2)



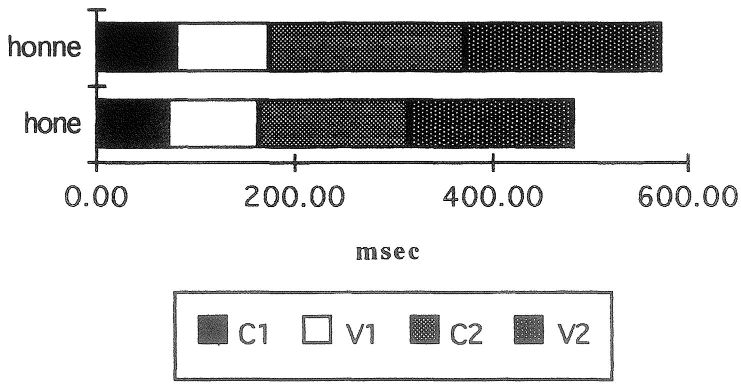
ANU Average (3)



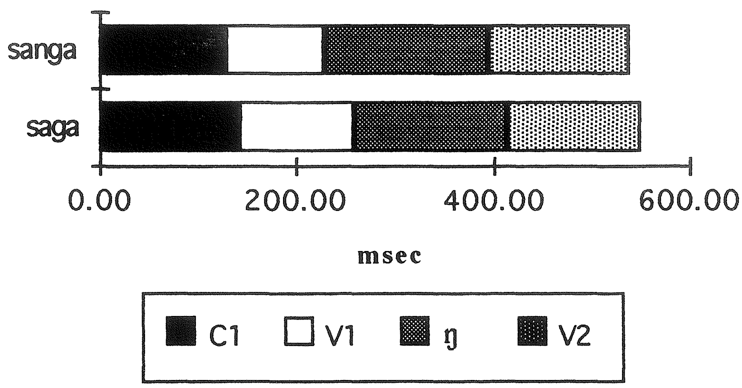
ANU Average (4)



ANU Average (5)



ANU Average (6)



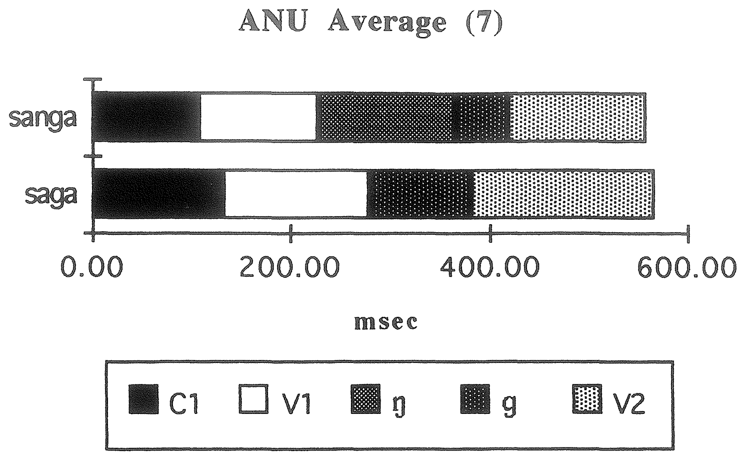


Fig.4.17 ANU average (1-7)

4.3.2.1 Word duration

Unlike Japanese NS, word duration in learners' speech does not always show a consistent pattern.

The word duration is underdifferentiated, especially in those with nasal C2. The average ratios are 100/122 ([koma/kommȧ]), 100/119 ([hone/honne]), 100/98 ([saŋa/saŋŋa]) and 100/99 ([saga/sanga]). In the average figure of SAGA series (Fig 4.17), the duration of two- and three-mora expressions is almost identical. This situation is different from NS data in which the corresponding average ratios are 100/140 (KOMA), 100/150 (HONE) and 100/129 (SAGA).

SERIES	RATIO
kate/katte/katee	100:125:129
rika/rikka/rikaa	100:133:146
iso/isso/isoo	100:125:151
koma/komma	100:122
hone/honne	100:119
saŋa/saŋŋa	100:98
saga/sanga	100:99

Table 4.6 Word duration (ratio)

Some individual learners' data also show a number of instances where the difference between the duration of two- and three-mora expressions is not statistically significant (i.e. underdifferentiation) (see figures shown below); indeed, the word duration of a two-mora expression was sometimes longer than that of a three-mora expression (i.e. reverse realization) (e.g. Fig.4.26 and 29).

For the reasons above, it can be said that learners do not control their timing in the same way as NS in respect of word-level speech production. This implies that beginners have not acquired the concept of mora timing at the level of speech production.

All informants had at least one example of problems of word duration in their speech (see Appendix 4.4 for individual speakers' ratios). Many problems of underdifferentiation were found in words with nasal C2. Some subjects did not have any problems with words which had obstruent C2, but had problems differentiating between the duration of words with nasal C2.

As far as these results are concerned, it can be said that acquisition of the differentiation of word duration between two- and three-mora expressions first takes place in words which have obstruent C2, and later for those with nasal C2, amongst beginning-level Australian learners.

Examples of problems in word duration observed in some of the subjects are illustrated as follows:

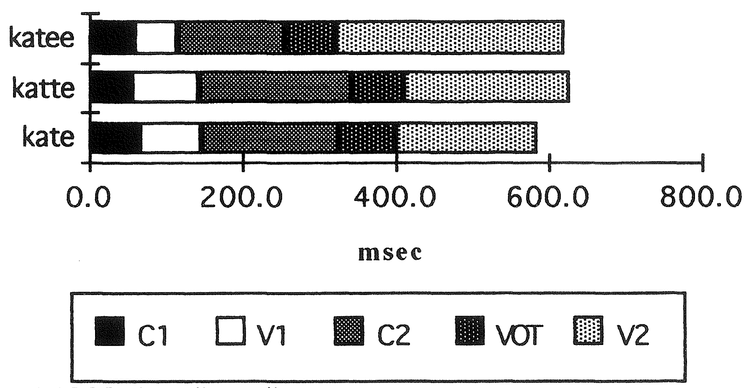


Fig.4.18 NP kate/katte/katee

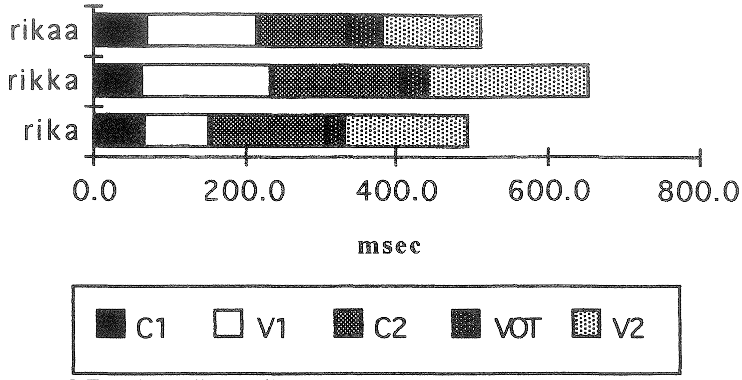


Fig.4.19 BD kate/katte/katee

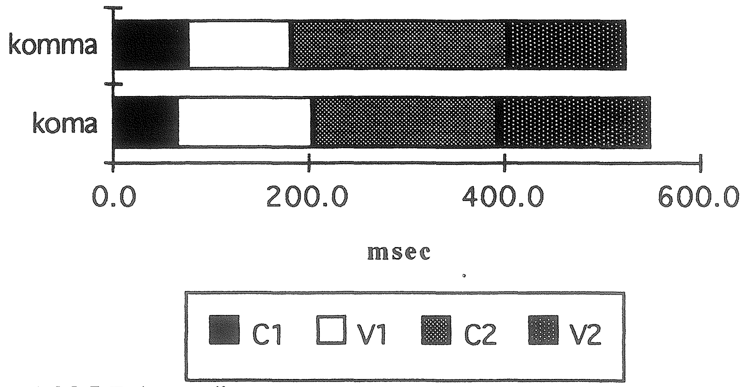


Fig.4.20 LB koma/komma

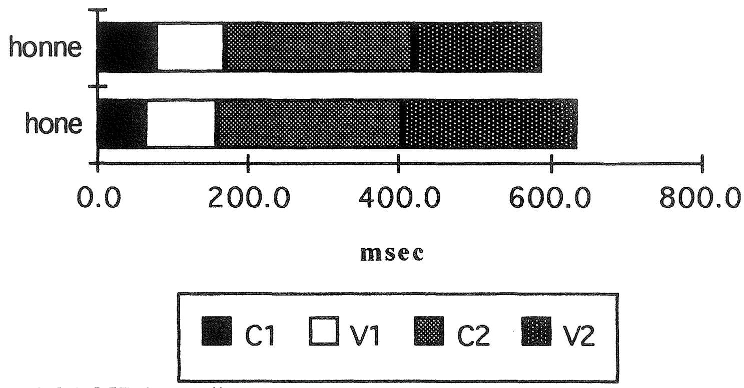


Fig.4.21 NP hone/honne

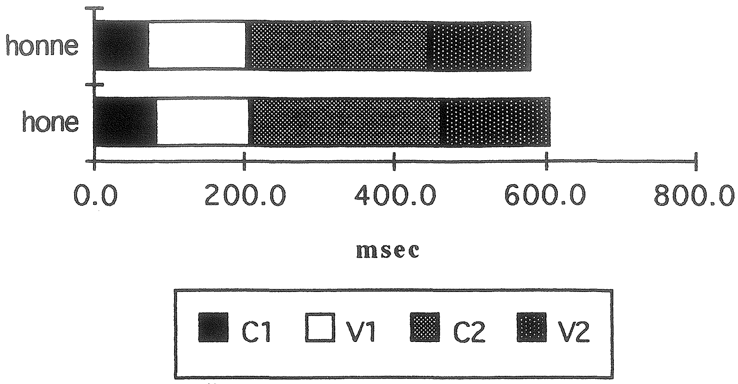


Fig.4.22 LB hone/honne

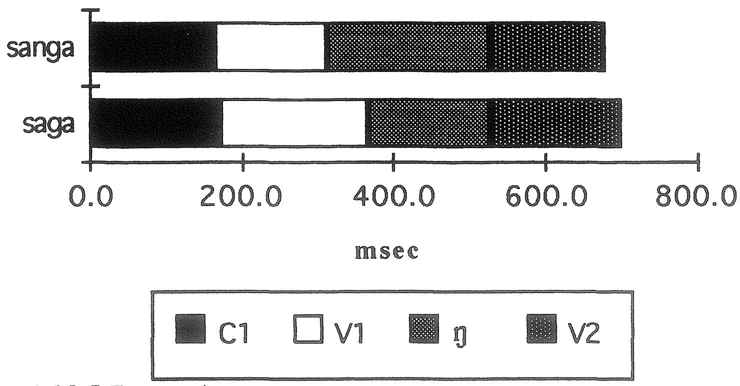


Fig.4.23 LB saga/sanga

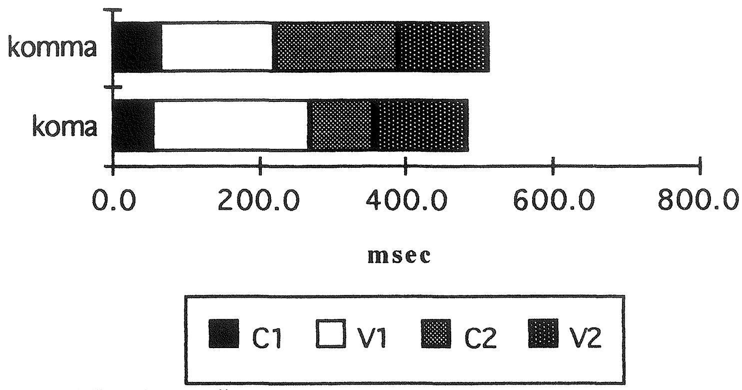


Fig.4.24 BD koma/komma

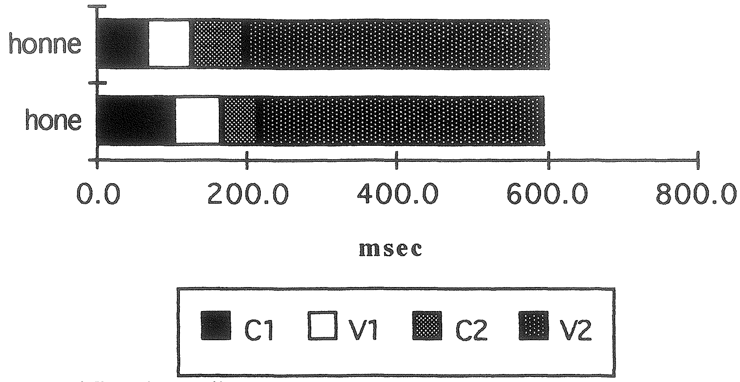


Fig.4.25 BD hone/honne

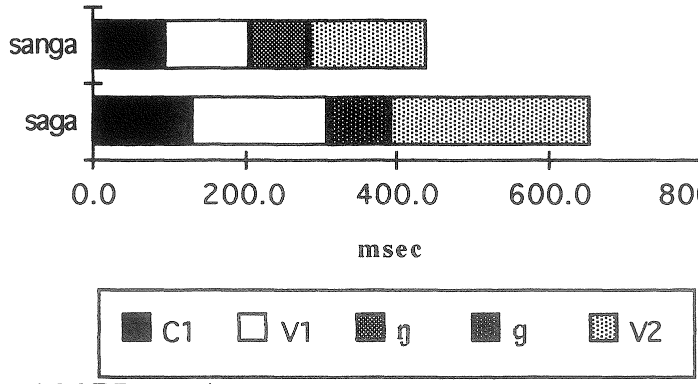


Fig.4.26 BD saga/sanga

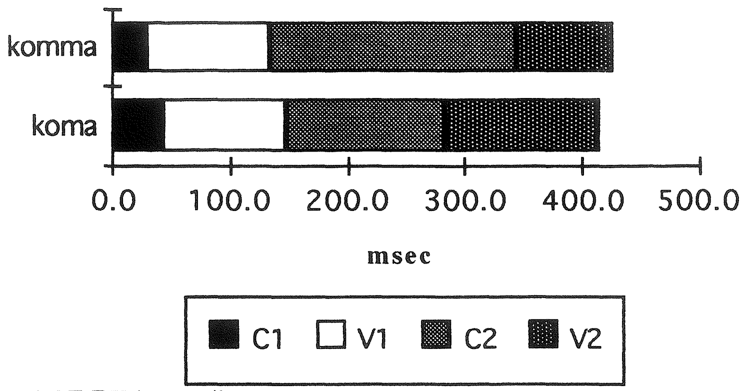


Fig.4.27 PF koma/komma

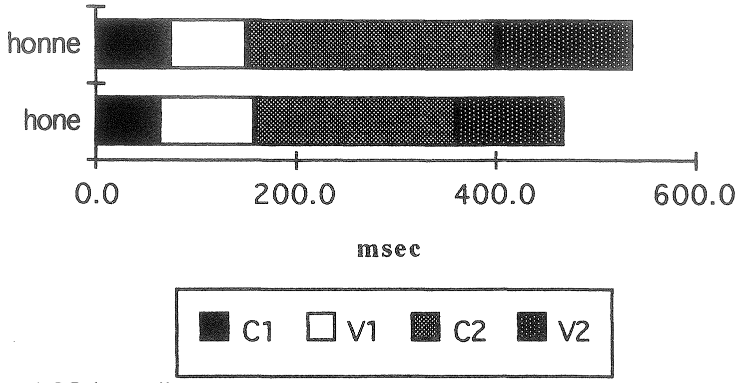


Fig.4.28 hone/honne

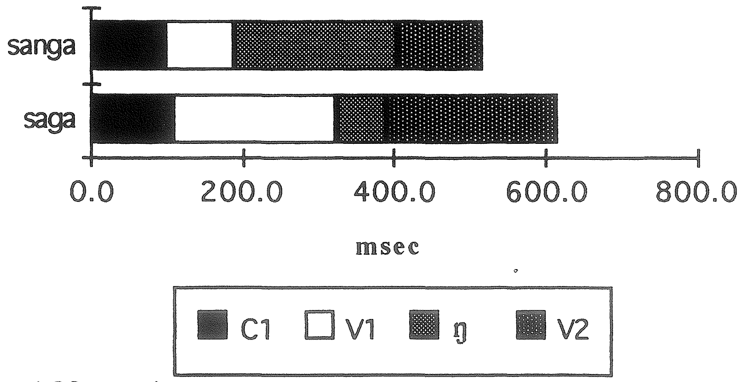


Fig.4.29 saga/sanga

4.3.2.2 Long/short contrast

As shown previously, it was found that there is a clear phonetic contrast corresponding to the phonological long/short opposition at the level of speech production by NS. The rate of differentiation varies across different speakers, but the differentiation can be as small as 2:1.

Table 4.7 shows the average ratios of long/short contrast for beginning learners.

	C2	V2
1 kate	69	48
2 katte	100	58
3 katee	63	100
4 rika	62	50
5 rikka	100	53
6 rikaa	72	100
7 iso	72	49
8 isso	100	55
9 isoo	83	100
10 koma	59	
11 komma	100	
12 hone	77	
13 honne	100	
14 saga	54	
15 sanga	100	
16 saŋa	93	
17 saŋŋa	100	

Table 4.7 Long/short contrast (ratio)

The long/short contrast shown in the beginning learners' data is smaller than those shown in the data of NS. The average ratio of long stops is only 1.5 times as long as their short counterparts. The rate of underdifferentiation of stop C2 for this group of beginning-level Australian learners is greater than Han's (1992a) American advanced-level subjects (2:1).

The ratios of long vowels, on the other hand, are 1.9 times longer than their short counterparts. The rate of underdifferentiation is more prominent in C2 than V2. This may imply that the acquisition of the durational contrast of V2 precedes that of C2.

The cause of underdifferentiation of the long/short contrast may be (1) the absolute duration of a long phoneme being too short compared with the short counterpart, or (2) the absolute duration of a short phoneme being too long in relation to the long counterpart.

It is generally believed that learners of Japanese tend to underdifferentiate the long/short contrast mainly because they tend to produce a shorter duration for geminate C2 (Han 1992). In this study, some examples of underdifferentiation due to a shorter geminate C2 were found as shown below (The ratios of long/short contrast for individual subjects are presented in Appendix 4.5):

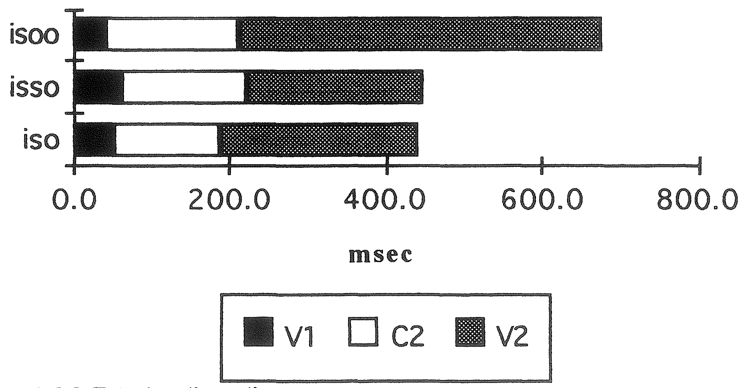


Fig.4.30 BD iso/isso/isoo

It is interesting to note that the longer duration of a single C2 was a more frequently occurring cause of the problem than the shorter duration of a geminate C2, as far as the subjects observed in this study are concerned. Also, a number of words with short C2 were perceived as having long C2 by the author (see Appendix 4.2 for my auditory impression).

Even with single C2 which phonetically has a long duration, it appears that some speakers try to make the single/geminate distinction in their speech production. Fig.4.31 shows that the learner tend to overexaggerate the geminate C2 (overexaggeration) where the duration of the geminate C2 is systematically longer than single C2. This implies that learner is aware of the phonological distinction and therefore try to achieve the long/short contrast, but has to overexaggerate the geminate C2 (to an extent that sounds unnaturally long to Japanese NS) to make the distinction because the single C2 is too long.

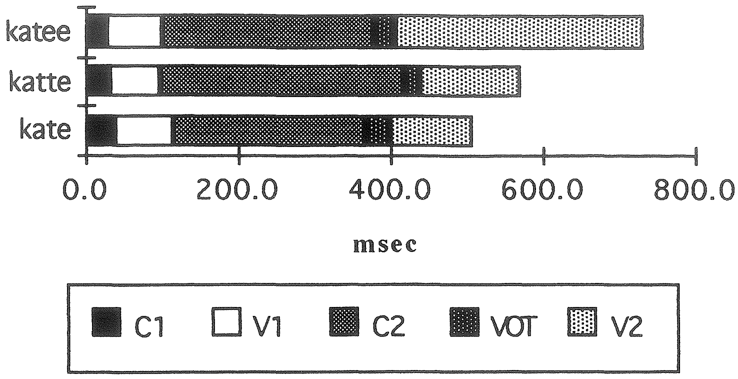


Fig.4.31 LD kate/katte/katee

The following figures show problems which are associated with a longer duration for single C2.

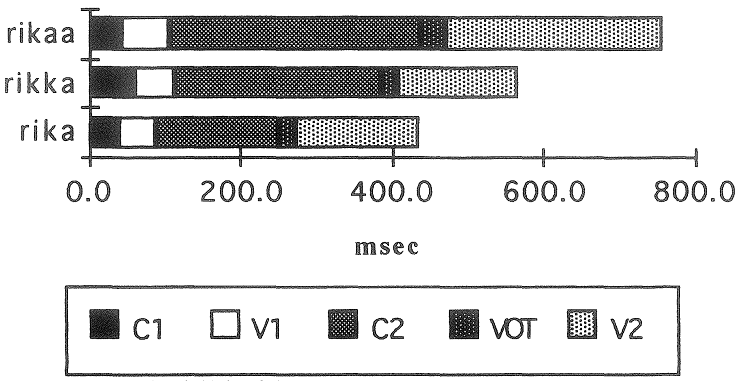


Fig.4.32 PF rika/rikka/rikaa

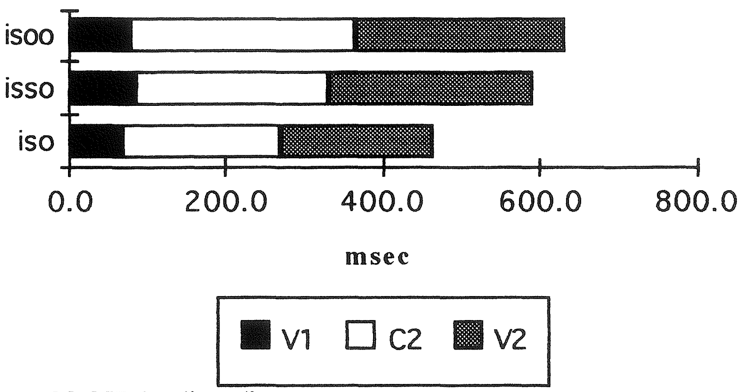


Fig.4.33 NP iso/isso/isoo

It has been mentioned that the beginning learners had considerable problems in the production of words with a nasal C2. Both word-duration and long/short contrast were underdifferentiated in many cases as shown below:

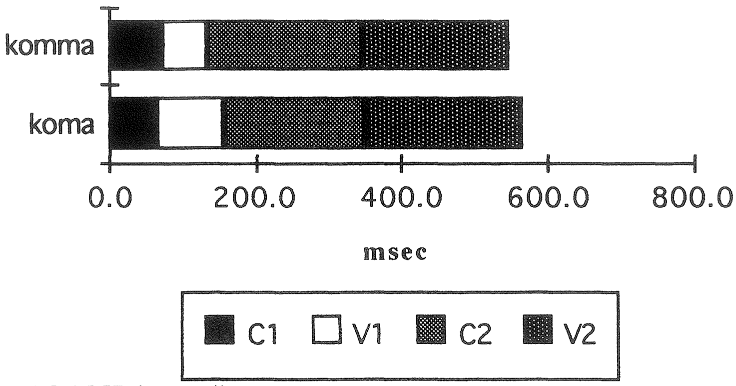


Fig.4.34 NP koma/komma

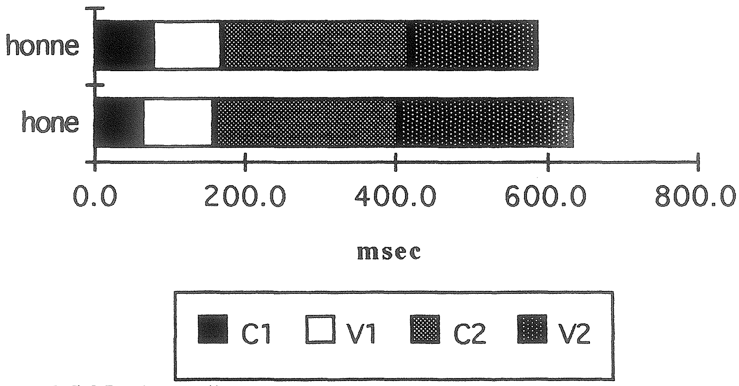


Fig.4.35 NP hone/honne

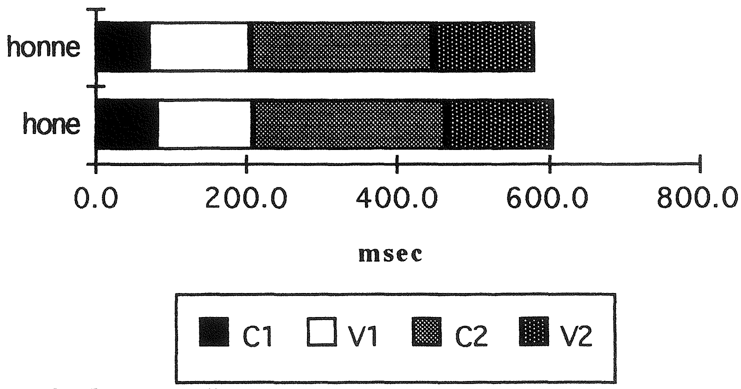


Fig.4.36 LB hone/honne

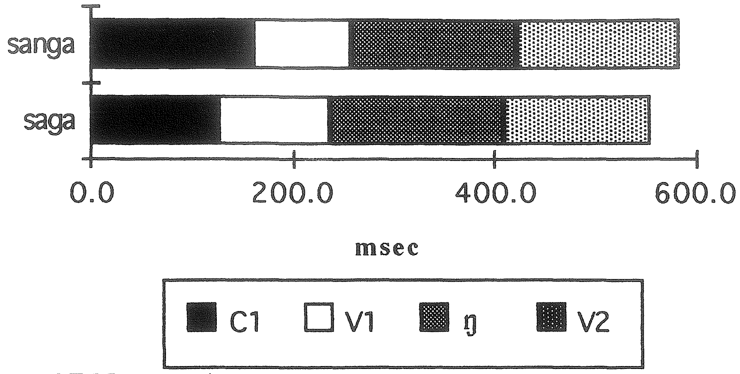


Fig.4.37 NB saga/sanga

Although some subjects used [ŋ] for both two- and three-mora expressions, it is interesting to note that four learners distinguished *saga* and *sanga* in terms of a velar stop [g] and a velar nasal [ŋ] while the durational difference was not significant (Fig.4.38-40).[33] From the viewpoint of IL phonology, we may be able to say that this is one strategy that learners use to make the distinction, instead of manipulating the duration. It may be that the beginning level learners, who do not have full command of manipulating duration to distinguish words, use phonetic distinctions such as [g] and [ŋ] as a strategy for this purpose. This strategy seems to be one easy way to make a clear distinction between the two words without having control of timing. This tendency was not observed in NS speech production.

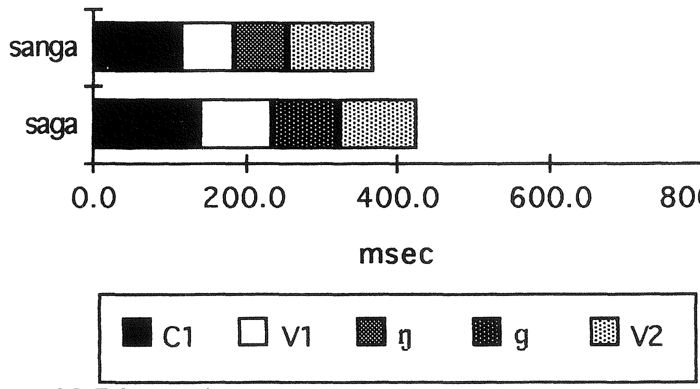


Fig.4.38 BS saga/sanga

[33] The number of beginners who pronounced SAGA series in the following ways is: saŋa (4), saŋŋa (8), saga (6) and sanga (2).

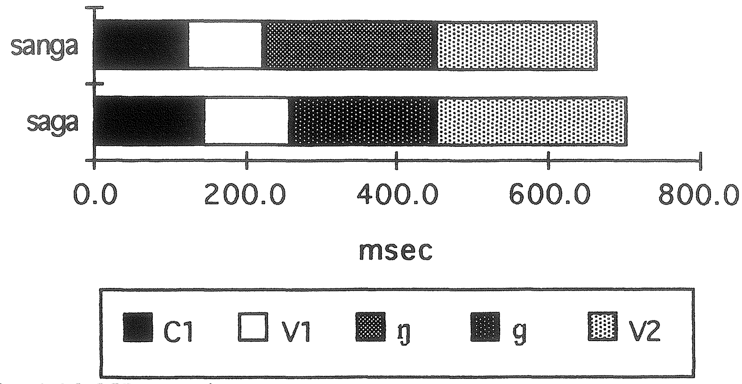


Fig.4.39 NP saga/sanga

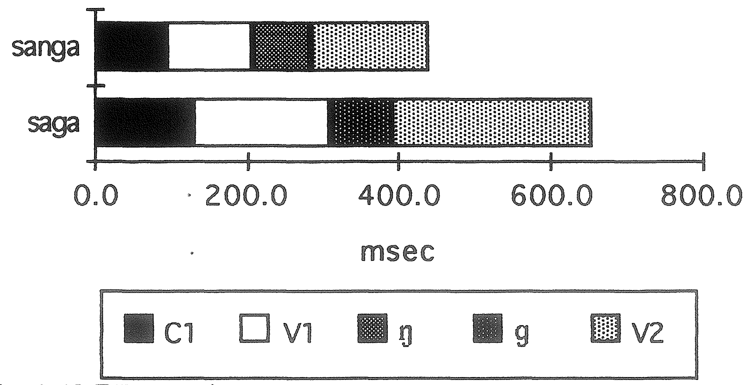


Fig.4.40 BD saga/sanga

With regard to V2, both (1) the absolute duration of a long phoneme being too short compared with the short counterpart, and (2) the absolute duration of a short phoneme being too long in relation to the long counterpart are observed. Underdifferentiation due to the shorter duration of the phonologically long V2 (i.e. (1)) was generally observed in the RIKA series; for example, LB (92/96/100), NP (91/91/100). The short V2 was heard as acceptable by the author (i.e. short) while long V2 gave the impression of being short. On the other hand, underdifferentiation due to the longer duration of phonologically short V2 (i.e. (2)) was often observed in the KATE and ISO series. For example, the problem regarding NB (73/64/100) and NP (73/96/100) in the ISO series was their short V2 which were perceived as long.

Pitch pattern may play an important role in the interpretation of the above data. In English, stressed vowels have longer duration, higher FO and greater intensity. If English-speaking learners tend to associate English stress with H tone in Japanese expressions, a longer duration of a syllable with H tone is likely to occur. In the present data, there is a tendency for underdifferentiation due to the shorter duration of the phonologically long V2 being likely to occur when the pitch pattern is HL; on the other hand, when the pitch pattern is LH, underdifferentiation due to the longer duration of phonologically short V2 seems to be more common.[34]

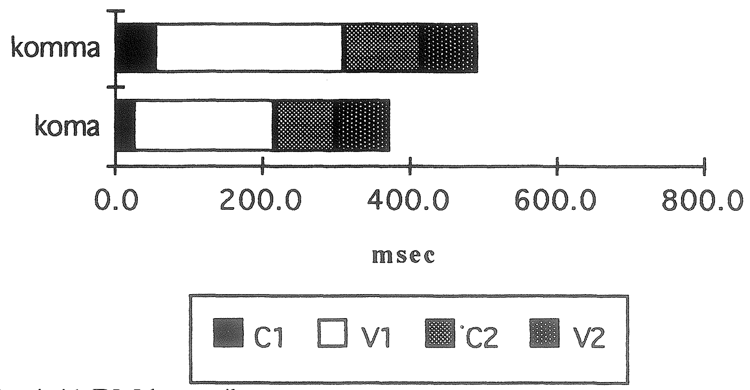


Fig.4.41 BM koma/komma

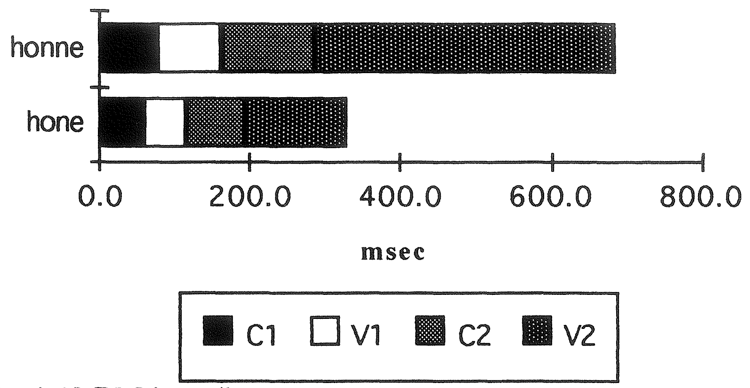


Fig.4.42 BM hone/honne

[34] The fact that the pitch pattern affects segmental duration in the speech production of Japanese learners has been reported by Sugito (1982). She measured the duration of V1 and C2 in [kami HL] 'god' and [kami LH] 'paper' pronounced by four NS and four English learners of Japanese, and found that not only the duration of the vowel, but also the duration of the intervocalic consonant [m], was longer in the syllable with a H tone in the learners' speech.

The above graphs show examples of the interaction between durational contrast and pitch pattern in the learners' speech production: the durational contrast of C2 is realized in either V1 or V2, depending on the pitch pattern of the words. The durational contrast of C2 seems to be realized in terms of the duration of V1 when the pitch pattern is HL (Fig.4.41), and in terms of the duration of V2 when the pitch pattern is LH (Fig.4.42). This is because not only the duration of V in the syllable with an H tone is longer compared to that in the syllable with an L tone, but also the duration of V before or after a geminate C2 is longer than that of V before or after a single C2 (e.g. [o] in *komma* is longer than that in *koma*). Based on this, it can be said that the realization of the durational contrast of C2 has been shifted to the preceding or following vowel. Therefore, it seems that learners' phonetic output is conditioned by both the segmental structure and the pitch pattern of the words.

The word-final glottal stop is another parameter which needs to be taken into account in the analysis of the durational contrast in V2. NS almost always produced a glottal stop in word-final position. In the learners' examples in which the short V2 was perceived as long by an NS, on the other hand, the glottal stop was not usually produced in word-final position. This seems to be another cause of problems related to the production of long/short contrast of V2.

Finally, observations of the learners' data reveal that there are cases where the duration of single phonemes is longer than that of their geminate counterparts (Appendix 4.5). For example, NB's data show that the duration of single C2 in [rika] and [rikaa] is 122% and 164% of the geminate C2 in [rikka] (see Fig.4.43). This situation can also be observed in V2. For example, BM's data show that the duration of the short V2 in [rika] and [rikka] is 124% and 162% of the long V2 in [rikaa] (see Fig.4.44). BD produced a similar result (107/122/100) for the same examples.

I term this phenomenon 'reverse phonetic realization of the phonological durational contrast', in order to distinguish it from underdifferentiation. It was also reported in Toda (1993a) concerning another group of adult subjects who were undertaking intensive Japanese training (also in Chapter 6).

Fig.4.43 and 44 figures show examples of reverse realization of phonological contrast:

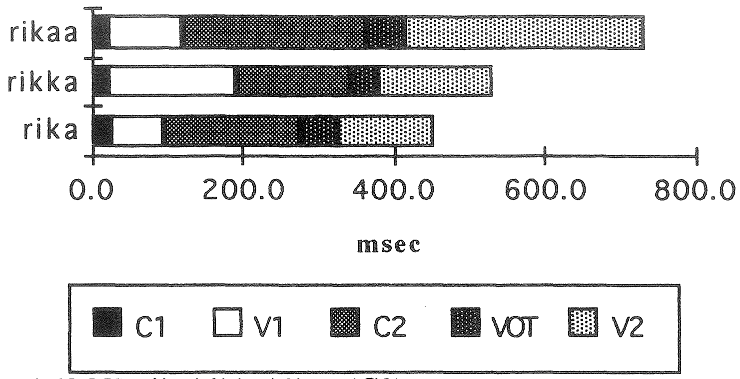


Fig.4.43 NB rika/rikka/rikaa (C2)

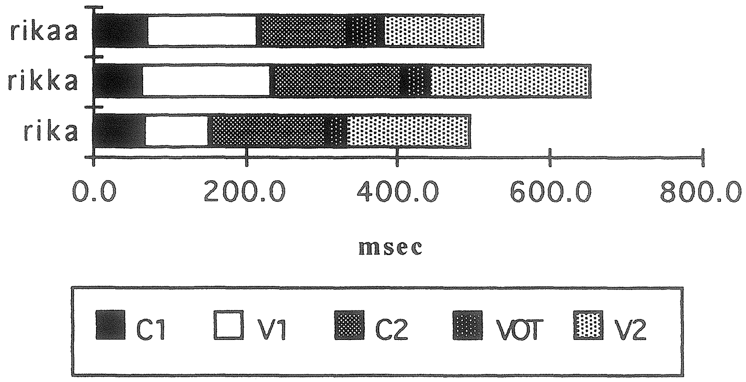


Fig.4.44 BM rika/rikka/rikaa (V2)

The following 4.3.2.3 speculates further about reverse realization of the durational contrast on the basis of C2/V1 ratios.

4.3.2.3 C2/V1 and V2/V1 ratios

Generally speaking, the C2/V1 ratios in CVCV and CVCVV obtained from beginning learners are greater than those obtained from NS. These results seem to suggest that the duration of single C2 is too long in relation to the duration of V1 in words with short C2 in the learners' speech. For example, the average C2/V1 ratios for the ISO series are 2.11/2.41/2.22 (Table 4.8) for beginning learners while they are

1.79/3.01/1.75 for NS. Due to the long duration of single C2, both CVCV and CVCVV are perceived as having geminate C2 in NS auditory impressions. This problem is more prominent than smaller C2/V1 ratios associated with CVCCV (i.e. short duration of geminate C2).

	C2/V1	V2/V1
1 kate	1.86	1.71
2 katte	2.15	1.67
3 katee	1.87	3.95
4 rika	2.46	1.96
5 rikka	3.06	1.63
6 rikaa	2.20	3.03
7 iso	2.11	2.16
8 isso	2.41	1.95
9 isoo	2.22	3.94
10 koma	0.97	1.01
11 komma	1.70	1.07
12 hone	1.64	1.83
13 honne	2.12	2.17
14 saga	0.69	1.26
15 sanga	1.59	1.18
16 saŋa	1.32	1.16
17 saŋŋa	1.65	0.88

Table 4.8 C2/V1 and V2/V1 ratios

Similar problems are found with regard to the V2/V1 ratios of short V2. For example, the average V2/V1 ratios for the ISO series are 2.16/1.95/3.94 for beginning learners, while they are 1.72/1.33/3.75 for NS. This means that the duration of V2 is too long in relation to that of V1 in CVCV and CVCCV.

In the data obtained in this section, it was found that the longer duration of single C2 (or V2) is also causing underdifferentiation as well as shorter duration for long C2 (or V2).

In the data of individual subjects (Appendix 4.6), the C2/V1 ratio for words with a long C2 is often smaller than that for words with a short C2. In other words, a discrepancy between phonology and phonetic realization is found in terms of the C2/V1 ratio as well as long/short contrast in the speech production of the subjects.

For example, the durational ratio of C2/V1 for:

The KATE series is 1.04/0.71/1.08 in BS (Fig.4.45)

The RIKA series is 2.62/0.87/2.54 in NB (Fig.4.46)

The ISO series is 2.03/1.63/2.46 in PF (Fig.4.47).

Examples of a smaller C2/V1 ratio due to the longer duration of V1 are shown below:

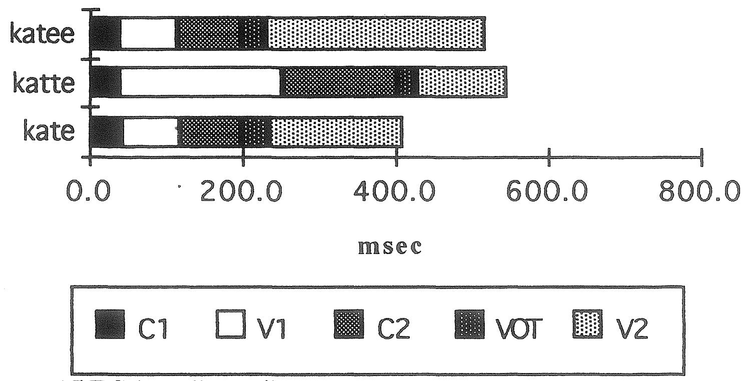


Fig.4.45 BS kate/katte/katee

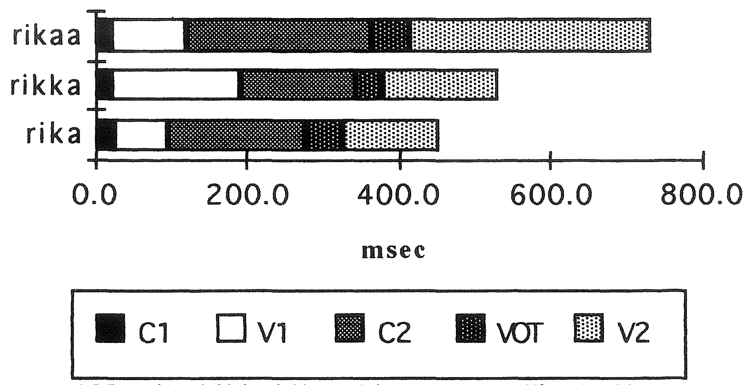


Fig.4.46 NB rika/rikka/rikaa (the same as Fig.4.43)

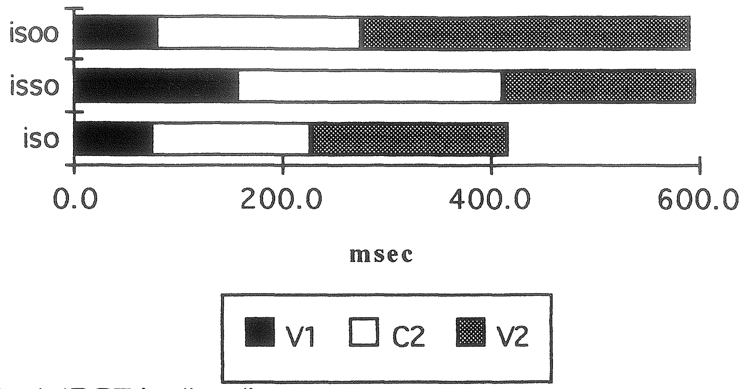


Fig.4.47 PF iso/isso/isoo

The possible causes of this problem are: (1) the absolute duration of geminate C2 being too short compared to the single counterpart, (2) the absolute duration of single C2 being too long in relation to the geminate counterpart and (3) the duration of the preceding vowel V1 being too long compared to the duration of geminate C2. In the previous section, underdifferentiation due to (1) and (2) was discussed; the tendency for the duration of V1 to be prolonged before phonologically long C2, which is shown as (3), is discussed below.

At first, the duration of the first syllable appears to be inconsistent in all three examples, but in some examples the duration of V1 is consistently prolonged before a geminate C2. The results of Anova tests show a significant difference for the duration of V1 before a geminate C2: BS's KATE series ($F=23.729$, $p=.0014$), NB's RIKA series ($F=19.179$, $p=.0025$) and PF's ISO series ($F=13.074$, $p=.0065$).

These results seem to reflect one of the strategies for distinguishing long and short C2. In other words, learners are aware of the durational contrast at the phonological level, and therefore try to achieve the distinction. Their implementation strategies and phonetic realization, however, are different from those of NS.

One possible interpretation of the data is that the subjects have certain degrees of L1 interference from English phonology, which is based on CVC-syllable structure with respect to timing control. As there is no case where a syllable-final C is prolonged in

CVC-syllables in English, perhaps learners find it easier to manipulate the duration of vowels, rather than controlling the duration of syllable-final C. Another possibility is that some learners regard longer V1 before geminate C2 reported in the literature (Fukui 1978; Beckman 1992; Han 1992; Smith 1995)) as a cue for geminate C2 and exaggerated it over C2 durational difference, which is more difficult to achieve. If this is a strategy some learners use, it is not an effective one from the viewpoint of speech perception on the part of Japanese NS, because their perceptual judgment of geminate C2 can be affected by the duration of the preceding vowel, as has been demonstrated in the experiments in Chapter 3. This means that if the duration of the V1 becomes longer before a geminate C2, the duration of the C2 needs to be even longer (instead of shorter as found in the speech production of these learners) in order to be perceived as geminate.

4.3.2.4 VOT

In NS data shown previously, it was found that VOT increased as the place of articulation moved toward the back. It was therefore claimed that phonetic universals influenced NS speech production. The other finding was that the difference between VOT of single and geminate C2 was not significant. In this section, VOT in the learners' speech production is observed.

The following table shows VOT of the beginning learners.

	t	k	t-test
C	36.47 msec (18.54)	36.41 msec (16.21)	t=.018, p=.9858
CC	29.88 msec (16.52)	32.94 msec (12.71)	t=-.804, p=.4246
t-test	t=1.646, p=.1033	t=1.025 p=.3081	

Table 4.9 VOT

1) The average VOT is always greater than NS.

2) The standard deviation is always greater than NS.

3) The difference in VOT between [t] and [k] was not statistically significant ($t=.018$, $p=.9858$). This situation was the same in the corresponding geminate C2 ($t=-.804$, $p=.4246$). Unlike with NS, VOT does not increase as the place of articulation moves toward the back. Therefore, as far as VOT is concerned, it is not possible to say that universal principles operate in this group of beginning learners' speech production in the same way as for NS.

4) Gemination of C2 does not affect VOT.

The longer duration of both C2 and VOT associated with learners' speech production affect the auditory impression by Japanese NS. It has been pointed out that long duration for single C2 and VOT tend to increase the level of unnaturalness (Minagawa 1994).

4.3.2.5 Pitch assignment

The assignment of pitch for words with a nasal C2 is often different from that of the NS. Phonologically speaking, the tone assigned to the second mora must be different from that of the first mora (Chapter 2). In the production of all the beginning learners, however, *honne* was perceived by NS with the syllable [hon] assigned an L tone and the syllable [ne] assigned an H tone (see Appendix 4.2). In *komma* and *sanga*, [kom] and [saŋ] are assigned an H tone, and [ma] and [ga/ŋa] an L tone.

The F0 pattern of examples from [komma], [saŋŋa] is illustrated in the following narrow-band spectrograms, which reflect the fact that the F0 does not show a significant fall before the onset of the final vowel (Fig.4.48 and 49). These examples sound unnatural to NS, since an H tone is expected to be assigned to the first morae /ko/ and /sa/, and an L tone to the second mora /N/ as well as to the third morae /ma/ and /ŋa/. This seems to be a common problem for all non-native subjects.

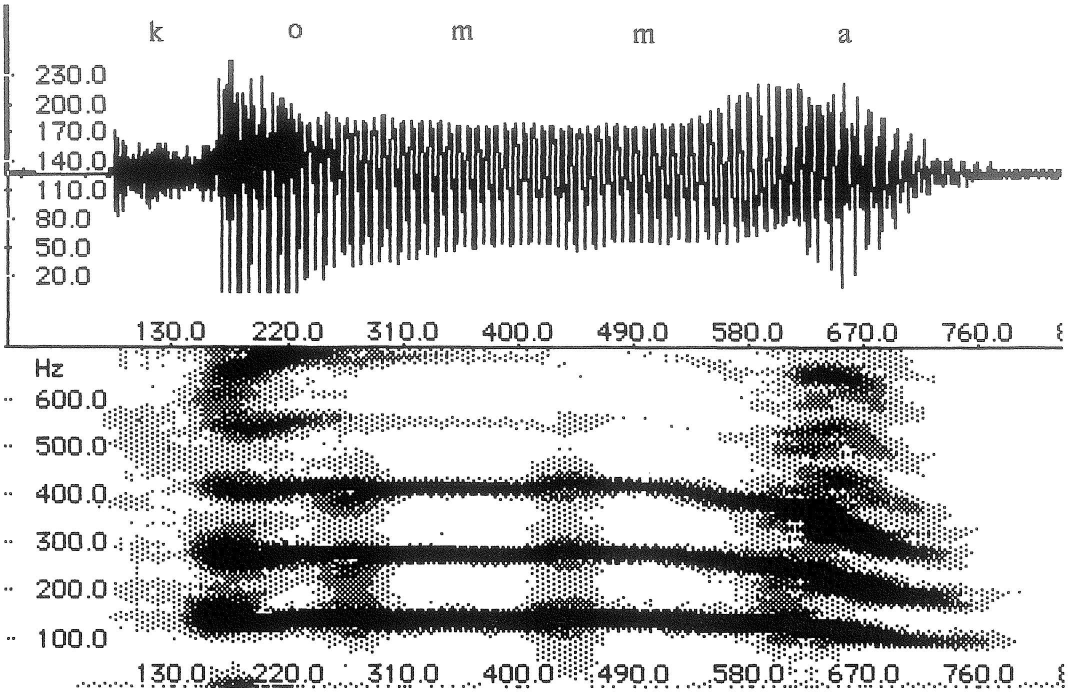


Fig.4.48 [komma]

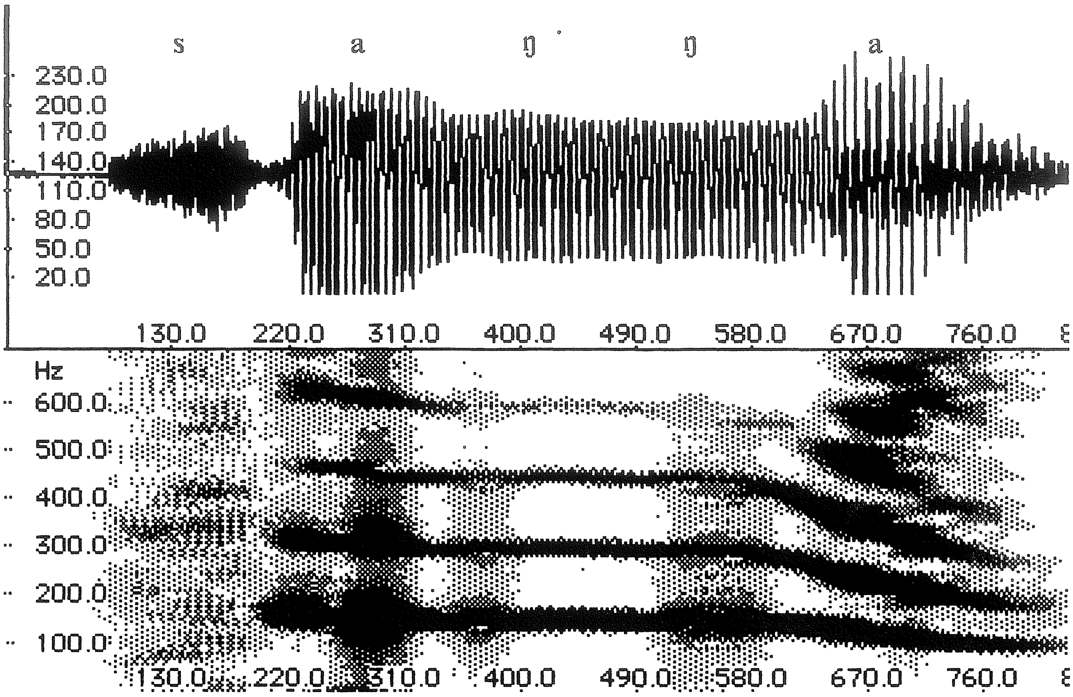


Fig.4.49 [saŋŋa]

4.3.3 Summary

The results obtained from the acoustic observations reveal some typological categories in the IL characteristics of Japanese timing organization by English-speaking learners.

Type 1. Underdifferentiation of the long/short contrast

This type can be subdivided into underdifferentiation caused by (1) a shorter duration of long phonemes, and (2) a longer duration of short phonemes. This study indicates that the latter problem is more prominent in respect of the data provided by the group of beginning learners observed.

Thus the widely accepted view of difficulties in the timing organization of geminate C2 is open to question. The majority of the subjects participating in this study produced geminate C2 with long duration. In fact, there are many examples in which the duration of C2 is clearly too long in relation to that of V1. For example, C2/V1 ratios of rika/rikka/rika was 3.59/6.54/3.10 (LB) and 3.75/5.41/4.58 (LD), while it was only 1.31/2.82/1.13 for NS.

In informal conversations during the recording sessions, some learners said that they interpreted CVCCV sequences as CVC-CV (e.g. [rik-ka]). As supporting evidence for this, a spectrogram of an utterance which shows a clear release of glottal closure between the syllable-final and syllable-initial [k] is presented in Fig.4.50. This is also illustrated as IL strategy 1 in Fig.4.51 shown below. This learner strategy may reflect their underlying knowledge of the occurrence of geminate consonants if the geminate-internal syllable boundary is identified with the English word (or foot) boundary in sequences such as *cattail* and *rock cake* (i.e. CVCCV identified as two words- CVC#CV), and this interpretation may have helped the students who produced long C2. This example demonstrates that the role of the syllable is important for English-speaking learners of Japanese, and in this respect, supports the claim that segmentation unit is determined by L1 (Mehler et al. 1981; Cutler et al. 1983, 1986; Otake et al. 1993). However, it is interesting to note that some learners who applied this strategy, successfully produced long duration for geminate C2, despite using the segmentation unit which was different from that of NS.

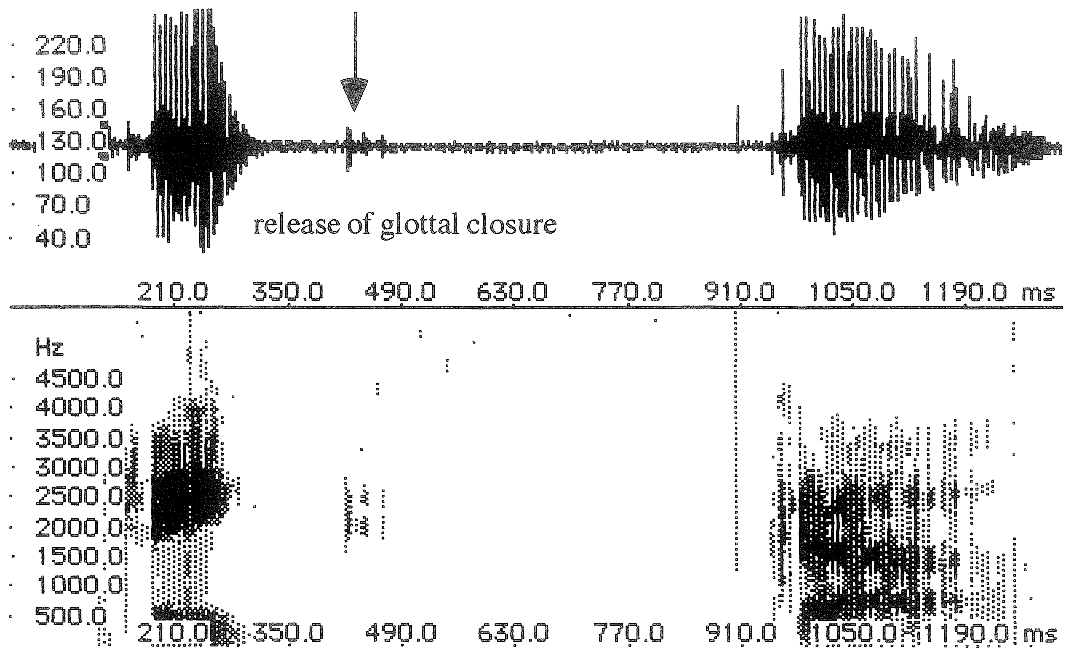


Fig.4.50 re-setting of the [k] in [rikka] (example from a beginner)

On the other hand, the timing control of words with single C2 is a greater problem for many learners which has not been discussed in great detail in the previous literature. It was found that the longer duration of single C2 in CVCV sequences was a common cause of underdifferentiation.

This finding is also important from the viewpoint of linguistic theory. If the timing organization of basic CVCV sequences is governed by 'universal rules' (Chomsky and Halle 1968:295), it is difficult to account for the difficulties that learners showed in the production of CVCV in Japanese. The findings of the present study show that English-speaking learners of Japanese controlled their timing structure in different ways to Japanese NS, and therefore seem to support the claim that Japanese timing structure is influenced by language-specific timing rules. The universal timing-implementation rules, such as the relative duration of VOT for stops, seem to operate across different languages (spoken by NS), and these may be universal.

The duration of closure and VOT varies across different languages. The recent study by Minagawa (1994) found that both the silent closure period and VOT of a single stop are longer in English, Welsh and Korean than in Japanese. Thus, as far as the data investigated in the present study are concerned, the longer duration of single C2 in Japanese CVCV and CVCVV sequences pronounced by English speakers may be considered as a case of negative L1 interference.

From a pedagogical point of view, the results obtained from this study imply that it is important for Japanese language instructors to be aware of students' difficulties in producing single C as well as geminate C, and that exaggerating the duration of geminate C or emphasising isochronous aspects of mora timing in the classroom may not necessarily benefit students (see Chapter 7).

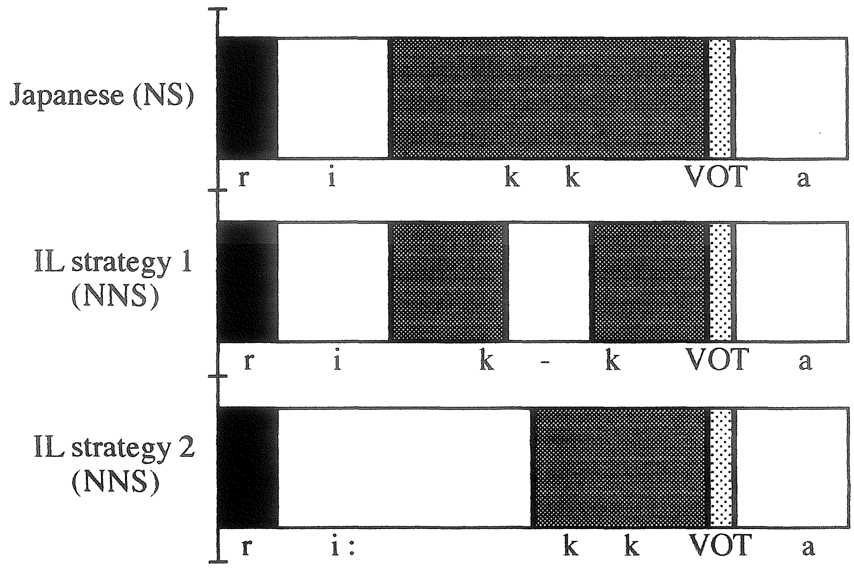


Fig.4.51 Timing compensation due to interference from CVC-phonology

Type 2. Reverse realization of the long/short contrast

The results obtained in this study suggest that the causes of underdifferentiation can be attributed, not just to the shorter duration of long phonemes and the longer duration of single C2, but also to the longer duration of vowels preceding geminate C2.

It was demonstrated that the duration of the V1 was statistically longer before a geminate C2. This was interpreted as the learners' attempt to distinguish the phonological contrast in the C2. Their implementation may reflect certain degrees of interference from CVC-phonology. The reason for this argument was the tendency for learners to manipulate the duration of vowels instead of C2, since no syllable-final C is prolonged in CVC-syllables in English. If we treat CVC as a unit of timing for the learners, a compensatory relationship between the duration of V and C in the first syllable may be illustrated as IL strategy 2 (Fig.4.51).

It was claimed that this strategy was not effective from the viewpoint of NS speech perception. In Chapter 3, it was demonstrated that NS perceptual judgment of geminate C2 can be affected by the V1 duration. Japanese NS use various phonetic cues in their judgment of phonological contrast other than the actual duration of C2. Without reference to these parameters it is impossible to state the appropriate closure duration of consonants. Thus a pedagogical suggestion such as 'make the duration of geminate consonants 3 times as long as the single ones' (Han 1992:126) seems too unidimensional to capture the dynamic reality of durational contrasts in the spoken language.

In the above observations, processes such as negative transfer (interference from English CVC-phonology upon Japanese syllable structure), overexaggeration (of geminate C2 due to the long duration of single C2) and unique syllable modification strategies were found to be operative in shaping the IL timing organization of Japanese.

It is argued that the timing control of CVCV is also difficult for learners, and thus teachers need to introduce practice on Japanese timing control which incorporates such important parameters as total word duration and the duration of the first vowel, as well as the long/short contrast (see Chapter 7 regarding the teaching of timing control).

Perception and production

As mentioned previously, there is a tendency for NS to produce words with short phonemes within the domain of their perceptual category of short phonemes (see 4.2.3). The beginners' results, on the other hand, do not demonstrate such a consistent relationship between speech production and the perception. The following is the main difference from the NS results.

In the beginners' data, the C2/V1 ratios of words with single C2 concerning speech production tend to fall in between the two ratios for speech perception (when the V1 duration is 70 and 140%). This is similar to the NS results of the C2/V1 ratios of words with geminate C2. For example, the C2/V1 ratio of [hone] (1.64) is between HO07N_E (2.77) and HO14N_E (1.32). Also in SAGA series, the C2/V1 ratios of both [saŋa] (1.32) and [saga] (0.69) are between SA07N_A (1.46) and SA14N_A (0.65). This situation also applies to V2/V1 ratios. For example, the V2/V1 ratios of [iso/isso] (2.16/1.95) are between I07SO_ (3.47) and I14SO_ (1.81) in those of perception.

This means that the C2(V2)/V1 ratios for words with short C2(V2) in the learners' speech production may override those of the sound stimuli in which they perceive long C2(V2). The results obtained in this observation suggest that beginners may produce words with short phonemes within the domain of their perception of long phonemes. This situation is different from the Japanese NS who generally produced words with short phonemes within the domain of their perceptual category of short phonemes.

4.4 Australian advanced learners

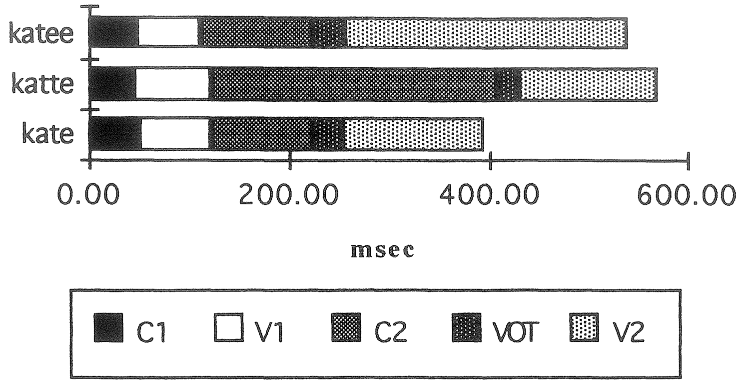
The following section presents the data of advanced speakers. The average values of eight advanced learners are presented, and comparisons are made between the data of NS and beginning learners, which were shown in preceding sections.

4.4.1 Subjects

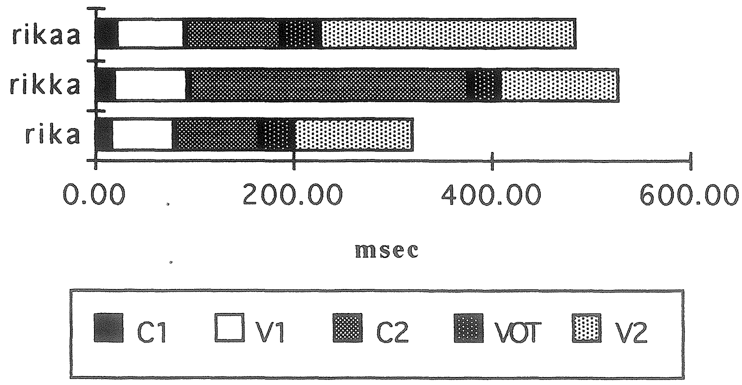
The eight advanced learners participating in this experiment are the same ones who participated in the perception test (Chapter 3). My auditory impression of the advanced learners' speech production is shown in Appendix 4.7.

4.4.2 Results

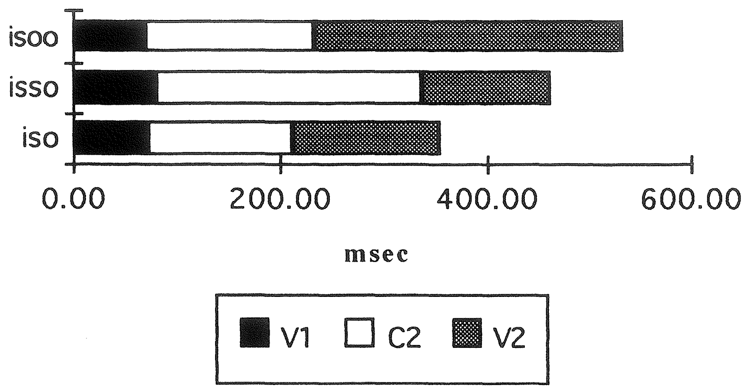
Advanced Average (1)



Advanced Average (2)



Advanced Average (3)



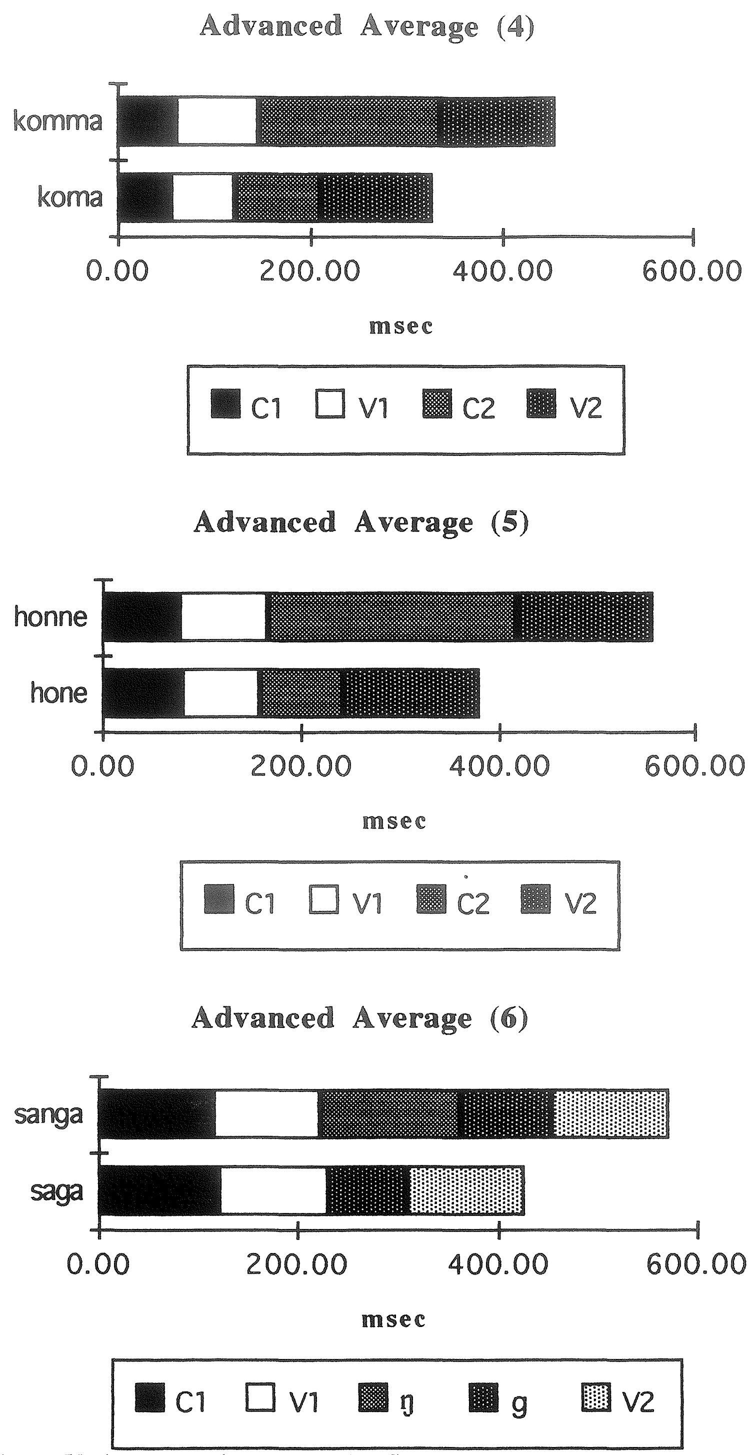


Fig.4.52 Advanced average (1-6)

4.4.2.1 Word duration

The data concerning word duration share similarities with NS: the duration of CVCCV and CVCVV (three-mora expressions) is always longer than that of CVCV (corresponding two-mora expressions. Appendix 4.8 presents the average values for the advanced level learners.

The average ratio of the three-mora expressions is 141% of the corresponding two-mora expressions, and this is similar to the ratio for NS (143%).

In the case of NS, all subjects pronounced the SAGA series as [saga/saŋga] and the ratio was 100:129. Most advanced subjects pronounced the SAGA series as [saga/saŋga], and the ratio was 100:135. There were two advanced learners who pronounced the examples with velar nasal [ŋ] for [saŋa] (BL and TN) and three for [saŋŋa] (AA, BL and TN), and their average showed the smallest ratio of differentiation with respect to word duration (100:111). Despite the small rate of differentiation, there was no problem at the level of auditory impression concerning the pronunciation of these advanced learners, which is probably due to their accurate assignment of tone (Appendix 4.7). This is an interesting contrast with the beginning subjects' data in which the durational contrast was greater, but the auditory impression was inappropriate due to incorrect assignment of tone.

Unlike the case of the beginning learners, word duration in advanced learners' speech does not show a reverse realization of word duration such as two-mora expressions being longer than three-mora expressions as mentioned above. It seems that word-level timing control, which is closely associated with the number of morae in the word, has been acquired by advanced speakers of Japanese. Beginning learners did not show the same level of performance.

SERIES	RATIO
kate/katte/katee	100:144:137
rika/rikka/rikaa	100:164:151
iso/isso/isoo	100:130:150
koma/komma	100:139
hone/honne	100:149
saŋa/saŋŋa	100:111
saga/sanga	100:135

Table 4.10 Word duration (ratio)

4.4.2.2 Long/short contrast

A clear phonetic contrast can be observed in the phonological long/short opposition. It is interesting to note that the contrast of C2 in the ISO series (55/100/63) was underdifferentiated, while that in the RIKA series (31/100/34) was overdifferentiated. The NS ratios of these groups were 48/100/49 and 39/100/37 respectively.

	C2	V2
1 kate	36	49
2 katte	100	49
3 katee	40	100
4 rika	31	48
5 rikka	100	47
6 rikaa	34	100
7 iso	55	48
8 isso	100	42
9 isoo	63	100
10 koma	47	
11 komma	100	
12 hone	34	
13 honne	100	
14 saŋa	41	
15 saŋŋa	100	
16 saga	34	
17 sanga	100	

Table 4.11 Long/short contrast (ratio)

4.4.2.3 C2/V1 and V2/V1 ratios

The C2/V1 ratio was about twice as large for words with a long obstruent C2 as for those with a short C2 in the NS data. The overall tendency of the advanced speakers' data is towards NS rather than beginning learners; the ratios with long phonemes are greater than those with short phonemes. There is a tendency for the C2/V1 and V2/V1 ratios of [rikka] and [rikaa] to be greater than NS, being 3.79 and 3.75 with advanced speakers and 2.82 and 2.93 for NS. This seems to imply a tendency for advanced speakers to produce longer C2 and V2 in relation to the duration of V1, and can thus be regarded as overexaggerating the long C2 and V2.

In the data of beginning learners, it was found that the ratios of C2/V1 in some examples were not significantly different between single and geminate C2, such as the ISO series (2.11/2.41/2.22). In the data obtained from advanced learners, no such example was observed.

	C2/V1	V2/V1
1 kate	1.41	1.92
2 katte	3.74	1.82
3 katee	1.82	4.56
4 rika	1.34	1.91
5 rikka	3.79	1.63
6 rikaa	1.39	3.73
7 iso	1.93	1.94
8 isso	3.24	1.58
9 isoo	2.32	4.25
10 koma	1.30	1.66
11 komma	2.21	1.38
12 hone	1.06	1.73
13 honne	2.79	1.55
14 saŋa	0.66	1.07
15 saŋŋa	1.99	1.12
16 saga	0.72	1.07
17 sanga	2.15	1.12

Table 4.12 C2/V1 and V2/V1 ratios

4.4.2.4 VOT

The above table shows the VOT of advanced learners.

1) The average VOT is always greater than that in the NS data, but smaller than in the beginning learners' data.

2) The standard deviation is always greater than that in the NS data, but smaller than in the beginning learners' data.

3) The difference in VOT between [t] and [k] is statistically significant (t=-2.317, p=.0227). This is also the case in the corresponding geminate C2 (t=-3.004, p=.0043).

4) The difference between VOT for single and geminate C2 is significant. The results of the t-tests are [t/tt] (t=3.96, p=.0002) and [k/kk] (t=2.515, p=.0142). The VOT of the geminate C2 is significantly smaller than that of the single C2 in advanced learners' data. This agrees with the universal hypothesis proposed by Han (1992), that there is a shorter duration of VOT for geminate stops than for single stops, but this tendency was not observed in the NS speech production.

	t	k	t-test
C	30.24 msec (10.48)	35.01 msec (9.58)	t=-2.317, p=.0227
CC	20.75 msec (7.50)	28.73 msec (10.65)	t=-3.004, p=.0043
t-test	t=3.96, p=.0002	t=2.515 p=.0142	

Table 4.13 VOT

The observation of VOT amongst NS, beginning and advanced learners provide clear evidence for phonetic approximation (Flege 1980). The longer VOT of beginning learners showed interference from English, but advanced learners had approximated a VOT which was intermediate between English and Japanese. As the proficiency level of the learners increased, the VOT values became smaller.

4.4.2.5 Pitch assignment

Some advanced learners showed errors in the assignment of pitch for words with a nasal C2, such as [honne] with the first CVC-syllable (i.e. [hon]) assigned an L tone and the second CV-syllable (i.e. [ne]) assigned an H tone, similar to the common errors seen in the beginning learners' data. This was found in the speech production of *honne* of four speakers, CI, RA, RW and SF. Also, RA, RW and SB showed errors in the pitch assignment of *sanga*.

Most errors in pitch assignment occurred in words with a nasal C2. These results show that pitch assignment is a problem even for advanced learners of Japanese, and the problem occurs more often with words with a nasal C2 than with those with an obstruent C2.

4.4.3 Summary

The previous section reported the data of advanced learners of Japanese. Generally speaking, advanced learners show more similarities with NS than with beginning learners, such as clear differentiation between two- and three-mora word duration, and between long/short contrast in C2 and V2.

It was found that advanced learners show a tendency towards overexaggeration of a geminate stop C2. On the other hand, the fricative C2 was underdifferentiated. In advanced learners' results, the duration of a geminate fricative C2 is less than twice the duration of a single fricative C2. Thus in advanced learners' speech production, it cannot be said that there is any phonetic evidence for the mora as a larger unit than a single C2, because a combination of a single C2 and a moraic consonant is no longer than the duration of two single C2.

Universal characteristics of VOT were found in the advanced learners' data: that is, a longer VOT is associated with a velar stop than with an alveolar stop. This is similar to speech produced by NS. Generally speaking, the average duration of VOT of advanced learners was somewhere between that of beginners and NS. Thus it can be said that VOT becomes smaller as learners' proficiency level increases, but not to the same extent as NS. This may be the limitation of phonetic approximation towards NS norms by advanced learners of Japanese.

Perception and production

Advanced learners' results demonstrate that there are some examples in which C2/V1 ratios (with geminate C2) obtained for production fall in between the two ratios of perception (when the V1 duration is 70 and 140%). The number of occurrences, however, is less than those found in the NS results. Instead, there is a tendency for the ratios obtained for production to override those obtained for perception in advanced learners' data. For example, the C2/V1 ratio of [katte] (with geminate C2) in the data of production (3.74) is bigger than KA07T_E (3.62) and KA14T_E (2.00) in those of perception. This tendency is also found in V2/V1 ratios. The C2/V1 ratio of [katee] (with long V2) in the data of production (4.56) is bigger than KA07TE_ (4.45) and KA14TE_ (2.39) in those of perception. This problem is observed in [katte], [katee], [isso], [isoo], [honne], [saŋŋa] and [saŋga].

The above results imply that advanced learners tend to produce longer duration for geminate C2 at the level of speech production than the duration of C2 perceived as geminate at the level of speech perception. One of the main causes of this tendency is overexaggeration of geminate C2 and long V2 by advanced learners. In this respect, it is not possible to say that advanced learners' results share similarity with the NS ratios.

As far as the ratios for single C2 and short V2 are concerned, the production data of advanced learners show that they are generally smaller than those obtained from their perception data. For example, the C2/V1 ratios of [kate/katee] (1.41/1.82) in the data of production are smaller than KA07T_E (3.62) and KA14T_E (2.00) in those of perception. Similarly, the V2/V1 ratios of [kate/katte] (1.92/1.82) are smaller than KA07TE_ (4.45) and KA14TE_ (2.39). Thus it can be said that advanced learners produce words with short phonemes within the domain of their perception of short phonemes, and this situation is similar to the NS results.

4.5 Discussion

1. It was found that the beginning learners apply various strategies to achieve word-level and phonemic-level phonological distinctions of Japanese. Various typologies of learner strategies were listed in 4.3.3.

2. The timing organization of advanced learners generally shared more similarities with that of NS than that of beginning learners (see 4.4.1).

3. Universal principles concerning VOT which is associated with the place of articulation operated in advanced learners' speech (as it did in NS speech), but not in beginning learners' speech (see 4.4.2.4).

4. Phonetic approximation was found in advanced learners' speech (VOT) as described in 4.4.2.4.

5. L1 interference is more prominent in beginning learners' speech (e.g. syllable structure) than in advanced learners' speech. Advanced learners, on the other hand, show a tendency towards overexaggeration.

6. There is a tendency for the tone to be assigned to a syllable rather than to a mora in the production of words with a nasal C2. This tendency is observed in both beginning and advanced learners' data.

These results are summarised in the following table which illustrates the characteristic features of the speech of beginning and advanced learners (O= similar to NS; X= dissimilar to NS) when NS is regarded as the target:

	BEGINNER	ADVANCED
word level distinction	X	O
long/short contrast	X	O
C2/V1	X	O
pitch assignment	X	O (obstruent C2) X (nasal C2)
VOT	X	O (t<k)

Table 4.14 Similarities shared with NS results (production)

4.6 Conclusion

We can now relate these findings to the problems involved in the production of durational contrast in Japanese. Toda (1994) reported that beginning learners of Japanese tend to have common problems with the production of durational contrast such as:

1) overly long duration of the single C2, and

2) overexaggeration of the long C2

The findings obtained through the observation of learners' speech production in this chapter seem to be closely related to the problems which occur at the level of speech perception discussed in Chapter 3. It has been demonstrated that there is a perceptual domain in which a sound is perceived to be short by learners while the same sound is heard as long by NS. The fact that the threshold values of beginning learners are generally higher than those of NS seems to account for learners allowing a longer duration for the production of a short phoneme. It is possible to assume a situation in which learners produce a short phoneme, according to their perceptual domain, but where this would be considered long by the NS because it may fall into the NS perceptual domain of a long phoneme. In order to produce a short/long contrast, the duration of a long phoneme needs to be overexaggerated since the duration of a single phoneme is already too long.

The interaction between speech perception and production is very complex (and cannot be described in simple terms) and in this dissertation, it is not my intention to present a theory regarding their interaction. From the viewpoint of teaching timing control in Japanese (Chapter 7), however, it is important to know where problems exist and what kind of problems they are. Through acoustically-based observations, problems at the levels of both production and perception (i.e. inaccurate timing control and perceptual targets) were found, and we have discussed the ways in which they are related. In conclusion, it can be said that one of the goals for learners who wish to obtain native-like performance in Japanese should be to try to lower the threshold boundary for both production and perception to the same level as NS. This seems to be an important objective for learners to achieve if they are to improve their timing control.

Chapter 5 Longitudinal study- Perception

5.1 Subjects

The four subjects who participated in the longitudinal experiment are employed at the Department of Foreign Affairs and Trade (DFAT) and undertook intensive Japanese language training for one academic year (February - November 1993) at the Canberra Institute of Technology before their posting to Tokyo. They had one female Australian and two native Japanese teachers (one male and one female) throughout the year. There were four contact hours every morning and they were expected to do three hours of self-study every afternoon.

PK, DS (male) and AP (female) are in their mid-twenties, and have Australian backgrounds. They speak Australian English as their L1. PV (male) is in his early forties, and has a Kenyan background. He was born in Kenya, and migrated to Australia with his family when he was young. He also speaks English as his L1. None of the subjects suffered from hearing or speaking disorders. All four subjects started learning Japanese as beginners. DS and PK had studied Japanese for two months prior to this intensive course. None of the subjects had ever lived in Japan, except for DS who had visited Japan for two months as an exchange student in 1985.

The aim of the participants in this intensive language training course was to achieve an S2/R2 proficiency rating (or higher) in Japanese language as assessed by the Language Training Unit, DFAT.[35]

The subjects were asked to participate in perception and production tests which were the same experiments described in Chapters 3 and 4. The tests were given three times throughout the year: February-March, July-August and November 1993. All the subjects were aware of phonological durational contrast in Japanese before the first experiment was conducted.

---

[35] S=Speaking; R=Reading/Writing; 5=native, 4=near native; 3=advanced; 2=intermediate; 1=beginner as mentioned previously (Chapter 3).

5.2 Procedure

The sound stimuli and procedure of this experiment are the same as for the perception test which was described in Chapter 3.

5.3 First experiment

In this section, the DFAT learners' results obtained in the first experiment (DFAT1) are discussed.[36] The first experiment was conducted approximately one month after the commencement of the intensive course (late February - March 1993).

Vowels

CV07CV_	KA07TE_	RI07KA_	I07SO_
A.S	165%	190%	160%
S.D	32%	41%	40%
D.S	185%	210%	183%
S.D	40%	43%	45%
CV14CV_	KA14TE_	RI14KA_	I14SO_
A.S	173%	193%	160%
S.D	28%	26%	28%
D.S	188%	208%	178%
S.D	37%	37%	39%

Obstruents

CV07C_V	KA07T_E	RI07K_A	I07S_O
A.S	170%	163%	188%
S.D	51%	38%	61%
D.S	230%	225%	218%
S.D	41%	30%	38%
CV14C_V	KA14T_E	RI14K_A	I14S_O
A.S	173%	170%	165%
S.D	38%	39%	53%
D.S	228%	228%	205%
S.D	48%	40%	33%

[36] Similarly, 'DFAT2' and 'DFAT3' refer to the learners' results obtained in the second and third experiments respectively. For individual learners, 'AP1' means the learner's results obtained in the first experiment.

Nasals

CV07N_V	KO07M_A	HO07N_E	SA07N_A
A.S	75%	78%	65%
S.D	12%	18%	28%
D.S	91%	85%	76%
S.D	10%	14%	16%
CV14N_V	KO14M_A	HO14N_E	SA14N_A
A.S	83%	65%	73%
S.D	25%	15%	29%
D.S	93%	84%	77%
S.D	18%	16%	18%

Tables 5.1 Average Values - DFAT1

A.S: ascending series

D.S: descending series

S.D: standard deviation

CV07CV_	KA07TE_	RI07KA_	I07SO_
T.V	175%	200%	171%
A.D	268.45	263.00	265.10
CV14CV_	KA14TE_	RI14KA_	I14SO_
T.V	180%	200%	169%
A.D	276.12	263.00	261.23
CV07C_V	KA07T_E	RI07K_A	I07S_O
T.V	200%	194%	203%
A.D	229.40	243.54	214.85
CV14C_V	KA14T_E	RI14K_A	I14S_O
T.V	200%	199%	185%
A.D	229.40	249.83	196.29
CV07N_V	KO07M_A	HO07N_E	SA07N_A
T.V	83%	81%	71%
A.D	176.72	147.71	115.33
CV14N_V	KO14M_A	HO14N_E	SA14N_A
T.V	88%	74%	75%
A.D	186.03	135.21	122.48

Table 5.2 Average Threshold Values - DFAT1

T.V: Average Threshold Values of ascending and descending series (%)

A.D: Absolute Duration (msec)

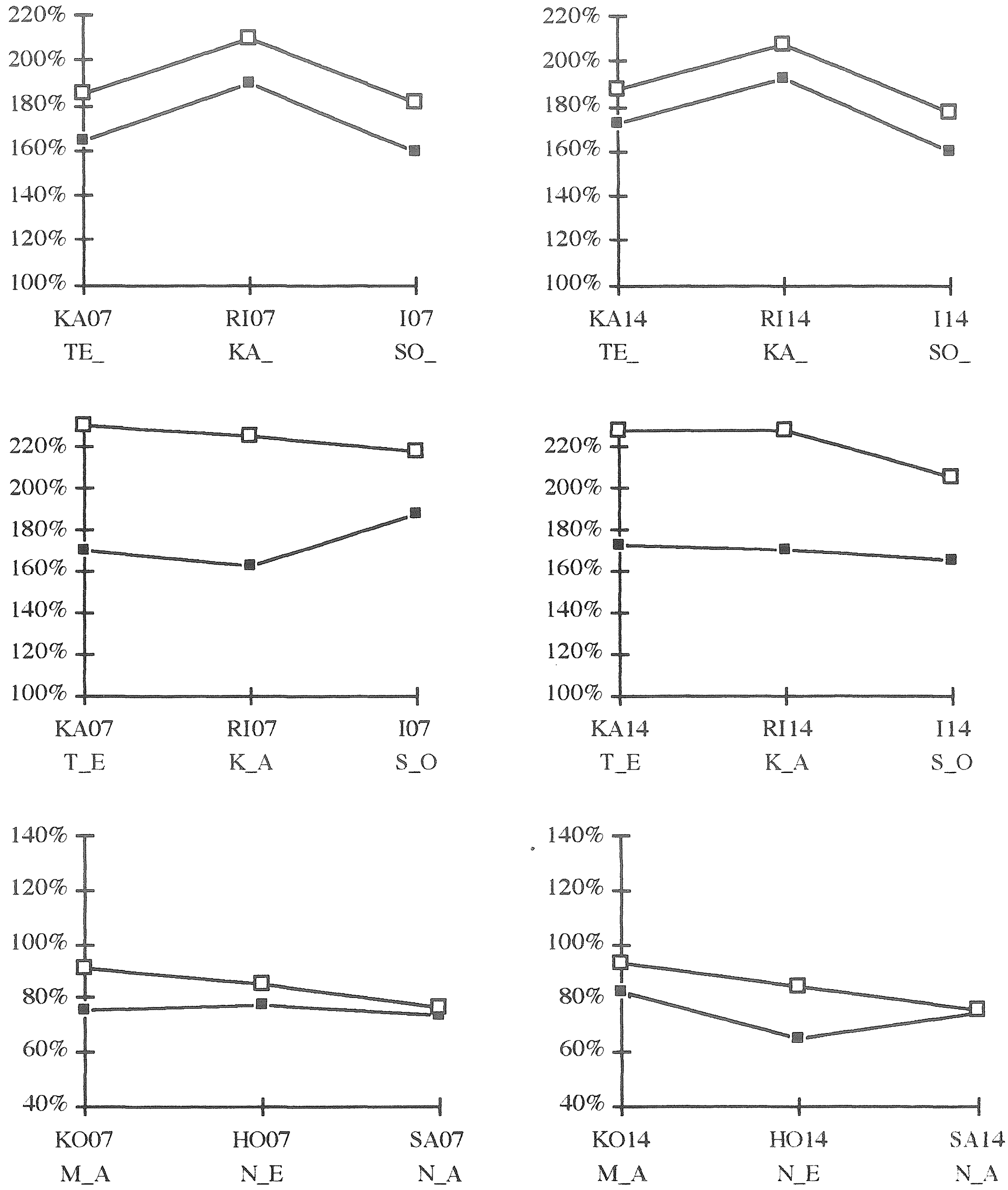


Fig.5.1 Graphs for perception by DFAT1 learners[37]
Ascending (black), Descending (white)

The results of the perception test for the DFAT learners are shown above. They bear remarkable similarities to the results obtained from ANU beginners which were presented in Chapter 3.

[37] Question marks did not appear in the DFAT1 results as they did in ANU results, although the same instruction was given for the use of a question mark before the experiment.

The main results are:

1. The average threshold values are generally greater than those of NS for both C2 and V2. Based on this result, we can say that the segmental duration needs to be longer to be perceived as long C2 and V2 by DFAT1 beginning learners (as well as ANU learners) than it does by NS.

2. Descending series almost always show greater values than the corresponding ascending series (Appendix 5.1). This is similar to the ANU beginners' data, and may be interpreted as evidence of the learners' different perceptual mechanisms from those of NS.

3. The distance between the ascending and descending series is generally greater than in the NS data. The differences between the two series were statistically significant in the following examples: KA07T_E, KA14T_E, RI07K_A, RI14K_A, HO14N_E and KO07M_A. This seems to imply that the degree of uncertainty (i.e. 'grey area') is greater for learners than for NS. These results share similarity with ANU1 results.

In comparison with ANU1, DFAT1 shows greater distance between the ascending and descending series for obstruent C2. This result seems to imply that a greater amount of uncertainty is involved in the perceptual judgment of C2 by DFAT1 learners. On the other hand, the performance of DFAT1 was better than that of ANU1 with regard to their perceptual judgment of V2, since the distance between ascending and descending series was smaller, and differences were not statistically significant (Appendix 5.1).

4. In the DFAT1 data, none of the examples showed significantly greater values when the duration of V1 was expanded to 140% in comparison with when it was shortened to 70% (Appendix 5.2). Thus it is not possible to say that the perceptual boundaries of DFAT1 learners are affected by the V1 duration in the same way as those of NS.

Overall tendencies of the DFAT1 data thus share a number of similarities with beginning learners, and dissimilarities with NS.

5.4 Second experiment

In this section, the results obtained in the second experiment (DFAT2) are presented and the differences from those obtained in the first experiment are discussed.

Vowels

CV07CV_	KA07TE_	RI07KA_	I07SO_
A.S	173%	208%	155%
S.D	34%	40%	28%
D.S	200%	200%	175%
S.D	21%	28%	33%
CV14CV_	KA14TE_	RI14KA_	I14SO_
A.S	170%	193%	158%
S.D	21%	32%	23%
D.S	188%	220%	183%
S.D	38%	15%	20%

Obstruents

CV07C_V	KA07T_E	RI07K_A	I07S_O
A.S	183%	173%	198%
S.D	55%	32%	33%
D.S	220%	210%	198%
S.D	30%	24%	25%
CV14C_V	KA14T_E	RI14K_A	I14S_O
A.S	173%	203%	188%
S.D	37%	60%	34%
D.S	198%	215%	178%
S.D	31%	41%	41%

Nasals

CV07N_V	KO07M_A	HO07N_E	SA07N_A
A.S	78%	91%	86%
S.D	16%	23%	24%
D.S	90%	74%	65%
S.D	15%	12%	24%
CV14N_V	KO14M_A	HO14N_E	SA14N_A
A.S	78%	79%	61%
S.D	14%	17%	20%
D.S	91%	78%	57%
S.D	15%	12%	15%

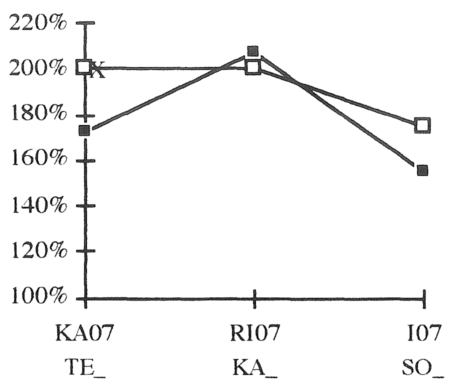
Tables 5.3 Average Values - DFAT2

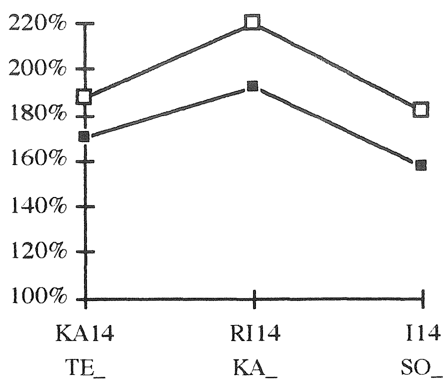
CV07CV_	KA07TE_	RI07KA_	I07SO_
T.V	186%	204%	165%
A.D	285.71	267.93	255.42
CV14CV_	KA14TE_	RI14KA_	I14SO_
T.V	179%	206%	170%
A.D	274.20	271.22	263.16
CV07C_V	KA07T_E	RI07K_A	I07S_O
T.V	201%	191%	198%
A.D	230.83	240.40	209.55
CV14C_V	KA14T_E	RI14K_A	I14S_O
T.V	185%	209%	183%
A.D	212.20	262.40	193.63
CV07N_V	KO07M_A	HO07N_E	SA07N_A
T.V	84%	82%	76%
A.D	178.05	148.85	123.50
CV14N_V	KO14M_A	HO14N_E	SA14N_A
T.V	84%	78%	59%
A.D	179.38	142.03	96.96

Table 5.4 Average Threshold Values - DFAT2

T.V: Average Threshold Values of ascending and descending series (%)

A.D: Absolute Duration (msec)





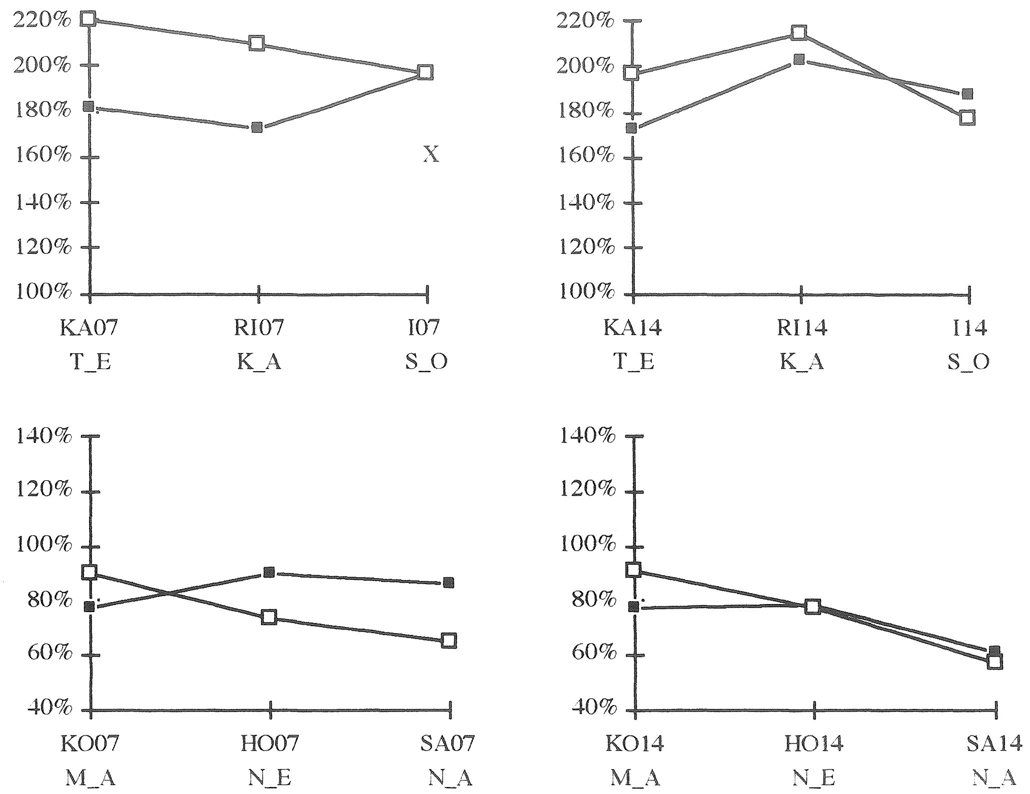


Fig.5.2 Graphs for perception by DFAT2 learners[38]

Ascending (black), Descending (white)

1. The average threshold values obtained in DFAT2 are generally greater than those obtained in NS.

2. The order of the ascending and descending series has changed in some examples. Unlike DFAT1, the values for the descending series are not consistently greater than those for the corresponding ascending series.

3. The number of words in which the differences between the two series are statistically significant has decreased. (See Appendix 5.4 for t-test results.) There are only three such examples (by comparison with six in DFAT1), namely: RI14KA_, I14SO_, RI07K_A. If we assume that the distance between these two series indicates the amount of uncertainty involved in learners' perceptual judgments, the fact that fewer examples show significant difference suggests that the amount of uncertainty

[38] X in the graphs shows where learners indicated a question mark before JND. Appendix 5.3 shows the values for which question marks were indicated by DFAT2 learners.

has decreased. The distance between the ascending and descending series has become notably smaller, especially with regard to obstruent C2.

4. V1 duration does not seem to have any effect upon perceptual judgment in DFAT2. There is no case in which the percentage increase in duration required for a JND is significantly greater when the V1 duration is expanded to 140% than when it is shortened to 70% (Appendix 5.5). This is similar to the results obtained from DFAT1.

DFAT2 may be summarised as sharing some similarities with DFAT1; the average threshold values are generally greater than those of NS, and there is also little influence of V1 duration upon perceptual judgment. DFAT2 differs from DFAT1, however, in that the distance between the ascending and descending series has become smaller, and in fact there are less examples in which the differences between the two series are statistically significant. These results seem to suggest that the learners' perceptual targets are more established than for DFAT1, in the sense that they demonstrate a smaller amount of uncertainty in their judgment. This does not mean, however, that the learners' perceptual boundary itself has become similar to that of NS, since the threshold values are still greater than those of NS.

5.5 Third experiment

In this section, the results obtained from the third experiment (DFAT3) are presented and the differences from those obtained from the first and second experiment are discussed.

This is the final test of the longitudinal experiment. The four participants who had been undertaking intensive Japanese language training were about to finish their course and leave for Japan at the beginning of the following year. All participants achieved their goal in terms of their proficiency rating (see 5.1).

Vowels

CV07CV_	KA07TE_	RI07KA_	I07SO_
A.S	190%	193%	185%
S.D	35%	35%	35%
D.S	183%	213%	165%
S.D	20%	26%	30%
CV14CV_	KA14TE_	RI14KA_	I14SO_
A.S	193%	198%	175%
S.D	21%	47%	28%
D.S	198%	200%	190%
S.D	27%	24%	43%

Obstruents

CV07C_V	KA07T_E	RI07K_A	I07S_O
A.S	183%	190%	185%
S.D	54%	37%	42%
D.S	195%	195%	200%
S.D	35%	50%	37%
CV14C_V	KA14T_E	RI14K_A	I14S_O
A.S	200%	203%	200%
S.D	32%	46%	47%
D.S	218%	213%	198%
S.D	46%	34%	29%

Nasals

CV07N_V	KO07M_A	HO07N_E	SA07N_A
A.S	79%	80%	80%
S.D	14%	12%	16%
D.S	81%	85%	79%
S.D	16%	18%	25%
CV14N_V	KO14M_A	HO14N_E	SA14N_A
A.S	85%	79%	73%
S.D	25%	15%	13%
D.S	85%	73%	73%
S.D	11%	21%	18%

Tables 5.5 Average Values - DFAT3

A.S: ascending series

D.S: descending series

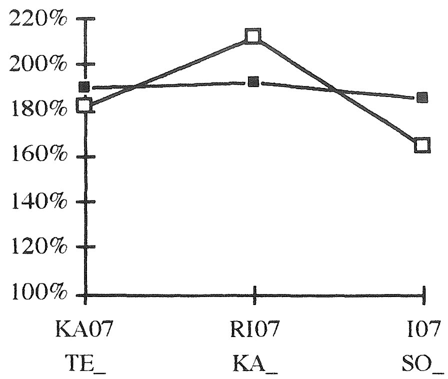
S.D: standard deviation

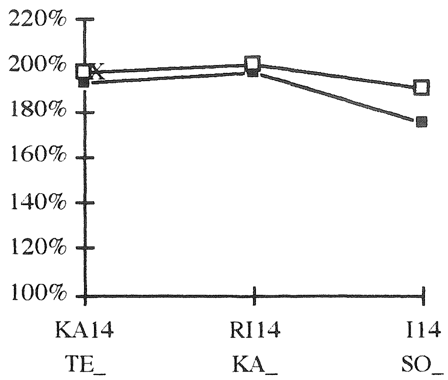
CV07CV_	KA07TE_	RI07KA_	I07SO_
T.V	186%	203%	175%
A.D	285.71	266.29	270.90
CV14CV_	KA14TE_	RI14KA_	I14SO_
T.V	195%	199%	183%
A.D	299.13	261.36	282.51
CV07C_V	KA07T_E	RI07K_A	I07S_O
T.V	189%	193%	193%
A.D	216.50	241.97	204.24
CV14C_V	KA14T_E	RI14K_A	I14S_O
T.V	209%	208%	199%
A.D	239.44	260.83	210.87
CV07N_V	KO07M_A	HO07N_E	SA07N_A
T.V	80%	83%	79%
A.D	170.08	149.99	129.62
CV14N_V	KO14M_A	HO14N_E	SA14N_A
T.V	85%	76%	73%
A.D	180.71	137.49	118.39

Table 5.6 Average Threshold Values - DFAT3

T.V: Average Threshold Values of ascending and descending series (%)

A.D: Absolute Duration (msec)





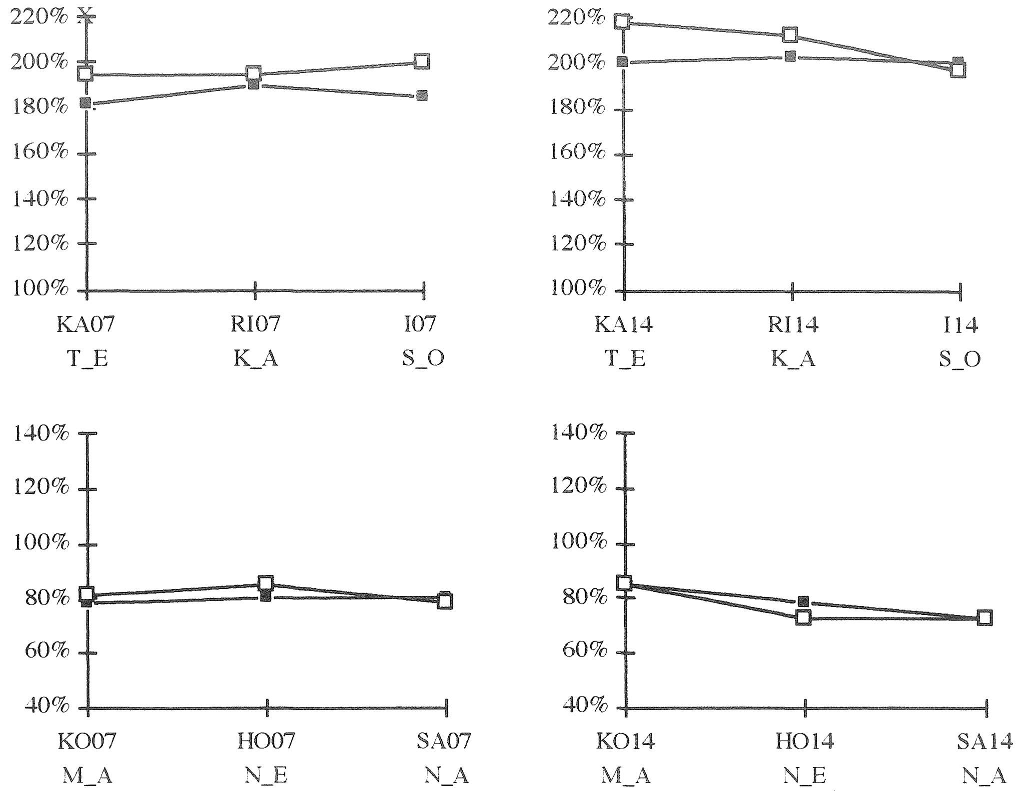


Fig.5.3 Graphs for perception by DFAT3 learners[39]
Ascending (black), Descending (white)

1. The main and crucial difference from the NS data is that most average threshold values obtained in DFAT3 are still higher, and in fact, they are not significantly different from the results obtained in DFAT1 (see Appendix 5.7 for dependent t-test results).

Throughout the three experiments conducted in the academic year, no significant change in the overall average threshold values was observed. If we assumed that the learners' perceptual target would have gradually become closer to the NS norm as their level of proficiency increased towards the end of one year's intensive training, the average threshold values would have become smaller. The DFAT3 results, however, demonstrated no such tendency. In other words, the learners still have an inaccurate perceptual target, and the duration of C2 and V2 still needs to be longer than NS in order to be perceived as a long phoneme.

[39] Appendix 5.6 shows the values for which question marks were indicated by DFAT3 learners.

2. In DFAT3, there is no consistent pattern in the order of the ascending and descending series. In DFAT1, there was a tendency for the descending series to be greater than the corresponding ascending series. This tendency has disappeared in DFAT3.

3. The distance between the ascending and descending series has become much smaller, and in fact none of the words demonstrated statistical significance in DFAT3 (see Appendix 5.8 for t-test results). It is interesting to note that the number of words which demonstrated a statistical difference between the two series has gradually decreased over the year: in DFAT1, six words were found to be significant, in DFAT2 there were three, and finally in DFAT3 there were none. The fact that the distance between the ascending and descending series has gradually become non-significant implies that the amount of uncertainty in their perceptual judgment has decreased over the year.

4. The relationship between V1 duration and the threshold values of C2 and V2 is still undefined. There are no examples in which the percentage increase in duration for a JND is significantly greater when V1 duration is expanded than when it is shortened (Appendix 5.9). This situation is the same throughout DFAT1-3. Therefore, as far as the present data are concerned, it cannot be said that learnability of a perceptual target can be extended to the perception of C2 and V2 as a function of V1 duration.

5.6 Discussion

The most crucial fact found from this longitudinal observation is that the average threshold values remained greater than those of NS throughout the year. After one year's intensive language training, the average threshold values obtained in DFAT3 did not become significantly smaller than DFAT1 (Appendix 5.7). On the other hand, while the distance between the ascending and descending series was significant in DFAT1 (Appendix 5.1), it gradually became smaller and finally became non-significant for all examples in DFAT3 (Appendix 5.8).

One possible interpretation of the above finding is that the learners had not established clear perceptual targets at the beginning of the year, and thus the influence from errors of expectation may have been greater (also see Chapter 3). The learners gradually developed perceptual boundaries in which they related short/long contrasts, and therefore the amount of uncertainty involved in their perceptual judgment decreased. The interesting fact is that, although the extent of the grey area decreased after one year, the actual threshold values which can be regarded as perceptual boundaries did not show a significant change. In other words, the learners established perceptual targets which were inaccurate, and their perceptual judgment became more and more accurate in the context of their inaccurate perceptual targets.

As mentioned above, the inaccurate perceptual targets learners have established involve a longer duration for the perception of a long phoneme compared to NS. This means that there is a perceptual domain in which the same stimulus is identified as short by learners but as long by NS. This discrepancy may be regarded as a cause of perceptual problems concerning long/short phonemic contrast for learners of Japanese.

If this is the case, the question is: when will the learners' perceptual boundary change to something which is closer to that of NS? It was demonstrated in Chapter 3 that advanced speakers of Japanese have a similar perceptual target to NS. The threshold values which were obtained from advanced learners' were significantly smaller than those obtained from beginning learners, and this suggests that the learners' perceptual boundary has been modified somewhere between the intermediate level to near-native advanced level in the process of language acquisition. Such perceptual development was not observed during the one year's intensive course. It would appear that the perceptual norms of NS are something which take a long time for learners to acquire. From the results described in this longitudinal study, it seems that beginning learners cannot be expected to acquire NS models by themselves without appropriate instruction, since they may establish an inaccurate perceptual target and their judgment may become more and more accurate only towards this inaccurate perceptual target.

It was also found that the perceptual boundary was not affected by V1 duration. Threshold values were not significantly greater when the duration of V1 was expanded to 140% than when it was shortened to 70%. This situation applied throughout DFAT1-3.

5.7 Conclusion

Longitudinal observation was conducted during one year with a group of DFAT learners to observe the details of the process of acquisition. It was found that the degree of uncertainty involved in their perceptual judgment decreased over one year.

The crucial problem was that the learners established an inaccurate perceptual target, which was not modified throughout the year. It was interesting to observe the way in which their perceptual judgment gradually became more accurate in relation to their inaccurate perceptual target.

Chapter 6 Longitudinal study- Production

This chapter reports the results for speech production obtained from the longitudinal study with DFAT learners. The main aim of the chapter is to observe the various strategies which learners apply at different stages of language acquisition, and the results are therefore presented separately for each learner.

6.1 Subjects

The subjects were those who participated in the longitudinal experiment for perception described in Chapter 5.

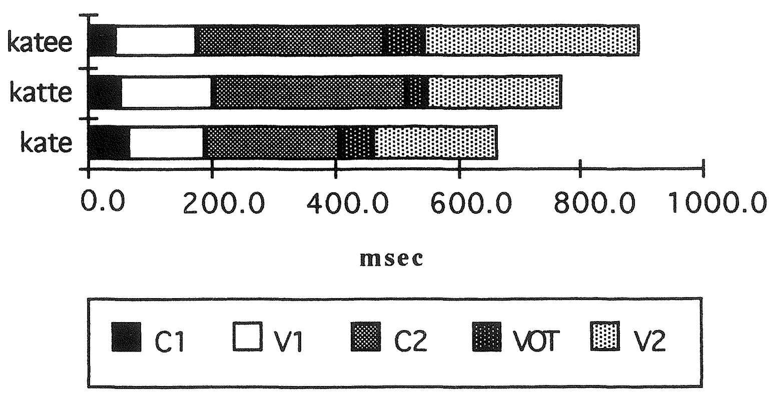
6.2 Procedure

The procedure for this experiment was the same as the perception test which was described in Chapter 4.

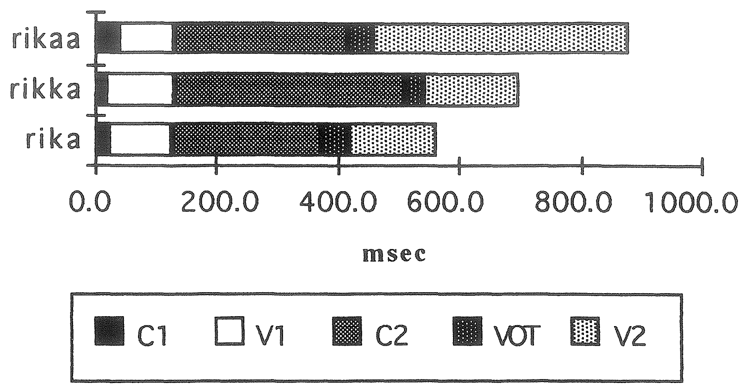
6.3 First experiment

The results obtained from the first experiment (conducted in late February - March 1993) are presented in the following section. The results obtained in this experiment bear a number of similarities with those of the ANU beginning-level learners discussed in Chapter 4. My auditory impression of DFAT1 is shown in Appendix 6.1.

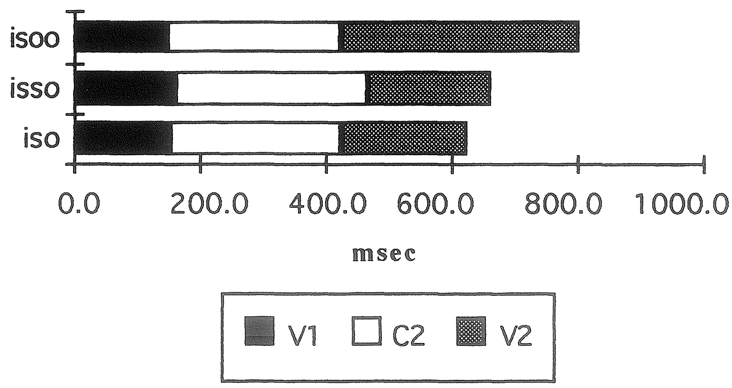
AP1(1)



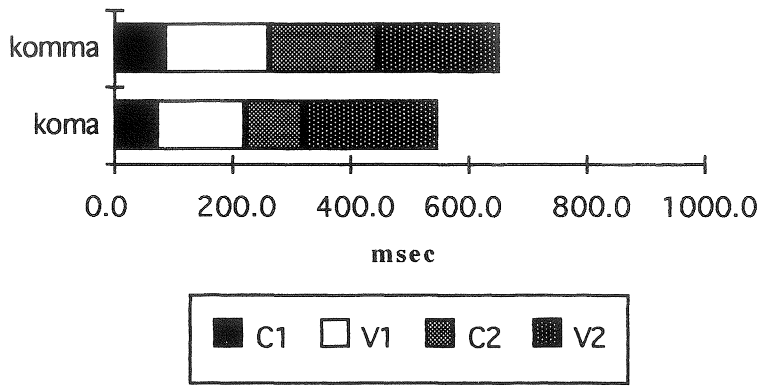
AP1(2)



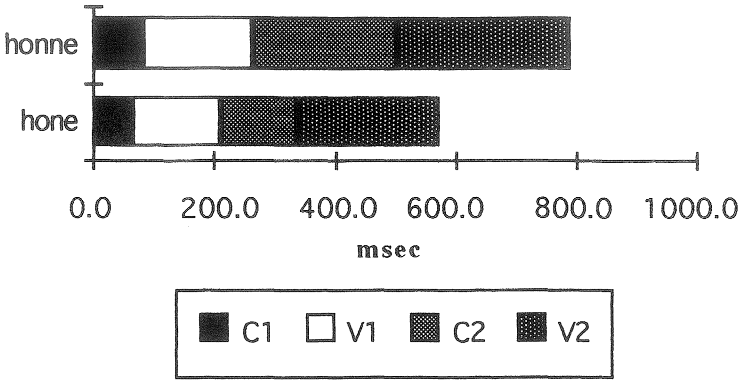
AP1(3)



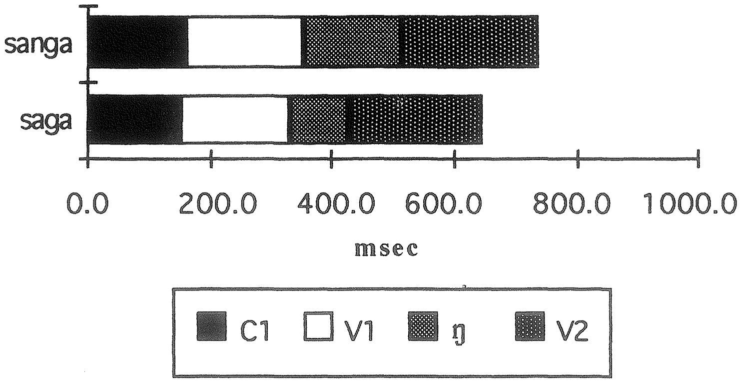
AP1(4)



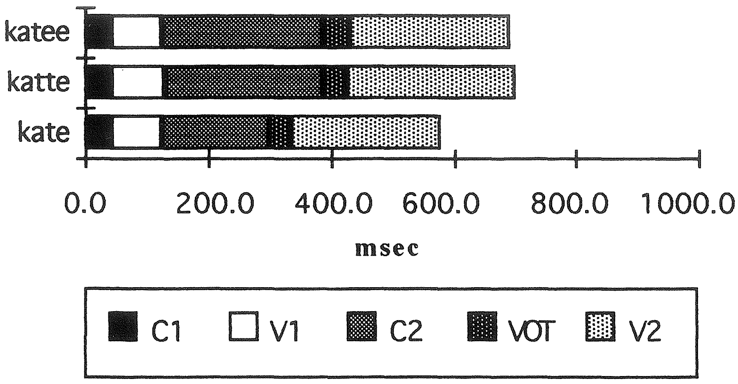
AP1(5)



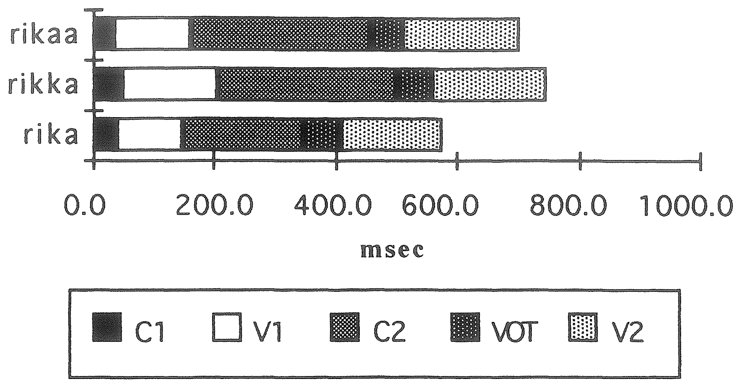
AP1(6)



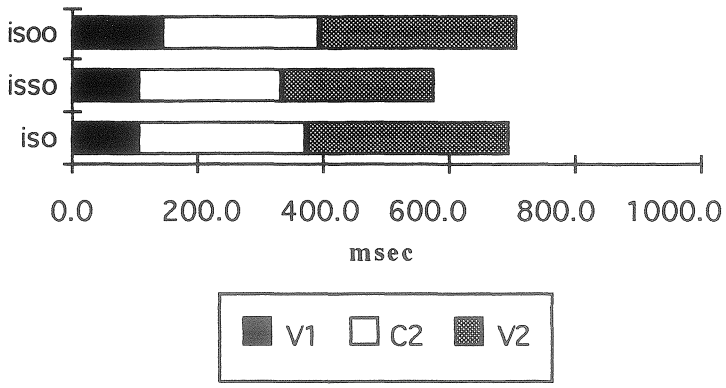
DS1(1)



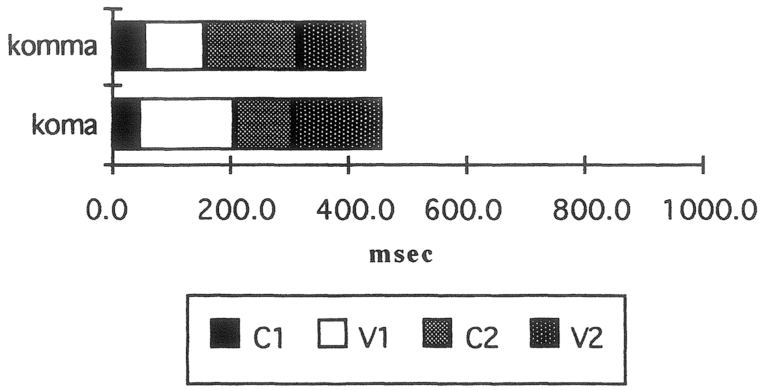
DS1(2)



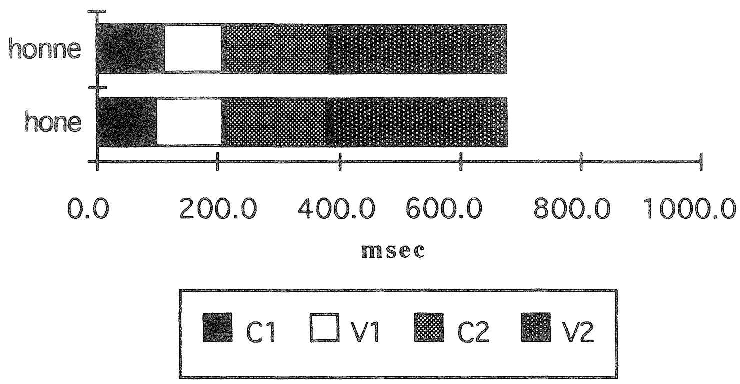
DS1(3)



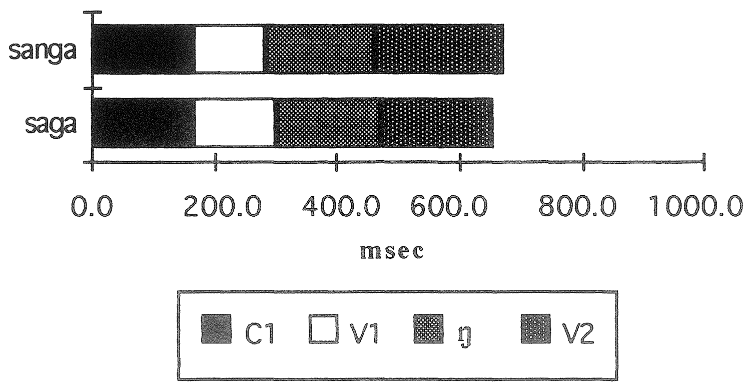
DS1(4)



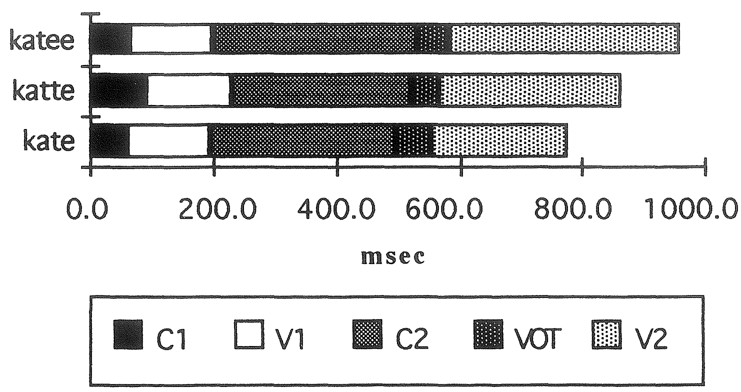
DS1(5)

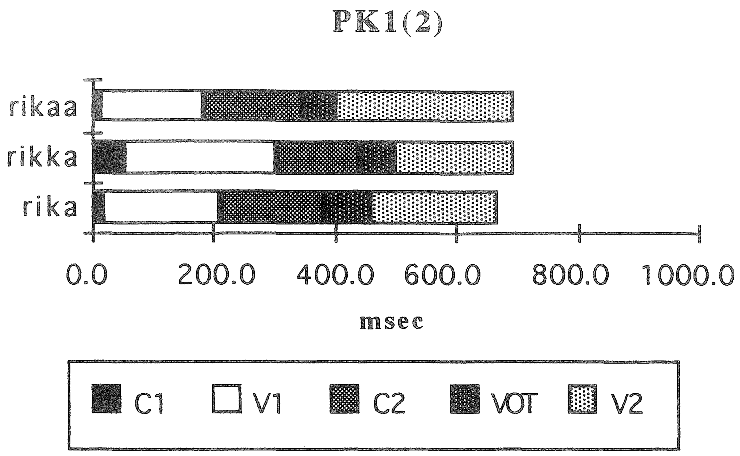


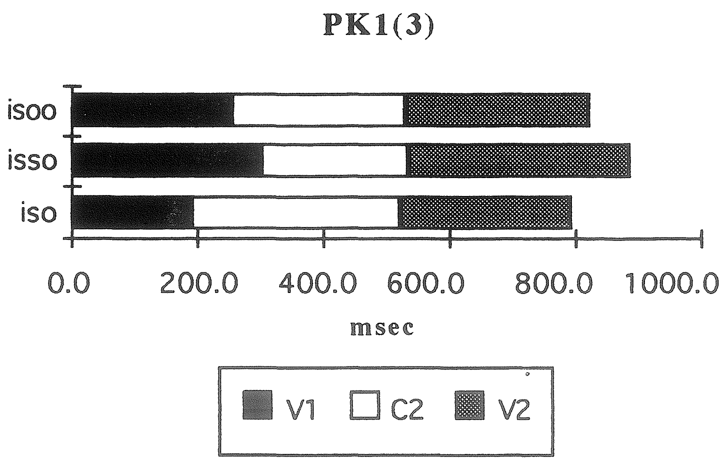
DS1(6)

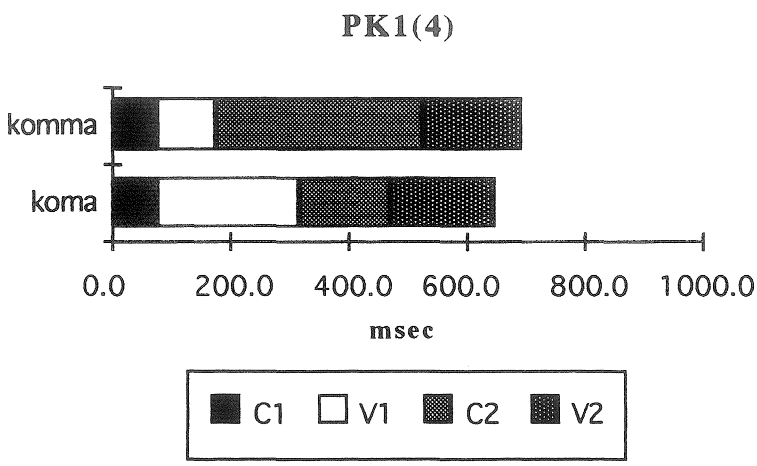


PK1(1)

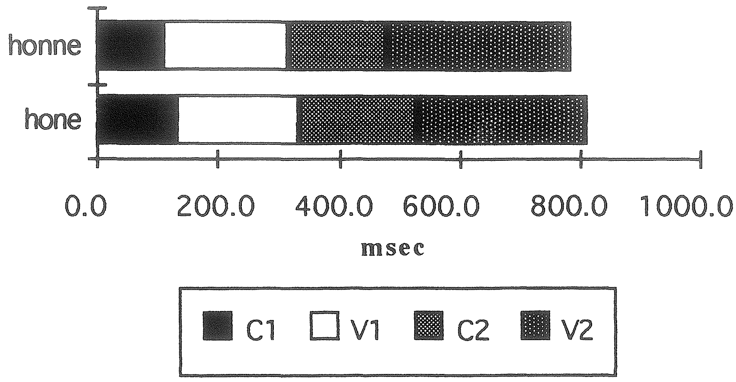



PK1(2)
rikaa
rikka
rika
0.0
200.0
400.0
600.0
800.0
1000.0
msec
C1
V1
C2
VOT
V2

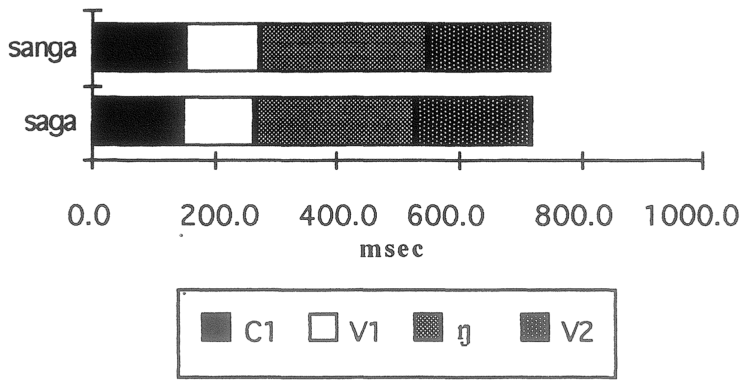

PK1(3)
isoo
isso
iso
0.0
200.0
400.0
600.0
800.0
1000.0
msec
V1
C2
V2


PK1(4)
komma
koma
0.0
200.0
400.0
600.0
800.0
1000.0
msec
C1
V1
C2
V2

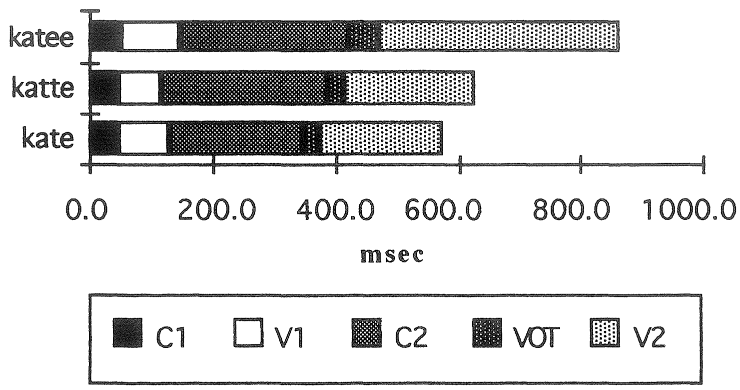
PK1(5)



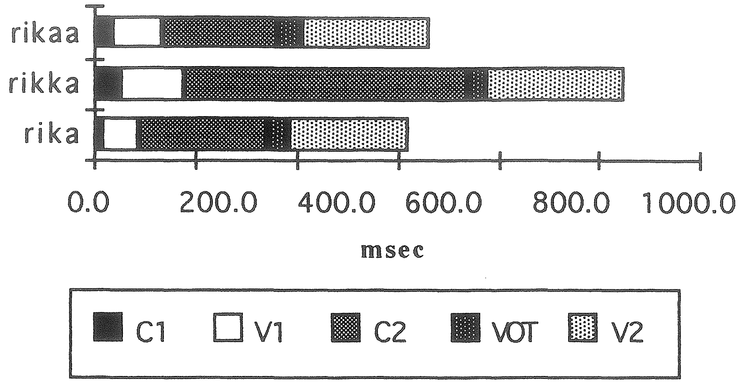
PK1(6)



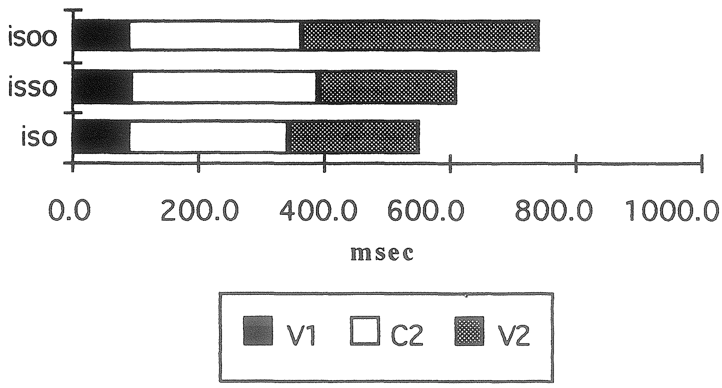
PV1(1)



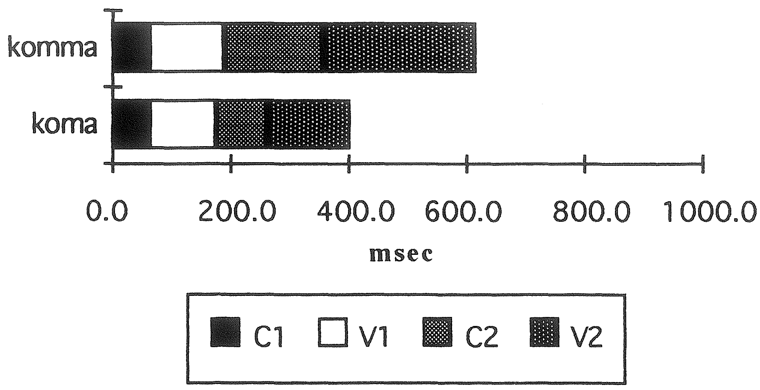
PV1(2)



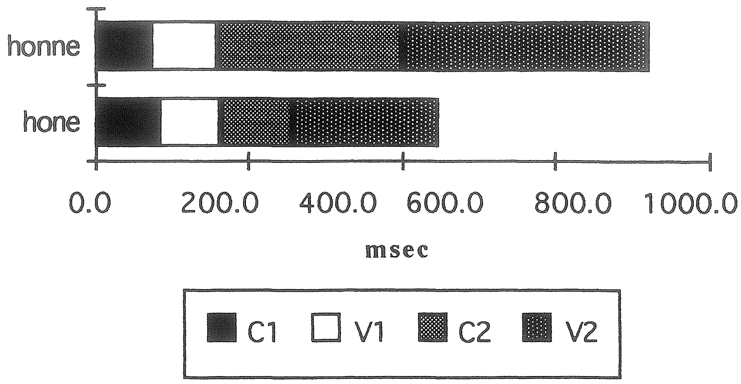
PV1(3)



PV1(4)



PV1(5)



PV1(6)

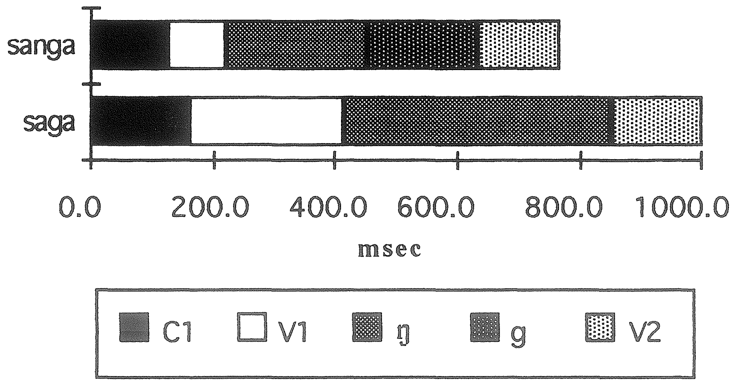


Fig.6.1 Production by DFAT1 learners

6.3.1 Word duration

Generally speaking, the relative duration of two- and three-mora words in DFAT learners' speech production is underdifferentiated (Appendix 6.3). In some cases, the duration of three-mora expressions is not always significantly longer than that of two-mora expressions; for example, the ratios of PK's RIKA series are 100/104/104, and the Anova result is not statistically significant (F=1.494; p=.2975). In other cases, the duration of two-mora expressions is longer than the corresponding three-mora expressions. This tendency is found in the examples such as DS1 iso>isso, PK1 hone>honne and PV1 saga>sanga. AP1 is the only subject who seems to differentiate word-level duration consistently, but the amount of differentiation is smaller than that of NS, especially in respect of CVCCV. For example, the ratios of word duration in

AP1's RIKA series are 100/124/155 while they are 100/151/140 for NS. The same tendency is observed in the KATE series: AP1's ratios are 100/116/134 while they are 100/142/134 for NS. AP1 seems to be able to control the timing of CVCVV in a native-like way, but not CVCCV. This example shows that the main problem is linked to CC length. In other words, CVCVV is three-mora for AP1, while CVCCV is not. This may imply that the acquisition of timing control of CVCVV sequence precedes that of CVCCV.

The equalized duration for two- and three-mora expressions can be observed as well as two-mora expressions being longer than three-mora expressions. This tendency is commonly observed in CVCV/CVNCV (e.g. DS1(4) (5) (6) and PK1 (4) (5) (6)).

Thus it can be said that DFAT beginning-level learners have problems in differentiating two- and three-mora expressions in terms of word duration, which is similar to the results for ANU learners described in Chapter 4.

6.3.2 Long/short contrast

The long/short contrast is often insufficiently marked for either C2 or V2 in learners' data (Appendix 6.4). This is described as 'underdifferentiation' by Han (1992), and is also found in the data obtained from the ANU subjects in Chapter 4. In this study, underdifferentiation of this contrast is further analysed into two sub-categories: (1) underdifferentiation caused by the absolute duration of a long phoneme being too short compared to the short counterpart, and (2) underdifferentiation caused by the absolute duration of a short phoneme being too long in relation to the long counterpart. It is claimed that these two categories are different in nature, and thus should be treated separately.

As mentioned in Chapter 4, it is generally believed that the main cause for underdifferentiation of the long/short contrast is shorter duration of the geminate C2. In DFAT1, this tendency was found in the HONE series for DS1 and PK1 (see DS1(5) and PK1(5)). However, longer duration of a single C2 was certainly a more frequently occurring problem than shorter duration of the geminate C2 in the DFAT1 data. This situation is similar to the ANU beginning-level learners' data. In my auditory impressions, all the words with a single obstruent C2 (i.e. [kate/katee], [rika/rikaa] and [iso/isoo]) were perceived as having a long C2.

Overexaggeration of long C2 was also found to be an attempt to make the distinction, and may be treated as a learner strategy for realizing the phonological contrast. The speakers are aware of the importance of the phonological contrast, and therefore try to make the durational contrast; however they have to overexaggerate the duration of the geminate C2 in order to produce the contrast because their single C2 is too long. In fact, despite the longer duration of single C2, the duration of a geminate C2 systematically overrides that of single C2, as illustrated in KATE (70/100/98), RIKA (65/100/75) and ISO (88/100/90) in AP1 (see AP1(1)-(3)).

In Chapter 4, I proposed the term 'reverse realization of phonological contrast'. This refers to the phenomenon where the duration of phonologically short segments is longer than that of their long counterparts, and was shown to be a common problem observed in the ANU data. In the present data, it was also found to be one of the most common problems observed in DFAT learners' speech. For example, PK1's data show that the duration of the single C2 was longer than the geminate C2 in the KATE (103/100/111), RIKA (127/100/120), ISO (142/100/119) and HONE (118/100) series (e.g. the single C2 in [kate] and [katee] is 103% and 111% respectively when the geminate C2 in [katte] is treated as 100%). DS1's results for the KATE (69/100/104), ISO (118/100/110) and HONE (101/100) series also show a similar problem. For PV1, the KATE (80/100/102) and SAGA (106/100) series involve the duration of a single C2 exceeding that of a geminate C2.

AP1 also underdifferentiated the long/short contrast, but her present data do not show the reverse phonetic realization of the phonological contrast. If VOT had been included in C2, however, this phenomenon would have been found: the duration of short C2 in [katee] is 107% of the long C2 in [katte], making the ratio of this series 78/100/107. The corresponding ratios are 70/100/98 when VOT is excluded, and thus reverse phonetic realization does not apply (see 4.1.2. for the treatment of VOT).

PV1 underdifferentiated the durational contrast of the C2 in the KATE series, but clearly differentiated that of the C2 in RIKA series. It seems that this speaker interpreted [rikka] as [rik-ka], and applied this strategy when producing a long [k]. In slow speech, the release of glottal closure was clearly observed (as shown in 4.3.3.), and this may be regarded as evidence of the segmentation of CVCCV as CVC-CV. The prosodic structure used by this speaker seems to be that of English syllable at stress foot boundaries (i.e. CVCCV is parsed as two words- CVC#CV). It is difficult to imagine a Japanese NS segmenting the sequence in this way, since geminate obstruents are invariantly word-medial in Japanese. Further investigation is needed in order to find out whether or not this can be regarded as solely due to interference from L1. Also it is not clear why this strategy was applied only for the production of the long velar stop [k] and not for the alveolar stop [t] in PV1's data.

With regard to V2, underdifferentiation due to both (1) shorter duration of phonologically long V2 and (2) longer duration of phonologically short V2 were observed (which is similar to the results obtained from the ANU subjects in Chapter 4). An example of (1) is PV1's RIKA series (see PV1(2)). The short V2 in [rika] and [rikka] were perceived as acceptable while the long V2 in [rikaa] was too short. On the other hand, (2) is more commonly observed in the KATE and ISO series. For example, the problem in the ISO series of DS1 (102/77/100) and PK1 (93/121/100) was their short V2 which was perceived as long. The second tendency was particularly prominent in words with an LH pitch pattern for DS and PK. It was found that the duration of V2 tended to be longer in such words than in words with an HL pitch pattern.

As discussed in Chapter 4, the above results may be closely related to the fact that English stress is associated with longer duration (as well as higher F0 and greater intensity). If learners interpreted a syllable with H tone as a stressed syllable (as it would be in English), the expected result would be a syllable with a longer duration. Another related problem which gave the impression of a short V2 being long, was that the word offset was often not associated with a glottal stop. This tendency was particularly prominent in DS1 who had no word-final glottal stops, a situation different from NS speech production in which word offset is associated with a glottal stop.

In summary, there is a tendency for a long V2 to be too short when the pitch pattern is HL; on the other hand, it is more common for the short V2 to be too long when the pitch pattern is LH. The nature of the underdifferentiation of the V2, therefore, seems to be quite different from that of the C2 which was discussed previously.

Generally speaking, underdifferentiation tends to occur more frequently with the C2 than the V2. This suggests that long/short differentiation is more difficult for the subjects to achieve with intervocalic obstruents than with word-final vowels. AP's data suggest that she has good timing control, because she makes a clear contrast in terms of absolute duration of the long V2. In AP1, all examples of long V2 in [katee], [rikaa] and [isoo] are longer than their short counterparts. In AP1's ISO series, for example, there is a clear differentiation between long and short V2 (54/52/100); on the other hand, the ratio for the C2 is underdifferentiated (88/100/90). Thus the data obtained in the present study imply that there is greater difficulty involved in the differentiation of the durational contrast in C2. The implication of this result for the teaching situation in the classroom is that more time is required for training in timing control of C2 than V2 if learners are to improve their pronunciation (see Chapter 7).

6.3.3 C2/V1 and V2/V1 ratios

In NS data, the C2/V1 ratio was generally about twice as large for words with a long C2 as for those with a short C2. Unlike NS, DFAT learners' speech often showed a smaller C2/V1 ratio for words with a long C2 than for words with a short C2. For example, the C2/V1 ratios for kate/katte/katee were 2.23/2.12/2.46 for PK1, 2.09/2.81/3.26 for DS1 and 1.69/2.00/2.24 for AP1 (Appendix 6.5). If the V1 duration were consistent, and the geminate C2 longer than the single C2, we would expect the C2/V1 ratio of CVCCV to be greater than that of CVCV or CVCVV. Contrary to expectations, the ratio for [katee] was greater than that of [katte] for all three subjects. The ratio for the learners' [katte], on the other hand, is much smaller than 3.21 for NS (the ratios are 1.68/3.21/1.62 for NS).

This situation is similar to the ANU learners' data (Fig.4.45-47). It can thus be said that a discrepancy between phonology and phonetic realization is found in the C2/V1 ratio, as well as in durational contrast, in the speech production of the DFAT informants.

The possible causes of this problem suggested in the previous sections are: (1) the absolute duration of a geminate C2 being too short compared to the single counterpart, (2) the absolute duration of a single C2 being too long in relation to the geminate counterpart, (3) the duration of the preceding vowel V1 being too long compared to the duration of the geminate C2. Generally speaking, (2) was the most common tendency observed in all informants. Also, DS1 and PK1 demonstrated a few examples of (3). For example, PK1's examples with an obstruent C2 consistently have the smallest values for CVCCV: 2.23/2.12/2.46 for the KATE series, 0.91/0.55/0.97 for the RIKA series and 1.75/0.77/1.09 for the ISO series.

A significant difference was observed for both PK1 ($F=5.917$, $p=0.0381$) and DS1 ($F=12.266$, $p=0.0076$) regarding the duration of V1 in [rikka] (see PK1(2) and DS1(2)). Similar results were observed in the duration of V1 in [isso] for PK1 ($F=8.584$, $p=0.0172$) (PK1(3)). DS1, on the other hand, showed a significantly longer V1 in [isoo] ($F=5.994$, $p=0.0371$).

The influence of segmental structure upon the duration of V1 may be one of the reasons for these results. One possibility is that the duration of V1 is lengthened before a geminate C2 (i.e. CVCCV) as discussed in Chapter 4 for the ANU learners. This seems to be the case in PK1's results. The other possibility is that lengthening of the V1 is due to an expectation of a long word-final V2 (i.e. CVCVV), which may apply in DS1's ISO series. In both cases, lengthening of the V1 seems to be due to the existence of a long segment in the word.

Additionally, unlike NS, the average duration of CV in the first syllable is not consistent in some cases. For example, the duration of [ri] in [rikka] is longer than in

[rika] and [rikaa] in PK1(2). This means that the duration of the first mora can be affected by the segmental structure of a word.

Some problems are found with regard to the V2/V1 ratios. For example, the V2/V1 ratios for the KATE series are 2.90/3.04/3.19 (DS1) and 2.33/3.16/4.11 (PV1), while they are 1.82/1.37/3.99 for NS. These results show that the duration of V2 is too long in relation to that of V1 in CVCV and CVCCV. ISO series demonstrate a similar problem where the V2/V1 ratios are 3.04/2.25/2.16 (DS1) and 2.37/2.31/4.23 (PV1), while they are 1.72/1.33/3.75 for NS.

The V2 duration tends to be prolonged in words with an LH pitch pattern (see 6.3.2). Thus the results of V2/V1 ratios are influenced by the pitch pattern of the test words.

6.3.4 VOT

	t	k	t-test
C	47.29 msec (13.60)	51.83 msec (16.24)	t=-1.05, p=.2994
CC	35.24 msec (8.98)	44.44 msec (17.48)	t=-1.622, p=.1191
t-test	t=2.771, p=.009	t=1.255 p=.218	

Table 6.1 VOT

The values of average VOT were significantly greater than those of NS in both KATE and RIKA.

The difference in VOT between [t] and [k] was not statistically significant (t=-1.05, p=.2994), and a similar situation pertains between [tt] and [kk] (t=-1.622, p=.1191). This result is different from NS speech production, in which the VOT of a velar stop was consistently longer than that of an alveolar stop. Therefore, it can be said that VOT shows no clear relationship with the place of articulation in the DFAT1 results, the same as in the ANU data. With regard to VOT of geminate C2, [tt] is significantly longer than [t] (t=2.771, p=.009), but the difference between [k] and [kk] is not significant (t=1.255, p=.218).

6.3.5 Pitch assignment

The pitch assignment for words with a nasal C2 was a problem common to all subjects, and was particularly prominent in words with LH pitch.

PV1's [honne] was segmentally similar to the NS norm, but with an incorrect pitch assignment. The first CVC-syllable ([hon]) was assigned an L tone and the following syllable ([ne]) an H tone. It appears that the L tone was assigned to the first CVC-syllable because the subject had not acquired the phonological concept of a moraic nasal as a tone carrier, unlike NS. AP had a similar problem in terms of pitch assignment for this word, which was pronounced as [honnee]. The first syllable ([hon]) was assigned an L tone and the following long syllable ([nee]) an H tone. The lengthened V2 may be due to the accentual pattern of the word as discussed previously.

DS1 and PK1 pronounced this word as [honee] and it was perceived to have an LH pitch pattern with the first syllable ([hon]) assigned an L and the following long vowel ([ee]) assigned an HH. This pitch assignment seems to reflect the English syllable structure of CVN-V, and is regarded as L1 interference. Their speech production here thus has several concurrent problems: a shorter duration of [n] so that it is not perceived as a moraic nasal, a lengthened V2 due to the accentual pattern of the word, and also incorrect pitch assignment.

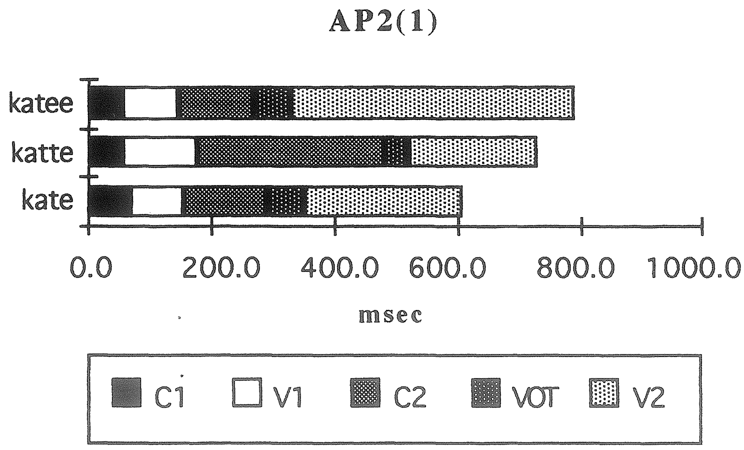
Phonologically speaking, the tone assigned to the second mora must be different from that of the first mora as discussed in the previous chapters.[40] The results obtained in the present study show that it is difficult for English-speaking subjects to acquire the concept of the moraic nasal as a tone carrier. PK1 had the same problem in words with an HL pitch pattern such as [komma] and [sanga], assigning an H tone to

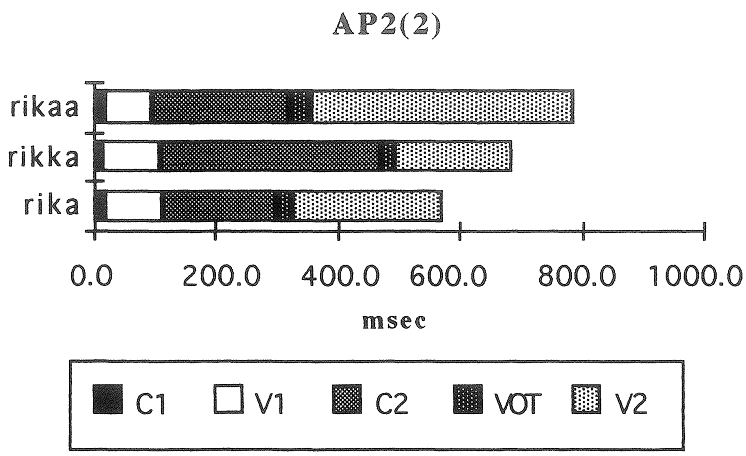
[40] In NS fast speech, however, [honne] with an LHH pitch pattern may be pronounced with an HHH pitch pattern as mentioned previously. This may be interpreted as anticipatory assimilation caused by the H tone on the second mora (see 4.2.2.5).

the first syllable (CVN) rather than an H tone to the first mora (CV) and an L tone to the following moraic nasal (N).

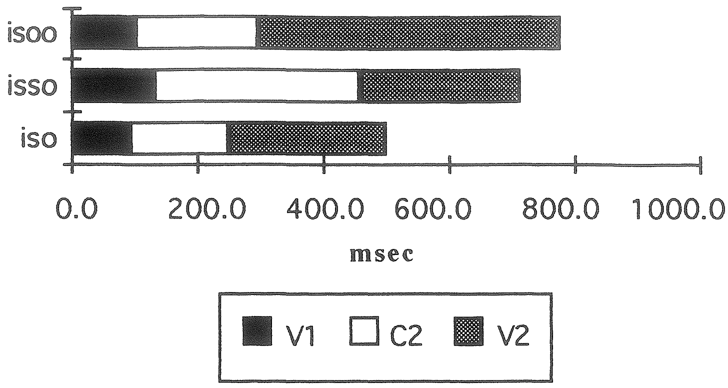
6.4 Second experiment

In this section, the results obtained in the second experiment are presented. The discussion mainly concerns the differences from the results obtained in the first experiment. My auditory impression of DFAT 2 is presented in Appendix 6.6.

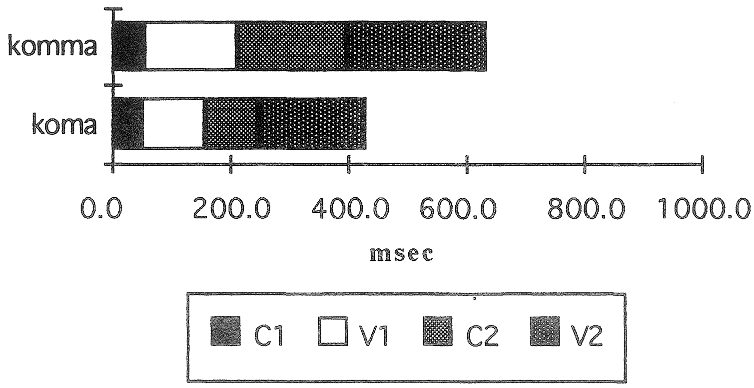




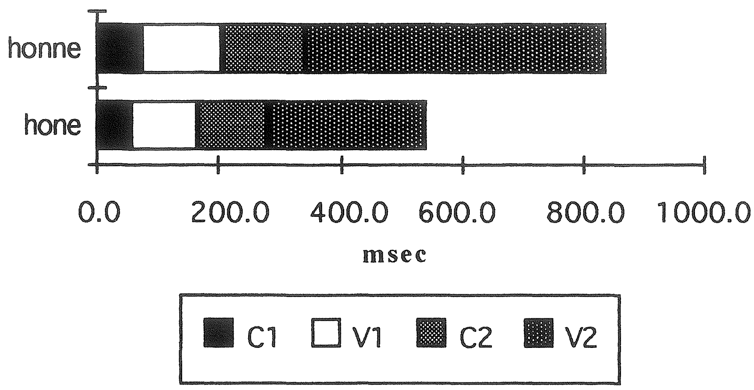
AP2(3)



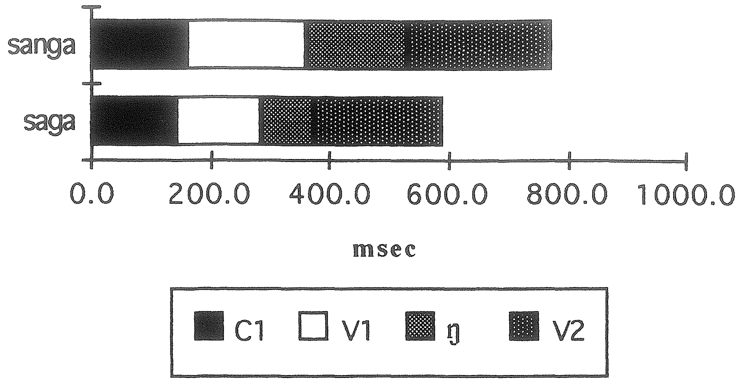
AP2(4)



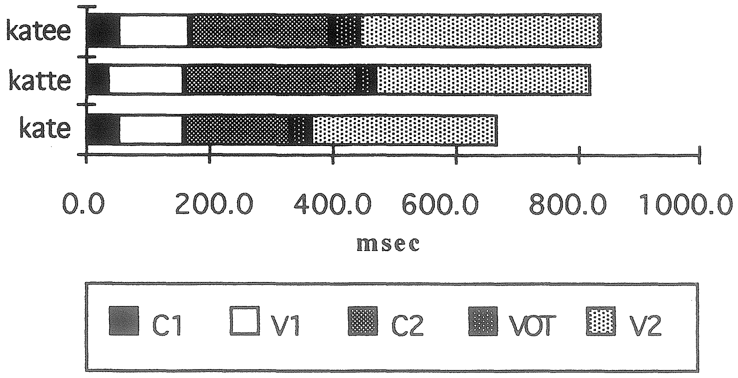
AP2(5)



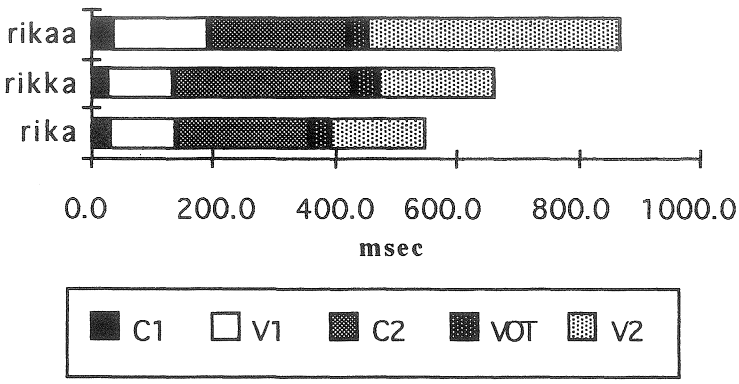
AP2(6)



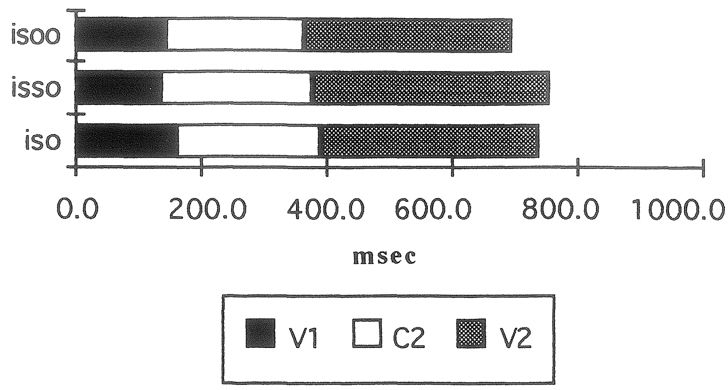
DS2(1)



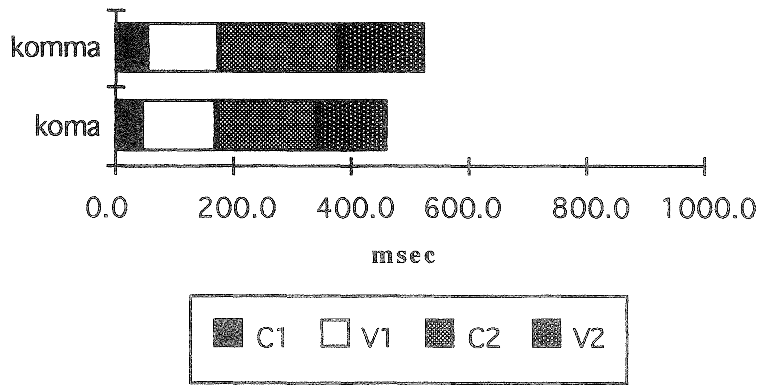
DS2(2)



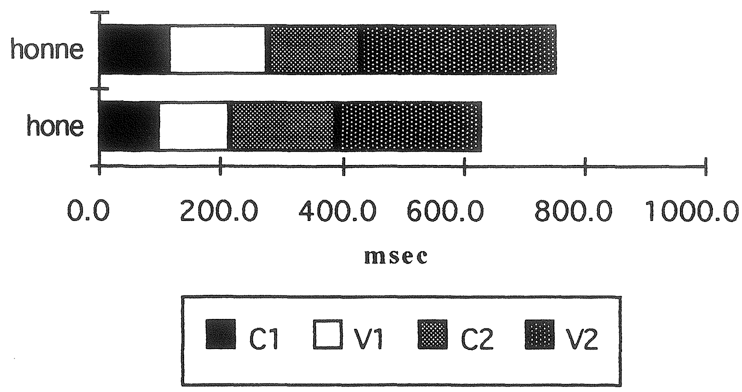
DS2(3)



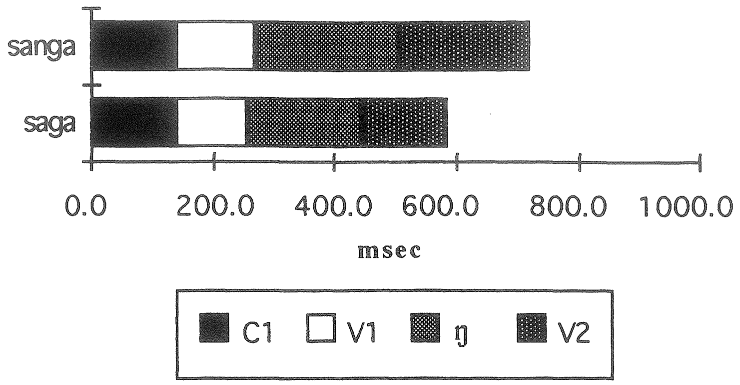
DS2(4)



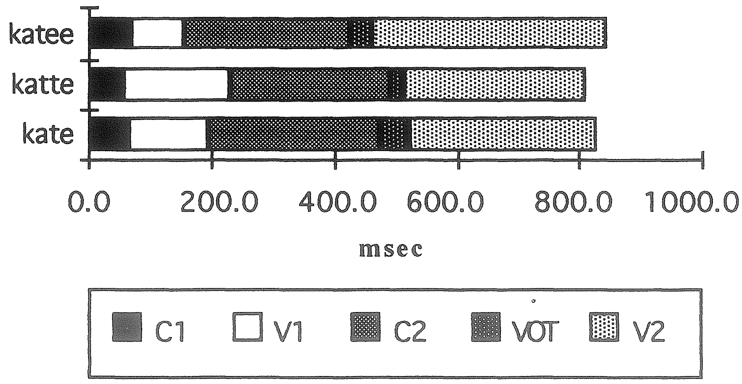
DS2(5)



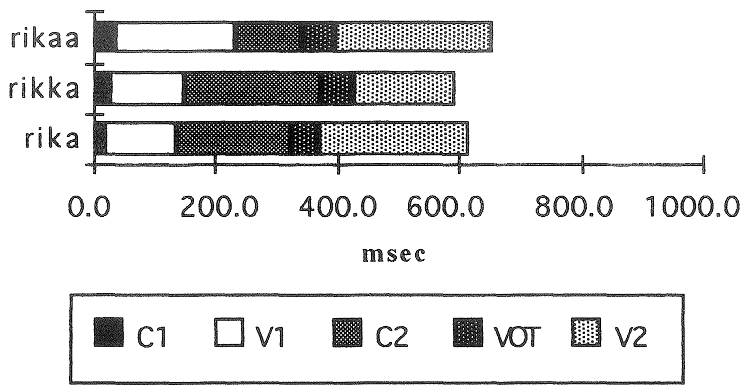
DS2(6)



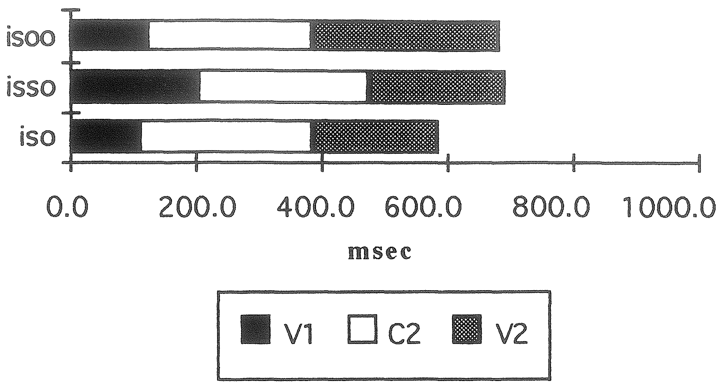
PK2(1)



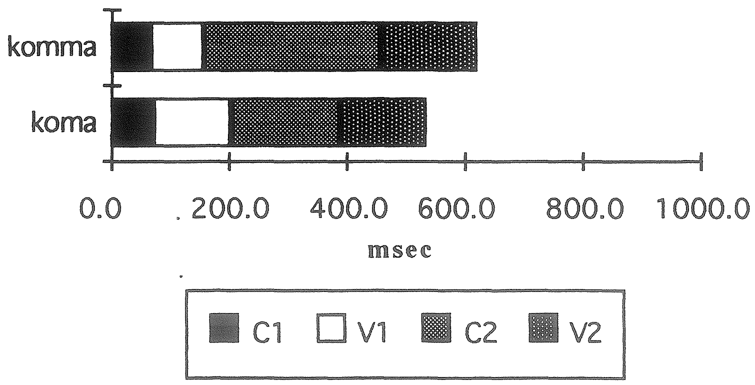
PK2(2)



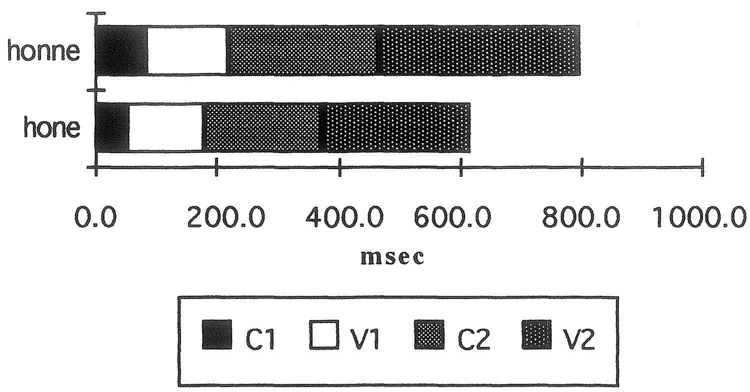
PK2(3)



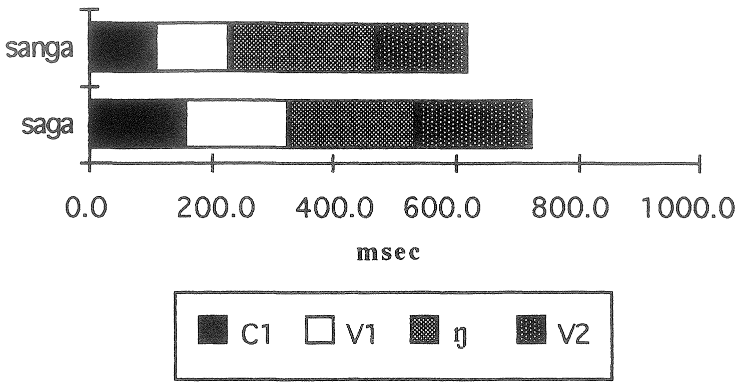
PK2(4)



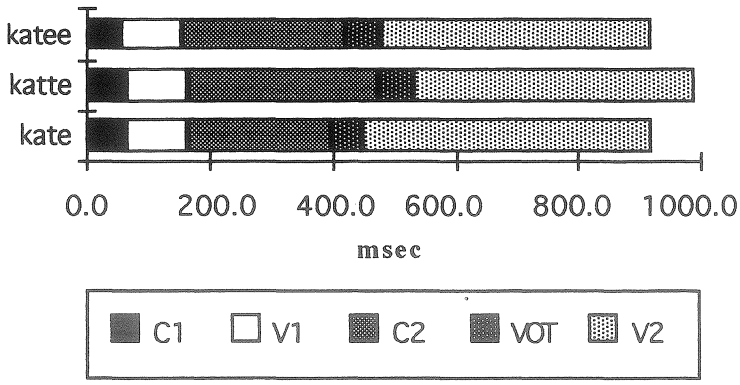
PK2(5)



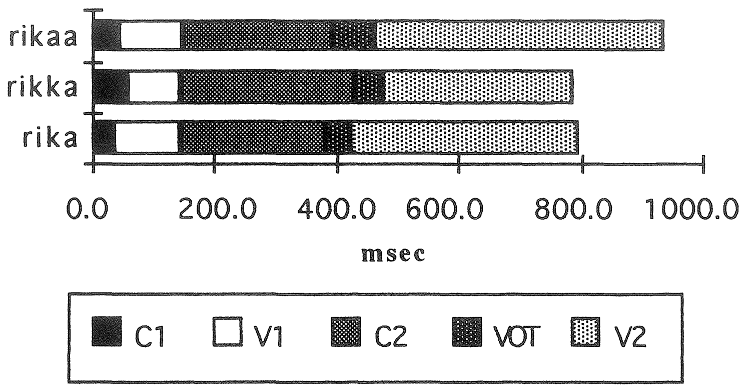
PK2(6)



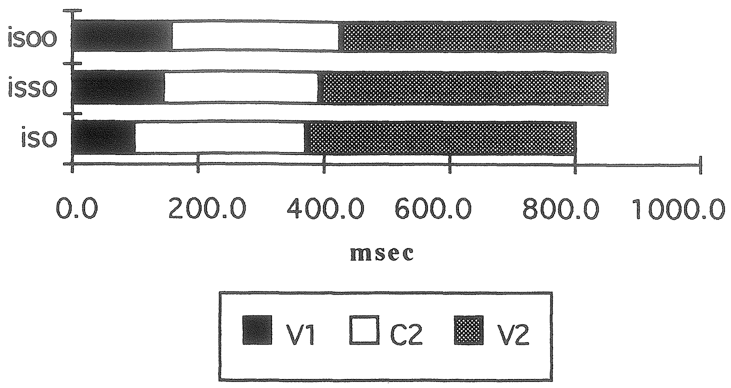
PV2(1)



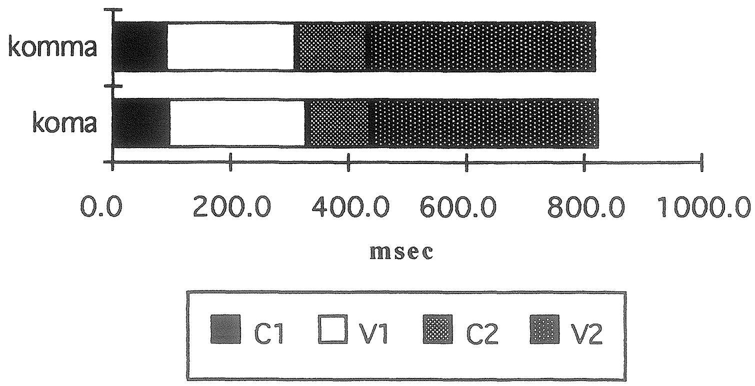
PV2(2)



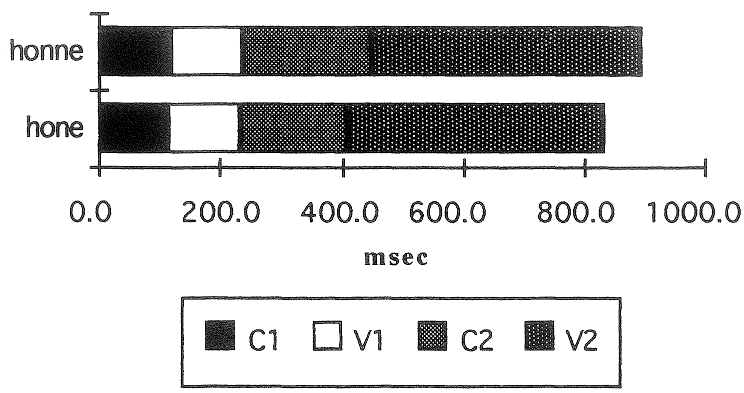
PV2(3)



PV2(4)



PV2(5)



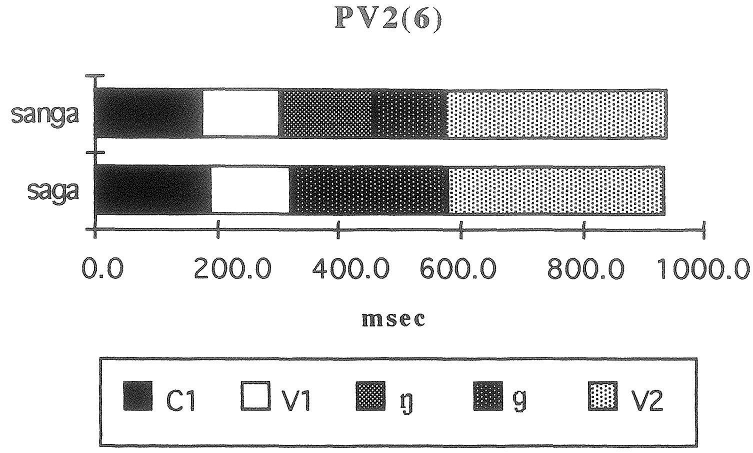


Fig.6.2 Production by DFAT2 learners

6.4.1 Word duration

The ratios of word duration obtained in the second experiment (Appendix 6.8) show some changes from the data obtained in the first experiment which was described in the previous section.

Firstly, the differentiation of word duration between two- and three-mora expressions has become clearer in the data of DS2 and PK2 (Appendix 6.8). In the first experiment, most subjects did not clearly differentiate between two- and three-mora expressions except for AP1. In the second experiment, DS2, for example, shows that the average duration of two-mora expressions in the KOMA (100/114), HONE (100/120) and SAGA (100/123) series is shorter than that of the corresponding three-mora expressions. This is an improvement since DS1 showed no word-level durational distinction in the words with a nasal C2. PK2 also shows a similar pattern in the KOMA (100/116) and HONE (100/130) series, but his SAGA series has a problem of a longer average duration for the two-mora expression [saŋa] than the corresponding three-mora expression [saŋŋa] (100/86). With regard to AP2, differentiation between the word duration of two- and three-mora expressions generally has become greater than for AP1. For example, the ratios of differentiation in the ISO series are 100/142/154 in AP2, while they were 100/106/128 in AP1.

Secondly, there is a tendency in some speakers' data towards identical word duration for two- and three-mora expressions. This is particularly prominent in the speech production of PV2. In comparison with the PV1 data, it has become more difficult to distinguish two-mora expressions from three-mora expressions based on word duration. PK2 also shows a similar tendency in the KATE series. The ratios are 100/108/100 and 100/98/102 in PV2 and PK2 respectively, and the differences between two- and three-mora expressions are very small.

6.4.2 Long/short contrast

In experiment 1, it was found that all subjects had problems in the production of the long/short contrast. Some improvement can be observed regarding this aspect of speech production in experiment 2 (Appendix 6.9). For example, AP2 seems to have acquired the durational contrast of the intervocalic C2 in the KATE and ISO series, mainly because the duration of the short C2 in CVCV and CVCVV has become shorter. With regard to the RIKA series in AP2, however, short [k] remained too long: all C2 in AP2's RIKA series were perceived as long [k] by an NS. This example seems to support the claim that the acquisition of alveolar stops precedes that of velar stops.

The long/short contrast is still not achieved in many cases. In fact, all obstruent C2 in DS2, PK2 and PV2 are perceived as long C2 by an NS except for [rikaa] in PK2. Also, the durational contrast in [honne] vs [hone], which was clear in AP1, has disappeared in AP2. All this implies that the subject's timing control is in a continuingly changing state and is still imperfect.

Reverse realization of the long/short contrast is no longer observed in DS2. The ratio shows that geminate C2 is consistently longer than single C2, unlike the results obtained in DS1: KATE (60:100:80), RIKA (75:100:77) and ISO (95:100:92). These examples may be regarded as an improvement (compared to DS1) although the durational contrast is still underdifferentiated and all single C2 are still perceived as long by an NS.

PK2 also shows some improvement (e.g. 84/100/50 for the RIKA series) although his long/short contrast in the KATE series still shows reverse realization (106/100/103). The overall frequency of the occurrence of reverse realization, however, has decreased from PK1.

With regard to PV2, it seems that the long/short contrast is less distinguishable in the RIKA and ISO series, where the ratios have changed from 45/100/41 (PV1) to 83/100/86 (PV2), and from 86/100/92 (PV1) to 110/100/107 (PV2) respectively. Thus the durational contrast is less distinctive in PV2. It appears that this tendency is closely related to the equalization of word duration described above. In this speaker's data, it seems that the durational contrast is lost in both word duration and long/short C2, and a new tendency towards equalized duration is observed.

With regard to V2, it seems that the acquisition of the long/short contrast in words with an HL pitch pattern precedes that for words with an LH pitch pattern. As far as the data obtained in the present study are concerned, all speakers demonstrated a durational contrast concerning V2 in the RIKA series. On the other hand, there are still some problems in the KATE (PV2: 106/105/100), and ISO (DS2: 105/113/100 and PV2: 98/104/100) series. This is because, when the second syllable is assigned an H tone, the duration of the V2 tends to be longer even when it is phonologically short. This tendency was also observed in the first experiment.

In PV2, it appears that the tendency to equalize duration (as seen in word duration and the C2) also applies in the interpretation of the data concerning the V2.

6.4.3 C2/V1 and V2/V1 ratios

A common tendency found in DFAT2 is that the C2/V1 ratio for geminate C2 is too small compared to NS. This implies that the duration of the geminate C2 is still too short in relation to that of V1. On the other hand, the C2/V1 ratio for single C2 is too big compared to that of NS. This implies that the duration of single C2 is too long in relation to V1 duration in the DFAT2 data.

Reverse realization of the C2/V1 ratio has disappeared from the subjects' data except for PK2 (Appendix 6.10). PK2 still shows a longer V1 duration before the geminate C2 in the KATE and ISO series, which is similar to the tendency observed in PK1. In the RIKA series, on the other hand, a lengthening of the V1 was observed in CVCVV (e.g. [rikka] was pronounced as [riika] in PK1, while [rikaa] was pronounced as [riikaa] in PK2). This makes the C2/V1 ratio of CVCVV the smallest (1.58/1.80/0.57). It is interesting to note that a longer V1 in CVCVV was also found in [isoo] for DS1. Therefore it seems that there are at least two causes for the lengthening of V1: one is due to the gemination of C2 (see Chapter 4), and the other to the influence of a long V2. The longer V1 in the CVCVV sequence found in this section may be interpreted as the result of anticipatory lengthening which is caused by the existence of a long V2.

With regard to V2/V1 ratios, those obtained in DFAT2 do not show any improvement from those in DFAT1. The V2/V1 ratios for the KATE series are 2.73/2.78/3.31 (DS2) and 4.58/4.71/4.42 (PV2), while they are 1.82/1.37/3.99 for NS. The V2 duration is still too long in relation to that of V1 in CVCV and CVCCV. ISO series also demonstrate a similar problem where the V2/V1 ratios are 2.15/2.73/2.31 (DS2) and 4.39/3.16/2.77 (PV2), while they are 1.72/1.33/3.75 for NS.

On the other hand, it is interesting to note that V2/V1 ratios of CVCVV are greater than those of NS in AP2. For example, they are 2.90/1.76/5.11 and 2.74/1.96/4.67 for KATE and ISO series respectively. These results imply that the V2 duration is too long in relation to the V1 duration in CVCVV due to overexaggeration of long V2.

6.4.4 VOT

The average VOT of the learners was longer than that of NS in the first experiment. Although this still remained the same in the second experiment in all examples, the difference between the average VOT of DFAT2 and that of NS has decreased. The decrease in VOT is particularly prominent in the realizations of [k] and [kk], which were 51.83 msec and 44.44 msec respectively in DFAT1, and 39.65 msec and 39.43 msec in DFAT2 (Table 6.2).

	t	k	t-test
C	49.23 msec (17.10)	39.65 msec (13.92)	t=2.13, p=.0386
CC	36.38 msec (17.97)	39.43 msec (16.98)	t=-.426, p=.6741
t-test	t=2.091, p=.0441	t=.042 p=.9664	

Table 6.2 VOT

Unlike NS, the VOT associated with [k] is not significantly longer than that associated with [t]. This situation is similar for VOT between [kk] and [tt]. Therefore, it can be said that VOT did not demonstrate as clear a relationship with the place of articulation in the DFAT2 data (this situation is the same in the DFAT1 data). It seems that, unlike NS, the phonetic universal principles still do not operate in DFAT2. With regard to the difference in VOT between a single and a geminate C2, it is significantly longer for [t] than for [tt], but this situation does not apply for [k] and [kk].

6.4.5 Pitch assignment

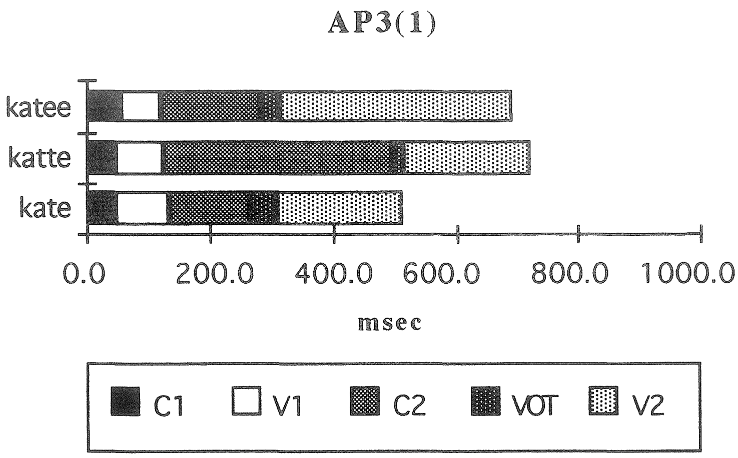
The distinction previously shown in the pitch patterns of KOMA series has disappeared in DS2. In DS1, [koma] and [komma] were pronounced as [kooma] HHL and [komma] HLL, and thus distinguished (although the phonetic realization of the distinction was different from that of NS). In DS2, however, both of these words were pronounced as [komma] HLL and it is no longer possible to perceive the differences between [koma] and [komma] in the auditory level.

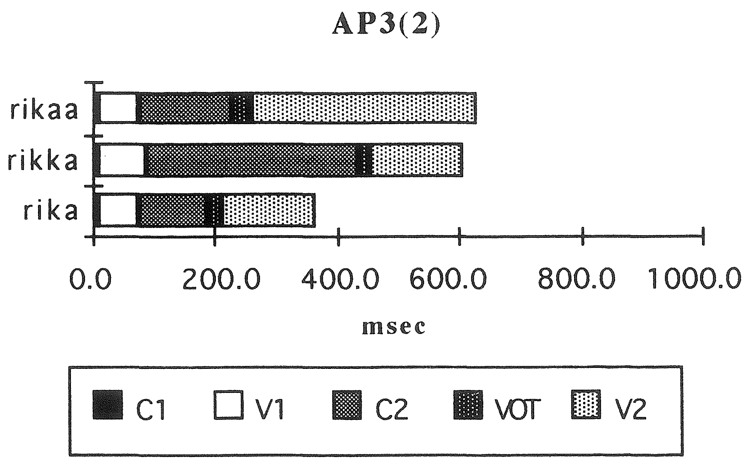
PK2 continues the incorrect pitch assignment for words with a nasal C2. The first CVC-syllable is assigned one pitch for all examples (e.g. [kom-ma] HH-L). This occurs consistently and is similar to PK1. This means that PK2 has not yet acquired the concept of a moraic nasal C2 as a tone carrier, and he does not allow a pitch change within the first syllable (e.g. [kom-ma] HL-L). In this respect, PK2 still has some difficulties with the acquisition of correct pitch assignment.

All subjects still have problems with the pitch assignment in the HONE series. They all give long duration to the V2 in [honne] which may be due to the LH pitch pattern, and AP2 and DS2 have short duration for the intervocalic nasal C2. This shows that the acquisition of timing control and correct pitch assignment is particularly difficult for words with a nasal C2.

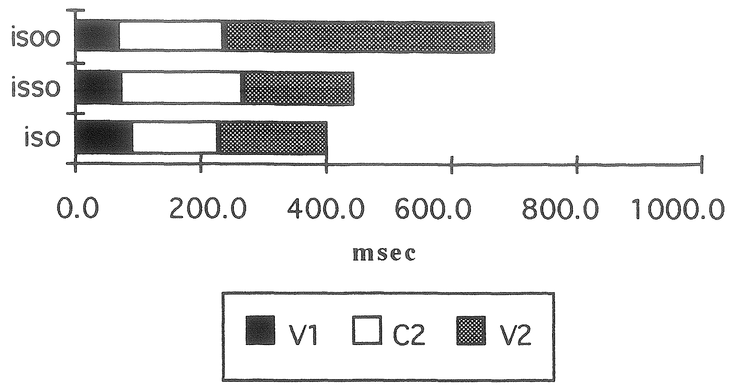
6.5 Third experiment

The final experiment was conducted in November 1993, approximately four months after the second experiment. The results are presented and the discussion will focus on the differences from the first and second experiments. My auditory impression of DFAT 3 is shown in Appendix 6.11.

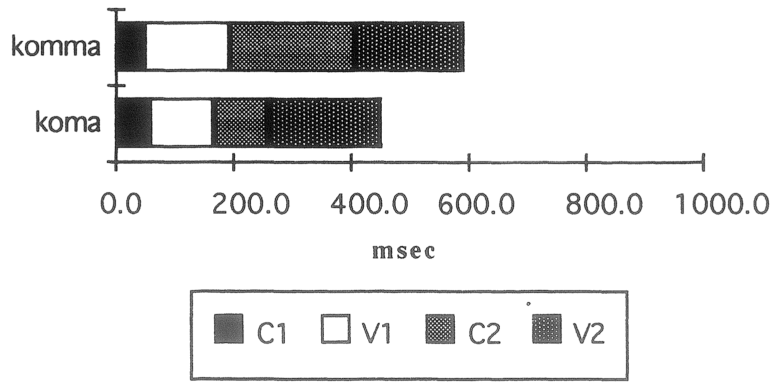




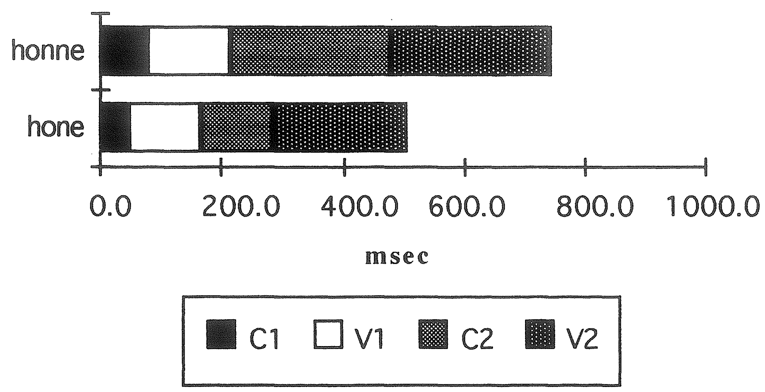
AP3(3)



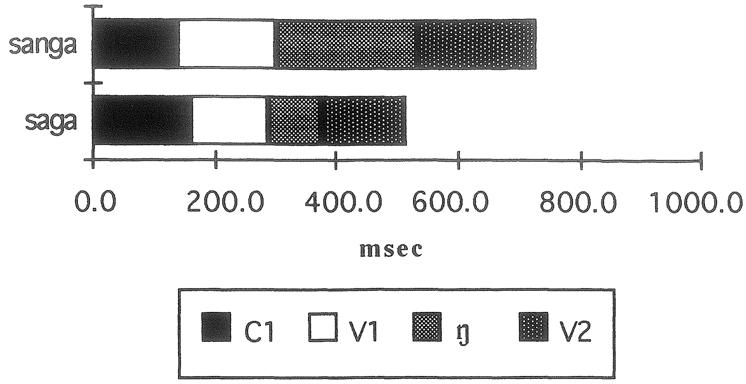
AP3(4)



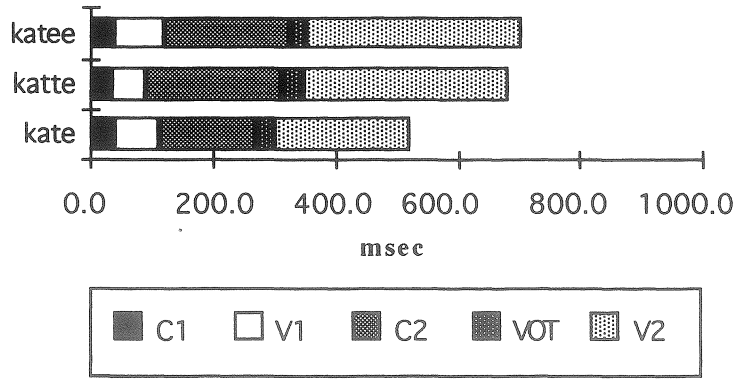
AP3(5)



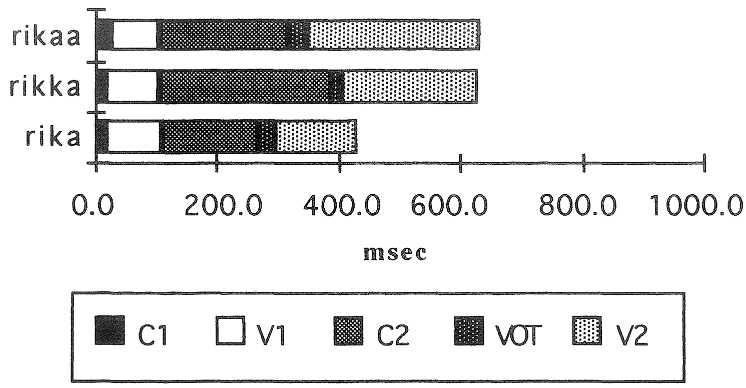
AP3(6)



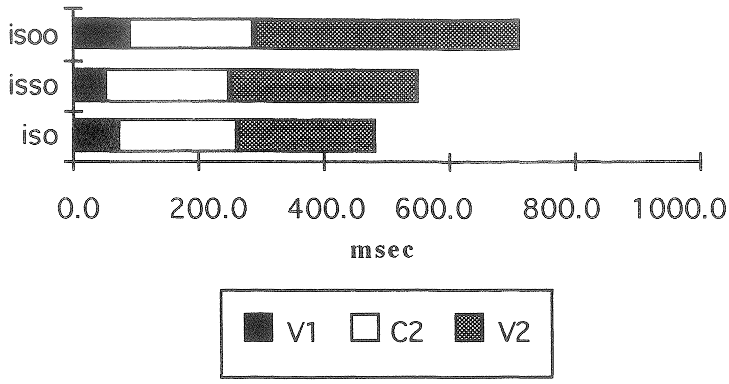
DS3(1)



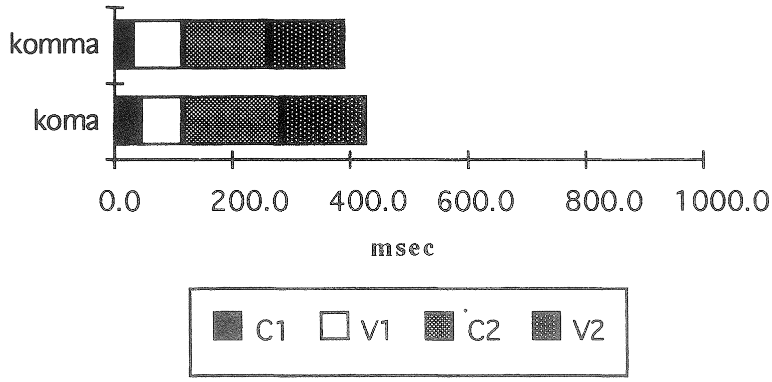
DS3(2)



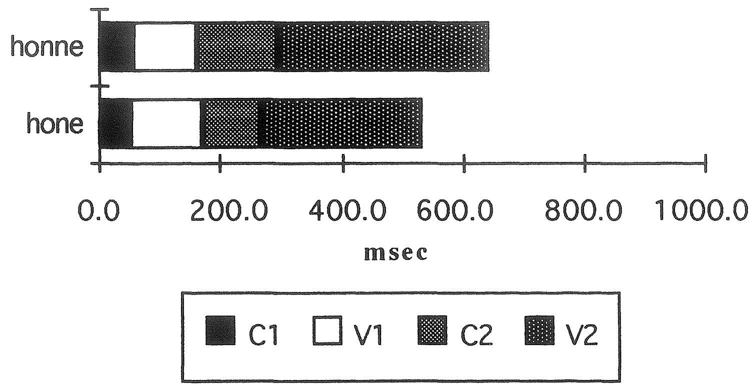
DS3(3)



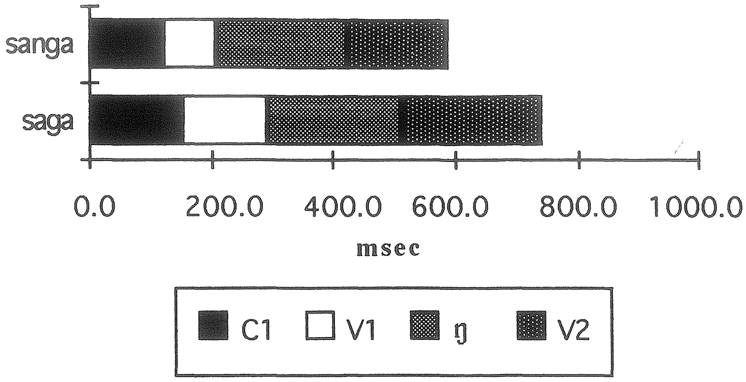
DS3(4)



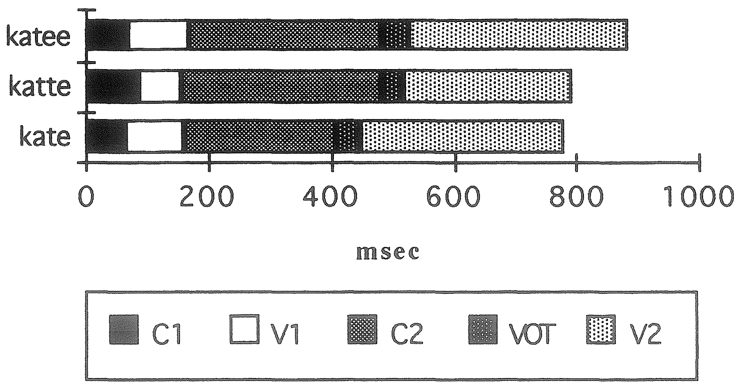
DS3(5)



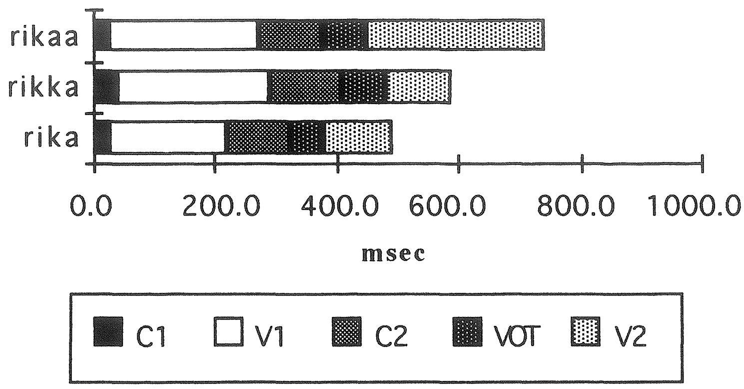
DS3(6)



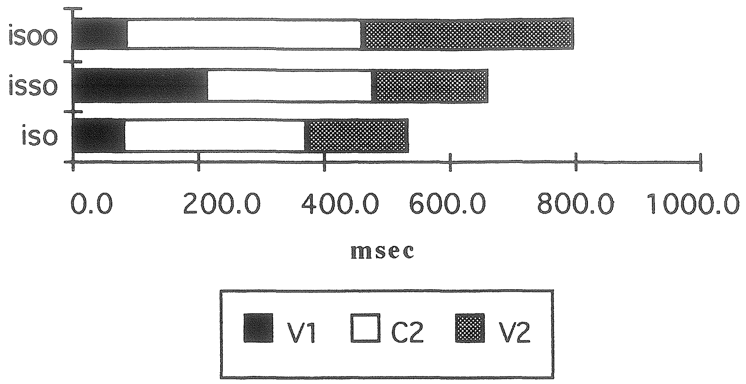
PK3(1)



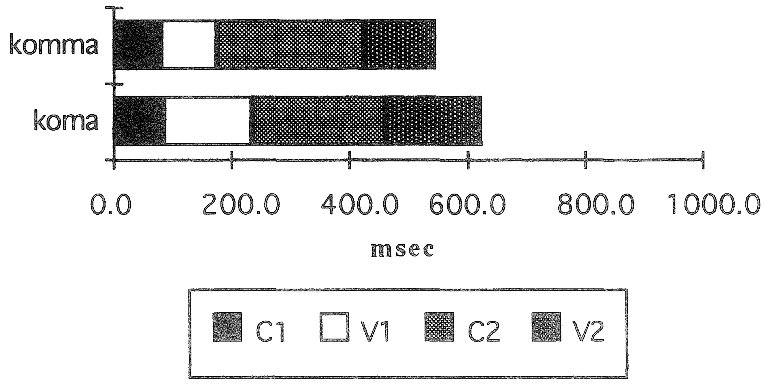
PK3(2)



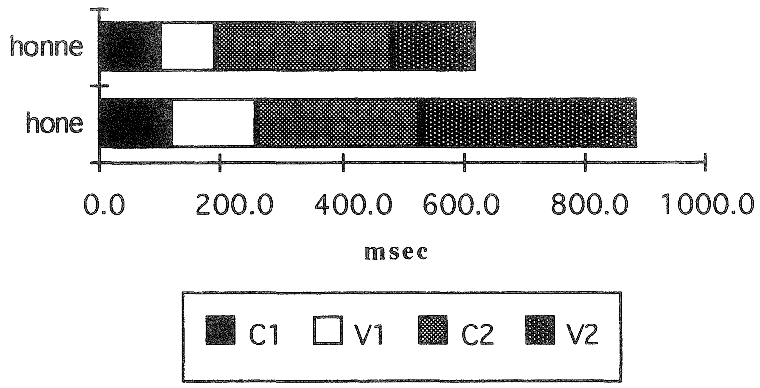
PK3(3)



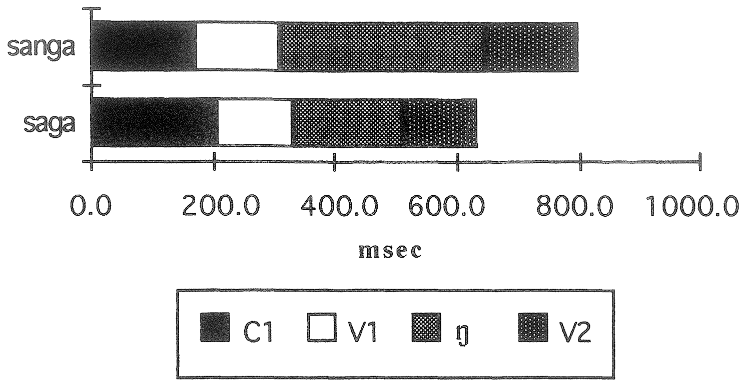
PK3(4)



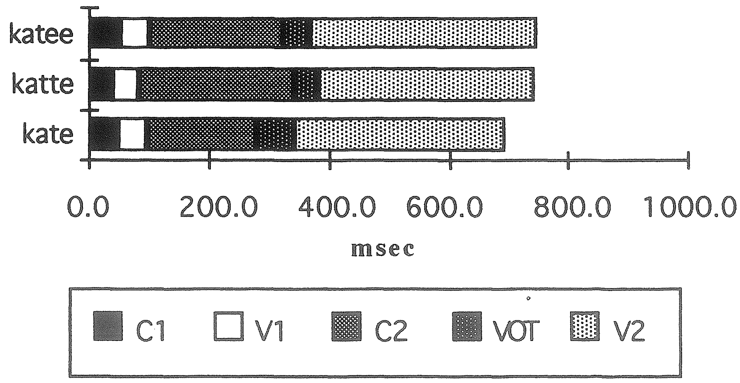
PK3(5)



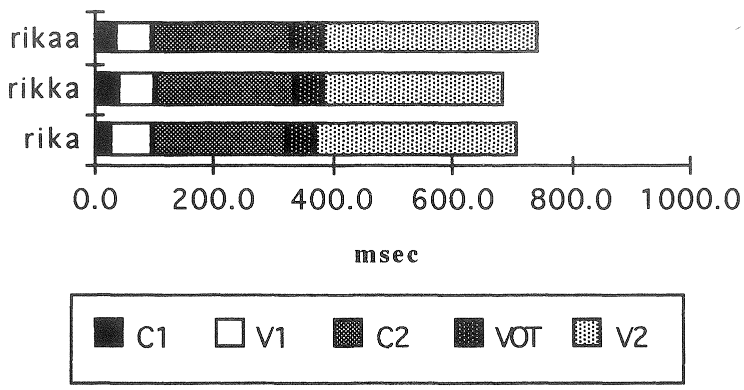
PK3(6)



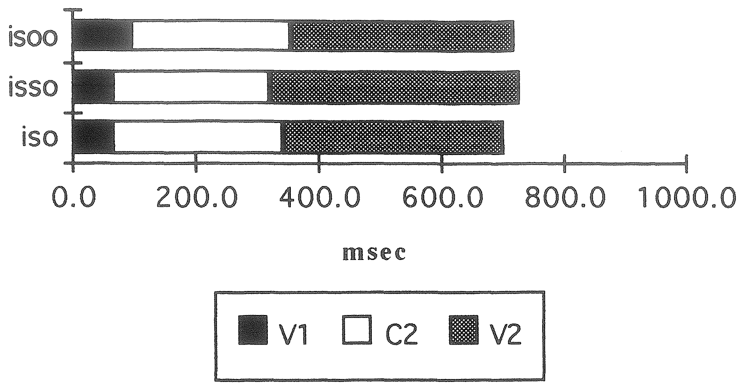
PV3(1)



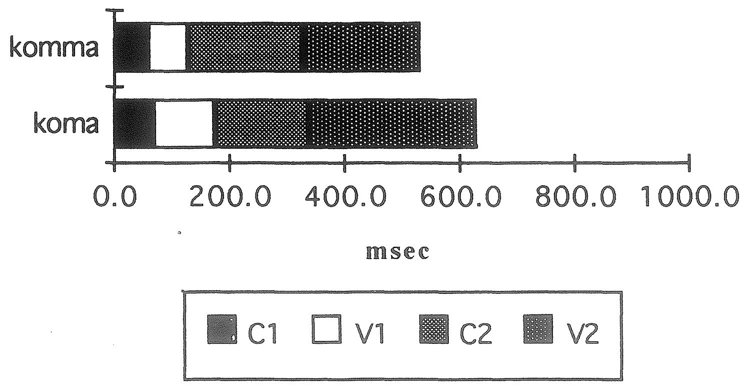
PV3(2)



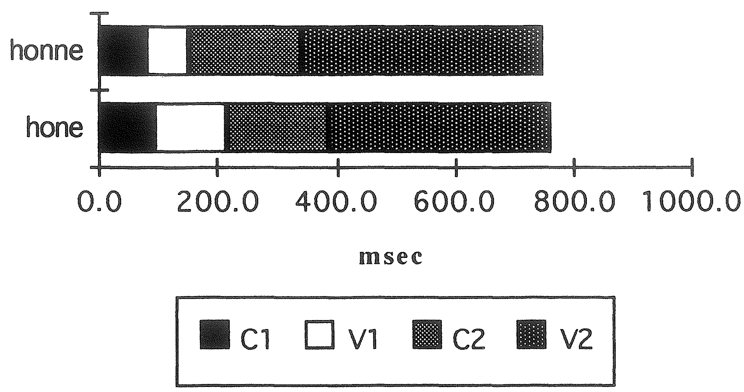
PV3(3)



PV3(4)



PV3(5)



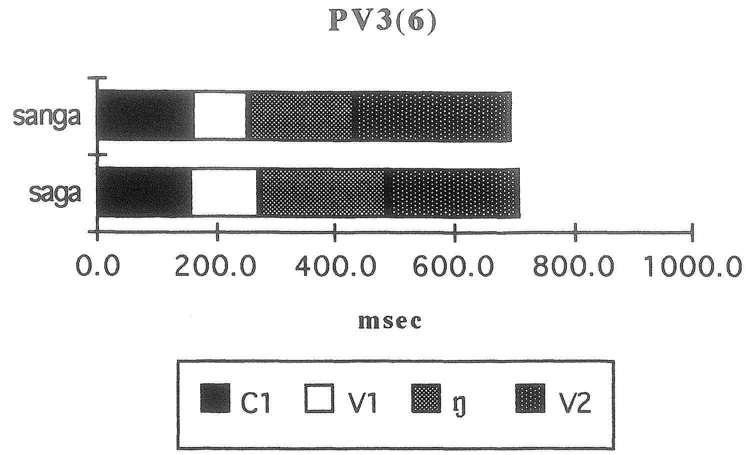


Fig.6.3 Production by DFAT3 learners

6.5.1 Word duration

Generally speaking, the durational contrast between two- and three-mora expressions is clearly shown in AP3 (Appendix 6.13). Therefore, it can be said that AP gradually improved her control of mora timing throughout her one year's training.

DS3 distinguishes the duration of expressions with an obstruent C2, but not words with a nasal C2. In DS2, it appeared that he had acquired timing control for words with a nasal C2, but contrary to this expectation, DS3 shows that the two-mora expressions in the KOMA (100/92) and SAGA (100/80) series are longer than the corresponding three-mora expressions. PK3 also showed this tendency in the KOMA (100/87) and HONE (100/70) series. Consequently it is not possible to say that these speakers have acquired accurate mora timing in Japanese during their one year's intensive training.

PV3 shows that the duration of two- and three-mora expressions has become even more equalized than it was in PV2. It appears as if this speaker intends to maintain an identical duration for all words within the same series. As a result, it has become impossible to distinguish two- and three-mora expressions based on word duration.

The data obtained from the experiments 1-3 show that most learners did not acquire correct timing control for word duration after one year's intensive training. Some two-mora expressions in DS3 and PK3 were still longer than the corresponding three-mora expressions. AP is the only subject who seems to have acquired timing control over word duration, but in her case, the duration of two- and three-mora expressions was already differentiated in the first experiment.

The results obtained in the present study show the difficulties involved in the acquisition of timing control in Japanese, especially concerning word-level durational contrast, in words with a nasal C2.

6.5.2 Long/short contrast

AP2 showed a long/short contrast for intervocalic C2 in the KATE and ISO series in the second experiment. In the final experiment, AP3 showed a clear durational contrast in the RIKA series as well (Appendix 6.14). The change in the durational contrast of C2 in the RIKA series is as follows: 65/100/75 (AP1), 51/100/62 (AP2) and 33/100/43 (AP3). Therefore, as far as the present data are concerned, it appears that AP has improved her timing control and gradually acquired the long/short contrast for C2 during the year. The order of phonetic acquisition is as expected: the acquisition of alveolar stops precedes that of velar stops. It was also found that the acquisition of durational contrast in word final V2 preceded that of intervocalic C2 in AP's data.

Reverse realization of the long/short contrast for C2 is still found for the ISO series for PK3 (110/100/141) and PV3 (110/100/101), the KOMA series for DS3 (113/100), and the SAGA series for DS3 (101/100) and PV3 (117/100). Also, underdifferentiation of the contrast is seen in examples such as the RIKA series for PV3 (95/100/100), the ISO series for DS3 (95/100/100), the KOMA series for PK3 (91/100), and the HONE series for both PK3 (92/100) and PV3 (89/100). In these examples, the difference is small and it appears PK3, DS3 and PV3 are losing the long/short contrast in their speech production.

As mentioned previously, the tendency towards the loss of long/short contrast was first found in the data obtained from PV2. After observing the data obtained in the third experiment, it now seems that this tendency occurs consistently in PV3. The other supporting evidence is that words which have an obstruent C2 and a long nasal C2 have become CVCCVV (e.g. [rikkaa], [kommaa]) and words with a short nasal C2 have become CVCVV (e.g. [komaa]). This seems to occur consistently, regardless of the pitch pattern. It appears as if PV has constructed his own systematic pattern in his timing organization, namely CVCCVV and CVCVV.

6.5.3 C2/V1 and V2/V1 ratios

The C2/V1 ratios of single C2 generally have become smaller while those of geminate C2 have become greater for AP3 (Appendix 6.15). This is due to overexaggeration of long/short C2 differentiation. The C2/V1 ratios of the KATE and RIKA series are 1.52/4.56/2.31 and 1.74/4.58/2.30 for AP3; the corresponding ratios were 1.69/2.00/2.24 and 2.30/3.27/3.12 for AP1. Thus the differentiation of the C2/V1 ratios of single C2 and geminate C2 is greater in AP3. In fact, the ratios in AP3 show a greater degree of differentiation than the corresponding ratios for NS, which are 1.68/3.21/1.62 and 1.31/2.82/1.13 respectively.

The average duration of V1 tends to be longer before a long nasal C2 in AP3 (this is similar to NS), while the other learners tend to show longer average V1 duration before a short nasal C2. Amongst all the subjects, the tendency towards longer V1 was most prominent for PK3 who demonstrated a long V1 in all the RIKA series.

With regard to V2/V1 ratios, AP3 demonstrates a similar tendency observed in AP2 (see 6.4.3); the V2/V1 ratios of CVCVV are greater than those of NS due to overexaggeration of long V2. For example, they are 2.43/2.60/5.56 and 2.00/2.41/6.47 for KATE and ISO series respectively, while they are 1.82/1.37/3.99 and 1.72/1.33/3.75 for NS. The V2/V1 ratios of CVCVV have become greater than those in AP2 in both KATE and ISO series.

Interestingly, the V2/V1 ratios in DS3 and PV3 still demonstrate the same problem found in DFAT1 and DFAT2. They demonstrate greater V2/V1 ratios for CVCV and CVCCV in the word with an LH pitch pattern. The V2/V1 ratios for the KATE series are 2.88/6.07/4.29 (DS3) and 7.01/7.89/7.37 (PV3) respectively, while they are

1.82/1.37/3.99 for NS. The V2/V1 ratios are too big compared to those of NS. This tendency is more prominent in DFAT3 than in DFAT1 in these speakers' data. ISO series also demonstrate a similar problem where the V2/V1 ratios are 3.13/5.80/4.69 (DS3) and 5.53/6.30/3.78 (PV3), while they are 1.72/1.33/3.75 for NS.

6.5.4 VOT

VOT results in DFAT3 does not show a clear relationship either with the place of articulation or with gemination of the C2. This result is the same as that of the ANU beginning-level speakers.

VOT for [t] and [k] were not statistically significant (t=-.084, p=.9332), and the corresponding geminate C2 shows the same tendency (t=-.684, p=.501). The gemination of the C2 does not affect VOT for either [t] (t=1.926, p=.0625) or [k] (t=.474, p=.6384).

	t	k	t-test
C	40.52 msec (11.69)	40.88 msec (17.55)	t=-.084, p=.9332
CC	32.62 msec (11.41)	37.63 msec (22.69)	t=-.684, p=.501
t-test	t=1.926, p=.0625	t=.474 p=.6384	

Table 6.3 VOT

Average VOT in the present data is still significantly longer than that of NS in all examples. In comparison with the data obtained from the first experiment, the average VOT decreased in DFAT3.

The results of dependent t-test show that this decrease is significant in KATE (t=2.09, p=.0422) and RIKA (t=2.244, p=.0297), and not statistically significant in KATTE (t=.626, p=.5377) and RIKKA (t=.823, p=.4191). Therefore, it can be said that the value for VOT decreased significantly in CVCV and did not decrease significantly in CVCCV during the one year's intensive course. This means that the

VOT of DFAT learners did not become similar to that of NS nor advanced-level learners.

6.5.5 Pitch assignment

Some errors are still observed in pitch assignment. For example, PK3 and PV3 show a consist pattern of assigning the first CVC-syllable one tone, such as [kom-ma] HH-L, [hon-ne] LL-H and [saŋ-ŋa] HH-L. In other words, these learners have not been able to develop the concept of a moraic nasal C2 as a tone carrier during their one year's intensive training. DS3 also has a similar problem with pitch assignment for the HONE series. The assignment of an LH pitch within a CVC-syllable (e.g. [hon-ne] LH-H) appears very difficult for the learners. It seems that more difficulty is involved in words with an LH pitch pattern than in those with an HL pitch pattern.[41]

AP3 showed no problems with pitch assignment. This is an improvement from AP1 where errors such as [hon-nee] LL-HH were observed.

6.6 Discussion

Three experiments were conducted in order to observe the speech production and perception of a group of four subjects who were undertaking an intensive course in Japanese over the 1993 academic year.

It was found that the results obtained from the acoustic observations in the first experiment bore similarities to those obtained from the ANU learners described in Chapter 4, such as underdifferentiation of the long/short contrast which was mainly due to the longer duration of short phonemes rather than the shorter duration of long phonemes.

The results described in this chapter provide further evidence of the difficulties involved in the timing control of single C2 in CVCV. All four subjects had this problem as evidenced by the results obtained from the first experiment; long C2, on the other hand, was produced correctly in a number of cases. The findings obtained in

---

[41] No speakers showed anticipatory assimilation caused by the H tone on the second syllable such as [honne] with an HHH pitch pattern, unlike NS.

this chapter provide further evidence that the timing organization of the basic CVCV sequence in Japanese is influenced by language-specific rules, and not governed by 'universal rules' (Chomsky and Halle 1968:295) as discussed in Chapter 4.

The results obtained from the three experiments revealed some interesting changes at different stages of the learners' acquisition of timing control. AP's timing control of the CVCV sequence gradually improved during the year. This indicates that Japanese timing organization is learnable. It was found, however, that the durational contrast became more and more distinctive in this speaker's data, to the extent that it was overexaggerated. The degree of AP's durational contrast was less than that of NS in the first experiment, but became greater than that of NS in the third experiment. Overexaggeration of phonological contrasts seems common to L2 learner's pronunciation, and a possible cause of this case is the speaker's awareness of the importance of the phonological contrast. If overexaggeration occurs, however, the speaker's pronunciation may sound unnatural to NS. Therefore, it is important to note that production of the durational contrast alone may not necessarily improve learners' pronunciation if the phonological distinction is too overexaggerated. From the viewpoint of teaching Japanese timing organization, instruction which merely focuses on the production of a clear differentiation may lead some learners, particularly those who are aware of the importance of the phonological distinction and try hard to achieve it, to overexaggerate the phonological contrast. The results obtained in this study imply that it is necessary to teach appropriate timing control rather than just focus on making the differentiation (see Chapter 7 concerning an application to the teaching of Japanese timing organization).

On the other hand, the other speakers seemed to have constructed incorrect patterns in their timing organization. Although their ability to control timing gradually improved during the year, the results of the final experiment were not similar to NS speech because their targets (i.e. the patterns they had constructed for themselves) were different from NS norms. For example, PV appeared to have constructed his own timing pattern which is perceived as CVCCVV[42] by NS and in the third experiment, all examples with an obstruent C2 have become CVCCVV (e.g. all of [rika], [rikka] and [rika] became [rikkaa]). This tendency was also found in PK's KATE series in the third experiment.

Although it is very difficult to account for this tendency, one possible reason is a decrease in the speakers' awareness of the importance of the durational contrast in Japanese. According to the course participants, phonological aspects of Japanese durational contrast were only taught at the beginning of their course. The learners had just been introduced to the phonological durational contrast when the first experiment was conducted, and thus the level of learners' awareness was high. No systematic practice which focuses on pronunciation was given to this group of learners for the rest of the year. Therefore, it is possible that when the speakers became less conscious of the phonological distinction, all CVCV, CVCCV and CVCVV merged to what an NS would perceive as CVCCVV. If this is the case, the results obtained in this study indicate the necessity for instructors to remind learners of the importance of the phonological contrast throughout the year, to maintain their awareness.

In the first experiment, PV1 seems to have interpreted CVCCV sequences as CVC-CV (e.g. [rik-ka]) in order to make the durational contrast. As discussed in the previous chapter, this could be interpreted as a strategy which reflects the speaker's underlying knowledge of geminate consonants occurring at morpheme boundaries in English (e.g. cat tail, rock cake), and this interpretation may have helped this subject to produce long C2. By the third experiment, however, it does not appear that this strategy was applied. Instead, PV3 tried to equalize word duration and the long/short contrast. This shows that learners' speech production and the strategies they use are continually changing.

With regard to VOT, it was found that although the average VOT gradually became smaller in some speakers' speech towards the end of the year, this decrease was not always statistically significant. VOT was still too long compared to that of NS and

[42] The C2 seems to become short when it is a single nasal C2 (ie. CVNVV).

advanced-level speakers. There was no consistent relationship between VOT and gemination of C2. Also, no evidence was found to support universal hypotheses concerning VOT and the place of articulation because the t-test results demonstrated that VOT for [k] was not significantly longer than that of [t] (the same situation was found with [kk] and [tt]) in DFAT1-3. It is interesting to note that NS and advanced-level speakers' data provided some acoustical evidence that VOT increased as the place of articulation moved back, but this type of phonetic universal was not found in the beginning-level learners' speech production. This means that although the beginning-level learners deliberately manipulate the duration of different segments in their speech production (not necessarily in the same way as NS), they do not incorporate universal principles in their speech production, unlike NS and advanced-level speakers.

Pitch assignment to words with a nasal C2 was a problem for many subjects. Most of the errors were related to the assignment of a tone to the first syllable. The pitch pattern of the HONE series (LH) was especially difficult for the learners.

Without longitudinal observations, the gradual increase in the rate of overexaggeration by AP, or the development of unique timing patterns by other speakers, would not have been discovered. A cross-sectional observation can only capture the characteristics of learners' speech produced at a particular time, and has its own limitations. It is necessary to examine longitudinal data when developmental aspects of language acquisition are discussed. The results obtained in experiment 1-3 demonstrated that acquisition processes of individual learners were completely different, consequently, they imply that the learners need to be instructed individually in the classroom to improve their pronunciation.

6.7 Conclusion

In this chapter, the longitudinal study was conducted and various findings were obtained which were unique only to the longitudinal observations. It is recommended that longitudinal data collection be conducted in order to observe independent learner strategies and the ways in which they change over time as described in this chapter.

Chapter 7 An application to the teaching of Japanese timing organization

This chapter describes a pronunciation lesson given to learners in order to investigate the extent to which efficient language teaching facilitates the acquisition of Japanese timing organization.

One of the fundamental questions addressed in the previous literature on SLA is whether or not second language instruction actually helps learners (see Long 1983 for review). It is generally believed that efficient language instruction assists the natural process of language acquisition, and that it facilitates rather than impedes learning (Krashen 1982:37). Pienemann (1985), however, claims that certain aspects of the given target language are not learnable if learners are not ready for them. Based on this assumption, he suggests 'Do not demand a learning process which is impossible at a given stage' (1985:63).

Learnability of a given target language depends on different aspects of the language. Long (1993) claims that the critical period for learners to acquire an NS-like pronunciation (i.e. phonology) is six years, whereas L2 morphology and syntax can be acquired before the age of fifteen. An interesting example given by Shibata (1958) may provide supporting evidence for this claim. The accent patterns of Japanese children evacuated from Tokyo to country areas were investigated after the Second World War. Amongst five hundred Japanese children, those who were younger than six years old when they moved to rural areas acquired the local accent, while those who were thirteen years or older retained their Tokyo accent.

While the acquisition process itself may be affected by the age of learners, we cannot simply assume that children learn faster than adults. In this sense one cannot say that younger is better (see Krashen et al. (1979) for the effect of age upon SLA). It is generally assumed that children can access their acquisition device in L2 learning, while adults tend to use learning strategies which are similar to ones employed in general problem-solving (Ellis 1990).

The question addressed in this chapter is to what extent Japanese timing organization is teachable to adults. A pronunciation lesson was given to a group of adult learners to investigate the role of instruction in the acquisition of Japanese timing control and perceptual categorization. Acoustic investigation was then conducted of both their speech perception and production after the lesson, in order to observe the extent to which instruction helped learners to improve their timing control. Additionally, identification tests on learners' pronunciation were conducted both with learners and NS. This was to examine how NS evaluation of learners' pronunciation, and learners' own evaluation of their pronunciation, had improved after the lesson.

As indicated in Chapter 1, this is an additional chapter describing one application suggested by the research findings from the experiments reported in the main body of this thesis. Because of limitation of space, it is beyond the scope of this study to discuss in detail the effects of different teaching methodologies.

7.1 Problems

Students of Japanese tend to make common errors concerning the perception and production of the following utterances:[43]

1. Kite kudasai.
   HL
   "Please come."
2. Kitte kudasai.
   HLL
   "Please cut (this)."
3. Kitte kudasai.
   LHH
   "Please give me a stamp."
4. Kiite kudasai.
   LHH
   "Please listen."

[43] The intended meaning of such utterances may be understood from their context in real-life situations. As mentioned in Toda (1993), however, learners (especially beginners) often encounter problems in context independent situations, such as when they ask the meaning of Japanese words in the classroom. I have experienced occasions where beginning learners have asked "What does ..... mean ?" or "How do you use .....?" In such situations, I did not understand which word they meant due to the fact that they confused long/short contrasts and/or pitch patterns.

Recent work by Han (1992:102) points out that:

> The duration of geminate consonants in Japanese intrigues language students, instructors and linguists for practical as well as theoretical reasons. Students of Japanese as a foreign language are puzzled and frustrated as they are repeatedly corrected on the pronunciation of geminate consonants. Experienced instructors know that the timing control of geminate consonants in Japanese is one of the most difficult phonetic skills for the students to master. For practical purposes, both students and teachers need to understand exactly how native Japanese speakers make the geminate/single contrasts. Imperfect timing control of these sounds gives even fluent speakers a foreign accent.

Han (1992) observes the production of phonemically contrastive long and short stops by advanced-level American learners of Japanese. She claims that the results obtained from her experiments indicate that Japanese NS show a clear durational contrast between long and short consonants with an average ratio of 2.8:1, while American learners have a ratio of 2.0:1 and thus underdifferentiate the contrast. Han concludes that the durational contrast in stops is one of the most difficult aspects of Japanese phonology for NNS, and comments that 'A suggestion to "make the geminate consonants 3 times as long as the single ones" may have practical, pedagogical merit' (1992:126).

In Chapter 4, it was argued that this suggestion was too unidimensional to capture the dynamic reality of the durational contrasts in the spoken language. It was found that underdifferentiation was only one of the processes involved in the learners' speech production, and that focussing only on durational contrast would not lead to a solution.

The question is, then, what other alternative methods can be applied to help learners improve their timing control. One suggestion proposed in the previous literature is to teach Japanese rhythm using musical notations (Yoshida et al. 1976; Kawakami 1984; Hirata 1988). These offer a graphic means to incorporate several dimensions of the phonetic aspects discussed in the previous chapters, such as duration, pitch pattern and vowel offset quality. Also, introducing the basic concepts of musical notation and usage circumvents the need to introduce many phonetic terms which might make the description unnecessarily complicated for learners.

Previous studies differ in the use of basic musical notation for the description of geminate consonants and whether or not they include the notion of pitch in their description. For example, Yoshida et al. (1976:9) incorporate pitch in their description and indicate the pitch rise in the second mora for all /CV/, /N/ and /Q/ as shown below:

ri N go 'apple' ki Q te 'postage stamp'

Thus, an H tone is assigned to the second mora in both /riNgo/ and /kiQte/, making the pitch pattern of these words LHH. This description is based on the Japanese phonological rule that the pitch assigned to the second mora has to be different from that assigned to the first mora (see Chapter 2). Phonetically speaking, however, the /Q/ in /kiQte/ is a silent closure period of [t] and there is no sound during the closure of this voiceless stop. In this respect, the phonetic properties of /Q/ are completely different from those of /N/. The phonologically-based musical notation above does not account for this phonetic reality, and additional thought needs to be given to the suggestion that learners raise pitch where there is no sound.

Prior to the present study, I conducted informal interviews with Japanese instructors teaching at tertiary-level institutions in Canberra in order to find out the teachers' awareness of their learners' problems concerning the acquisition of Japanese pronunciation.[44] One university lecturer (NS) said that one of her beginning-level students asked her a question regarding the pitch change marked on Japanese words

[44] The general impression I obtained from the interviews was that there was no systematic practice given on a regular basis for learners to improve their pronunciation in most cases, and this is largely due to lack of time. A number of instructors felt that, as there was a large number of grammatical items which needed to be introduced in the classroom, they could not afford to spend time on pronunciation practice.

in their textbook, Introduction to Japanese (Mizutani & Mizutani 1987). According to this lecturer, the student asked if the pitch went up at the location where the small *tsu* in *kana* orthography was located since the pitch rise was marked on the small *tsu* in the word. (The small *tsu* in *kana* orthography in the above example is an equivalent of a silent closure period of a voiceless geminate obstruents.) When told that the pitch would go up at the location of the small *tsu*, the student looked puzzled and said, "but how?". The teacher realized the problem and told the student that this was the way Japanese NS would *feel*, and that the student should try feeling that way even if there is no actual voicing associated with the pitch rise. This example demonstrates two interesting points: firstly, this learner found the phonologically-based description confusing because it did not agree with the phonetic reality; secondly, even an experienced teacher was not able to give either an appropriate explanation or instructions to help the learner to improve his pronunciation except for referring to the psychological reality of NS. It seems doubtful that this explanation was particularly useful for this student. In L2 teaching, I believe that NS instructors cannot expect learners to understand explanations which are based on the psychological reality of NS, because the learners' L1 is different from their target language.

Additionally, the results obtained in Chapter 4 and 6 of this study show that the duration of V1 tends to be too long in relation to the duration of the geminate C2 in some learners' data. This may imply that if students are too conscious of producing the suggested pitch pattern, their awareness may lead them to prolong the duration of the V1 even more as a result of the pitch adjustment (i.e. from L to H, or H to L) from the first to the second mora. This is a situation which teachers should try to avoid in students' pronunciation.

For these reasons, musical notations which indicate pitch rise during the silent closure period of a voiceless geminate stop do not seem to be the most appropriate way of introducing the practice of the timing control and pitch pattern of geminate stops to learners with a different language background (i.e. who do not have the same phonology as Japanese NS in their L1). Therefore in the present lesson plan, indication of pitch rise associated with geminate stops used in Yoshida et al. (1976) is avoided.

By contrast, neither Kawakami (1984) nor Hirata (1988) incorporate pitch patterns in their musical notation and focus only on durational aspects. Making a clear distinction between descriptions of the silent closure period of a geminate consonant (staccato crochet) and moraic nasals or long vowels (crotchet) appears a valid strategy to avoid learners' confusion. At the same time, the fact that these descriptions focus purely on durational aspects, with pitch patterns not being taken into account seems somewhat biased.

In the previous chapters, I have argued that the acquisition of correct pitch patterns is challenging for both beginners and advanced learners. It was claimed that more practice on this aspect is needed since the acquisition of correct pitch assignment is as crucial as that of timing control. Therefore multi-dimensional descriptions which also incorporate pitch patterns seem preferable for the improvement of learners' pronunciation. In the text used here for pronunciation practice, pitch is therefore incorporated together with duration (see Appendix 8.1).

7.2 Lesson plan

The aim of this exercise was to give learners opportunities to improve their timing control in Japanese with the aim that they would be able to perceive and produce the phonological distinction in similar ways to NS.

The subjects were the ANU beginners who participated in the experiments described in Chapter 3 and 4. After introduction of the use of musical notations (see 7.2.2), the pronunciation exercise was given to students. A classroom instructor read the examples in the text (Appendix 8.1) while students repeated after the instructor. In this lesson, acoustic-phonetic explanations such as 'make the duration of double consonants longer' were avoided. Some body actions (e.g. arm movements) were also used to encourage the participants to learn the prosodic characteristics of the utterances. The Verbo-Tonal Method introduced in Roberge and Kimura (1990) is used as a reference for the types of body movements.

After the practice exercise in the text, the instructor presented some sentences of familiar grammatical patterns, and instructed the students to practice the rhythm of the utterances. For example, 'rokuji ni okimasu' (I get up at six o'clock) was pronounced as 'roku-jini-oki-masu', based on the concept of bimoraic rhythm (Chapter 2).[45]

To provide further individual practice, a Hypercard application was designed specifically for this project (Appendix 8.2). This Hypercard program, labelled 'Let's Learn Nihongo no Rizumu (Let's learn Japanese rhythm)', is used on a Macintosh computer. In this application, the user first clicks the sound button on the screen to hear a word. Two or three words are shown on the screen, and the user clicks one word he or she heard when the sound button was clicked. If the answer is correct, the user hears a 'clink-klank' sound and scores a point, and if it is incorrect, the whole computer screen scrolls down with a slowed down 'boing' sound. The correct and wrong answer button scripts are as follows:

<u>"Correct Answer" Button Script</u>	<u>"Wrong Answer" Button Script</u>
on mouseUp	on mouseUp
play "clink-klank"	play "boing"
visual effect zoom open	visual effect shrink to bottom very slow
go this card	go this card
end mouseUp	end mouseUp

This twenty-page Hypercard stack took ten to fifteen minutes for each learner to use. 'Let's Learn Nihongo no Rizumu' can be used in a computer-lab as a group or used after the class for self study.

Total time required for the pronunciation exercise is one hour.

---

[45] The beginners told me that they found it difficult to treat 'ji (o'clock)' and 'ni (at)' as one unit (at the rhythmic level) due to the fact that there is a morpheme boundary between the two. It is common for beginners to pronounce the sentence with a pause between 'rokuji' and 'ni'. This is also because beginners are often unsure about the appropriate particle to be used in the sentence, and this hesitation leads to inappropriate rhythm (eg. rokuji (pause) ni (pause) okimasu).

7.2.1 Use of musical notation

The musical notation used in the text incorporates the crotchet, quaver and rest. The number of musical symbols is limited to three in order not to make the notation unnecessarily complicated. Three symbols used are as follows:

1. Quaver	♪	short syllable e.g. CV in /ka te/
2. Crotchet	♩	long syllable with no pitch change e.g. CVV in /(yu) soo/
3. Quaver rest	𝄾	first part of geminate C2 e.g. /Q/ in /ka Q te/

A quaver is used to indicate the basic moraic unit (CV or V). For example, *kokoro* 'heart' is represented by three quavers, and the LHL pitch pattern is simply represented by using high and low versions of the note, as follows:

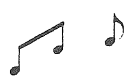
/ko ko ro/ 'heart'

A crotchet is introduced as a basic one beat which is based on the concept of bimoraic rhythm (see Chapter 2). For example, *yusoo* 'transportation' has a long second vowel, but there is no pitch change within this syllable. Thus the first syllable is described in terms of a quaver and the second syllable in terms of a crotchet.

/yu soo/ 'transportation'

The musical convention of 'slur' is used to indicate where a pitch change occurs within a syllable. Thus, /hoN/ (syllable with moraic nasals) is represented as follows:

/hoN ne/ 'true feeling'



The notation used for /hoN/ is the equivalent of two quavers. By introducing this concept, a long syllable with a pitch change is still regarded as one unit rather than a sequence of two separate units. From the viewpoint of English-speaking learners of Japanese, especially at the beginning-level, the concept of mora is difficult to understand and they are more familiar with the concept of syllables since English is a syllable-based language. In addition to making the situation easier for learners, the evidence of NS regarding long syllables as one unit (Nakajoo 1992; Sato 1994) also seems to support such a description.

Finally, a quaver rest is used to indicate the first half of geminate stops.[46] The reason for using the rest is that it symbolises silence in music and thus emphasises the importance of the silent period of stop closure. By introducing this concept, durational contrast is clearly represented:

/ka Q te/ 'selfish'	/ka te/ 'provision, food'
♪ 𝄾 ♪	♪ ♪

7.2.2 Use of minimal pairs

Minimal pairs are used in the test. The sentences used as examples are made short and simple so that students can focus on the durational contrast.

Stevick (1982:59) recommends the use of this technique for teaching of pronunciation as follows:

> After Mimicry-Memorization, the use of 'minimal pairs' is perhaps the most primitive technique for the teaching of pronunciation. And it is indeed a technique for 'teaching' in the narrow sense of promoting 'learning' (Chapter 3) but doing very little directly for 'acquisition'. Even if you are concentrating on acquisition, however, you can use minimal pair drills occasionally in order to sharpen your students' hearing of some of the sounds that they are in the process of acquiring.

I am aware that there are some questions about the use of minimal pairs for teaching pronunciation (Ootsubo 1990), such as that minimal pairs usually have quite different

---

[46] A quaver rest is also used in the description of the first half of long fricative C2. Phonetically speaking, long fricatives such as [ʃ] do not have a silent closure period, unlike stops. Instead, they have a continuous high frequency fricative noise. The description suggested here does not capture this phonetic reality for geminate fricatives since they are described in the same way as geminate stops. It is interesting, however, to note that many Japanese NS seem to regard all of the first part of geminate [p], [t], [k], [ʃ] and [tʃ] as the same unit as mentioned in footnote 9.

meanings and thus do not tend to appear in the same context. It seems, however, that the use of minimal pairs is one systematic and efficient way of clearly demonstrating the importance of phonological distinctions by indicating the relationship between certain phonological contrasts and the corresponding differences in meaning. It is particularly important to increase learners' awareness on this matter at the very early stage of their learning, as discussed in the previous chapters. English-speaking learners of Japanese do not have the same phonological distinction in their L1, and, as I have demonstrated, acquisition of the durational contrasts is a challenging task for learners. Therefore, it is necessary to give learners the opportunity to practice this aspect of Japanese phonology in the classroom. Without a formal instruction on the durational contrast, it seems difficult to achieve this goal. For example, a lesson based on the natural approach (Terrell 1977) may not necessarily give learners an opportunity to practice this important part of Japanese grammar.

The lesson conducted in this section aims to increase learners' awareness of the importance of the durational contrast, and to facilitate formal learning rather than natural acquisition based on L1.

7.2.3 Order

The presentation of the example is in the following order: minimal pairs in isolation (e.g. [tʃizu] 'map': [tʃiizu] 'cheese'); minimal pairs together joined by *to* 'and' (e.g. [tʃizu to tʃiizu] 'map and cheese'); simple sentence practice at the end of each example (e.g. *Teeburu no ue ni* [tʃizu] *ga arimasu.* 'There is a map on the table.' *Teeburu no ue ni* [tʃiizu] *ga arimasu.* 'There is some cheese on the table.') and combined sentence practice at the end of the groups. All examples are arranged in this sequence so that students can start from an easy level to a more complex and difficult level.

The instructor reads out the examples in the given order, then the learners repeat each item. The instructor checks each student's pronunciation and makes sure that they can handle word-level production before they move to the sentence practice.

7.2.4 Style

Sentence exercises are in either plain or *-masu* style.

The target minimal pairs are within a similar sentence structure, for example: *Tsukue no ue ni chizu ga arimasu.* 'There is a map on the table.' *Tsukue no ue ni chiizu ga arimasu.* 'There is some cheese on the table.' This is so learners can focus on the difference between the main nouns: *chizu* and *chiizu*.

7.3 Perception

The results obtained after the pronunciation practice sessions are termed ANU2, and are presented in the following section. ANU2 results are compared with ANU1 results which are obtained in the first experiment before the lesson.

Vowels

CV07CV_	KA07TE_	RI07KA_	I07SO_
A.S	168%	186%	159%
S.D	25%	23%	31%
D.S	193%	191%	177%
S.D	30%	31%	27%
CV14CV_	KA14TE_	RI14KA_	I14SO_
A.S	172%	192%	171%
S.D	27%	21%	22%
D.S	199%	206%	185%
S.D	30%	28%	32%

Obstruents

CV07C_V	KA07T_E	RI07K_A	I07S_O
A.S	162%	154%	161%
S.D	32%	40%	33%
D.S	195%	193%	186%
S.D	42%	44%	39%
CV14C_V	KA14T_E	RI14K_A	I14S_O
A.S	157%	157%	166%
S.D	29%	36%	50%
D.S	178%	194%	159%
S.D	39%	38%	43%

Nasals

CV07N_V	KO07M_A	HO07N_E	SA07N_A
A.S	73%	79%	80%
S.D	11%	19%	25%
D.S	86%	79%	84%
S.D	11%	16%	16%
CV14N_V	KO14M_A	HO14N_E	SA14N_A
A.S	85%	74%	80%
S.D	21%	24%	22%
D.S	82%	76%	71%
S.D	13%	20%	17%

Table 7.1 Average Values - ANU2

A.S: ascending series

D.S: descending series

S.D: standard deviation

CV07CV_	KA07TE_	RI07KA_	I07SO_
T.V	181%	189%	168%
A.D	276.89	247.88	260.06
CV14CV_	KA14TE_	RI14KA_	I14SO_
T.V	186%	199%	178%
A.D	284.56	261.69	275.54
CV07C_V	KA07T_E	RI07K_A	I07S_O
T.V	179%	174%	174%
A.D	204.74	218.09	184.08
CV14C_V	KA14T_E	RI14K_A	I14S_O
T.V	168%	176%	163%
A.D	192.12	220.60	172.41
CV07N_V	KO07M_A	HO07N_E	SA07N_A
T.V	79%	79%	82%
A.D	167.95	143.17	133.09
CV14N_V	KO14M_A	HO14N_E	SA14N_A
T.V	83%	75%	76%
A.D	176.46	135.90	123.29

Table 7.2 Average Threshold Values - ANU2

T.V: Average Threshold Values of ascending and descending series (%)

A.D: Absolute Duration (msec)

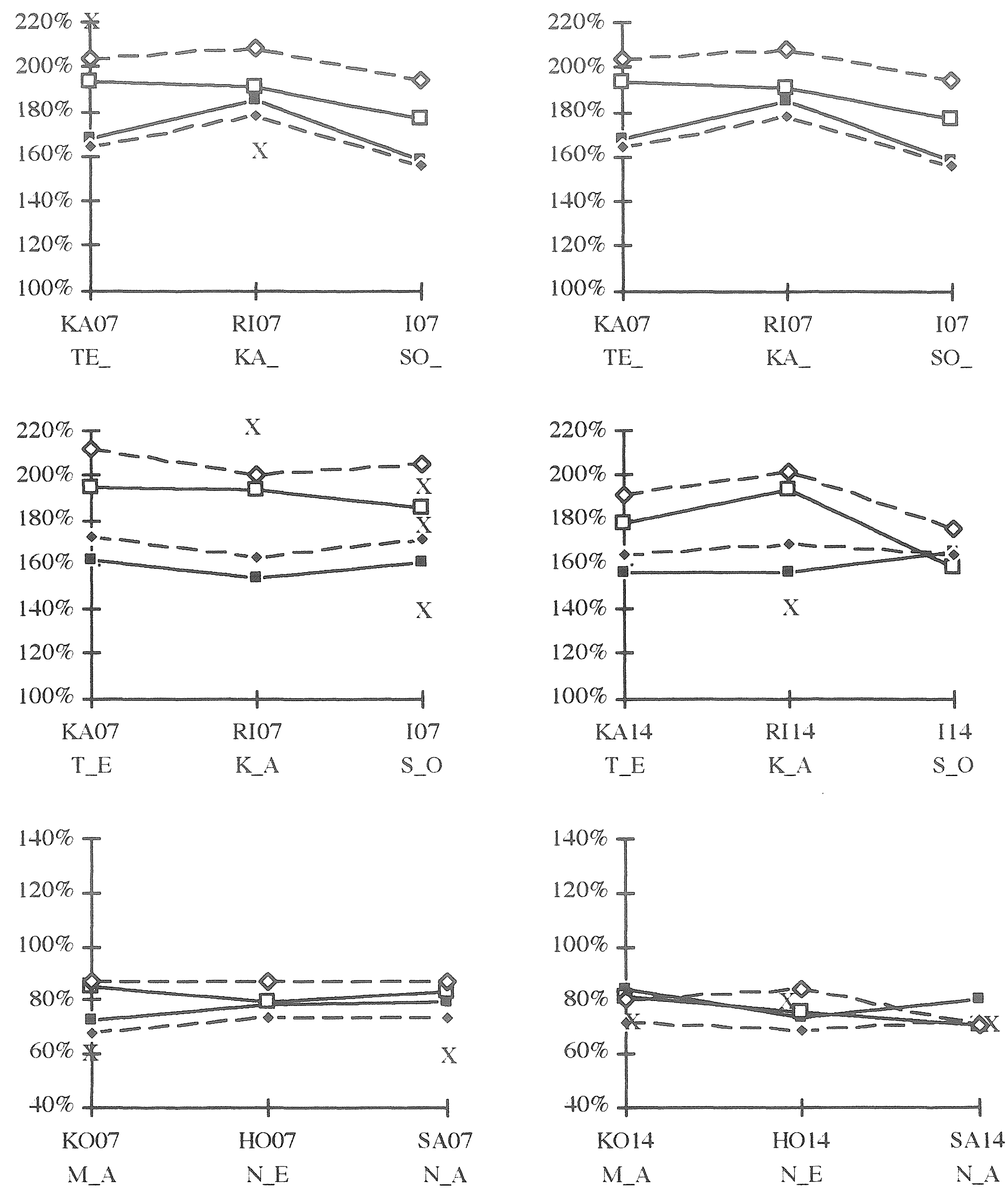


Fig.7.1 Graphs for perception by ANU2[47]

Ascending (black), Descending (white)

The results obtained before the lesson for the same data are indicated with dotted lines for comparison.

[47] X in the graphs shows where learners indicated a question mark before JND. Appendix 7.1 shows the values for which question marks were indicated by ANU2 learners.

7.3.1 Threshold values

Threshold values obtained in ANU2 are still greater than those of NS. This means that the learners' perceptual target is still inaccurate, and that the duration of C2 and V2 has to be longer to be perceived as a long phoneme by ANU2 than by NS.

Dependent t-tests were conducted to observe if the differences in the average threshold values before and after the lesson were significant (Appendix 7.2). It was found that in most examples, the differences were not statistically significant. KA07T_E and I07S_O are the exception to this; threshold values after lesson have become significantly smaller than those before the lesson.

These results imply that the lesson was effective in respect of modification of learners' perceptual targets only for KA07T_E and I07S_O, and that it cannot be said that the lesson was effective for all examples.

7.3.2 Ascending and descending series

In ANU2, the distance between the ascending and descending series has generally become closer than in ANU1, which may be interpreted as a reduction in the degree of uncertainty. In other words, the degree of perceptual categorization has improved after the lesson. This situation is particularly noticeable in examples with V2 and nasal C2.

The differences between the ascending and descending series were not statistically significant in a number of examples; RI07KA_, RI14KA_, I07SO_, I14SO_, KA14T_E, I14S_O, KO14M_A, HO07N_E, HO14N_E, SA07N_A and SA14N_A (see Appendix 7.3 for the t-test results). This differs from ANU1 where the differences between the ascending and descending series were statistically significant in most examples. Thus ANU2 has much fewer examples in which the differences were significant by comparison with ANU1. These results imply that the amount of uncertainty has decreased compared to those obtained before the lesson.

The distance between the ascending and descending series in ANU2 has become smaller than in ANU1 (for example, the perception of nasal C2 where the distance between the two series in ANU2 is even smaller than that of NS), but is still greater than that of NS in respect of the obstruent C2.

Generally speaking, the order of the ascending and descending series still shows the same tendency as ANU1, that is, the descending series has higher values than the ascending series.

7.3.3 V1 duration

2-factor Anova results show that none of the examples with an expanded V1 duration had significantly greater threshold values than those with a shortened V1 duration in the ANU2 data (Appendix 7.4). This is similar to the results which were obtained in ANU1.

7.3.4 Discussion

1. The differences in the threshold values are generally not statistically significant before or after the lesson, and therefore it can be concluded that the learners' perceptual targets were not modified. It is thus not possible to say that the mechanism of learners' perception has become similar to that of NS after the lesson.

2. The distance between the two series, on the other hand, became smaller, and a number of examples showed that the differences between the two series were not statistically significant. This may be interpreted as a reduction in the degree of uncertainty. In other words, the degree of learners' categorization has improved and in this respect, it can be said that the lesson was effective.

3. V1 duration did not affect learners' judgment either before or after the lesson.

It appears that learners' perception first becomes more categorical, and then the perceptual boundaries become similar to those of NS. These two processes do not seem to occur simultaneously. Additional evidence for this claim are the results obtained from the advanced learners' experiments (Chapter 4) and those obtained from the longitudinal study (Chapter 5). At the end of their one-year intensive course, the beginners' categorical perception had improved but their perceptual boundaries showed no statistical change. On the other hand, advanced learners' perceptual boundaries shared more similarities with those of NS. Some reasons which may account for the above results are:

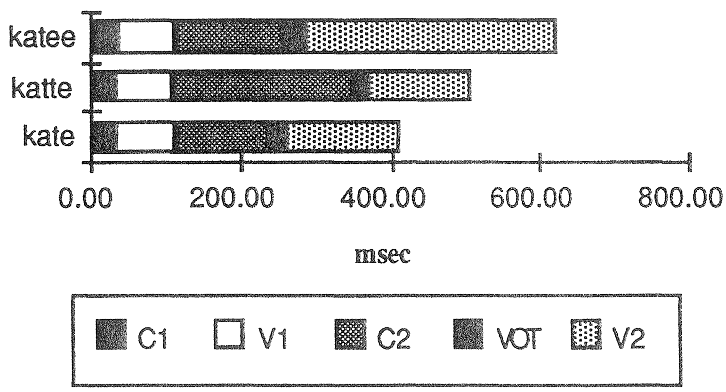
1. The beginning to intermediate level learners were not yet in the acquisitional stage (Pienemann 1985) of NS perceptual boundaries. When this experiment was conducted, they were in an acquisitional stage in which they were able to improve their degree of categorical perception, but not reach the NS norm of perceptual boundaries. This assumption is based on the finding that the acquisition of a higher degree of categorization precedes that of NS perceptual boundaries. If this is the case, it is worth investigating the details of intermediate to advanced level learners in order to find out when learners reach the acquisitional stage of NS perceptual boundaries.

2. The time spent for the lesson was not sufficient to achieve the expected results. If additional lessons had been conducted and the learners given opportunities for more practice, they may have improved both degree of categorization and their perceptual boundaries. The relationship between time spent in class on pronunciation exercises and the acquisition of perceptual boundaries needs further research.

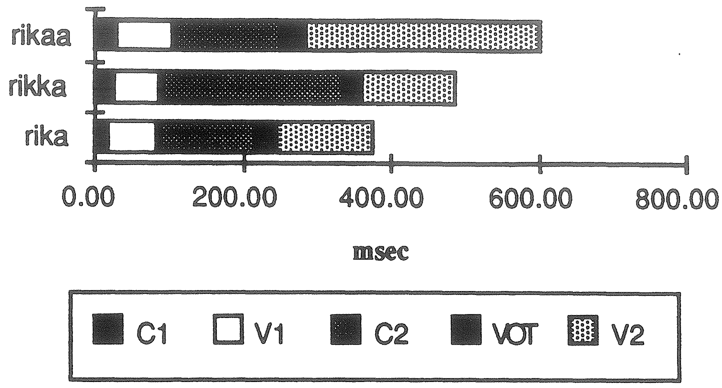
7.4 Production

My auditory impression of the speech production of ANU2 is shown in Appendix 7.5. The average values are presented in Appendix 7.6, and the average figures of ANU2 are shown below.

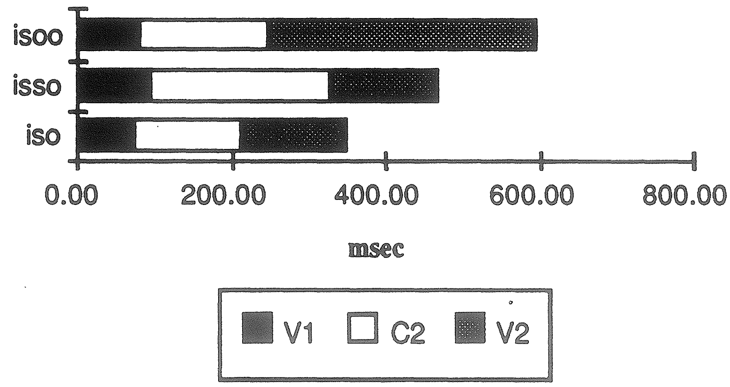
ANU2 Average (1)



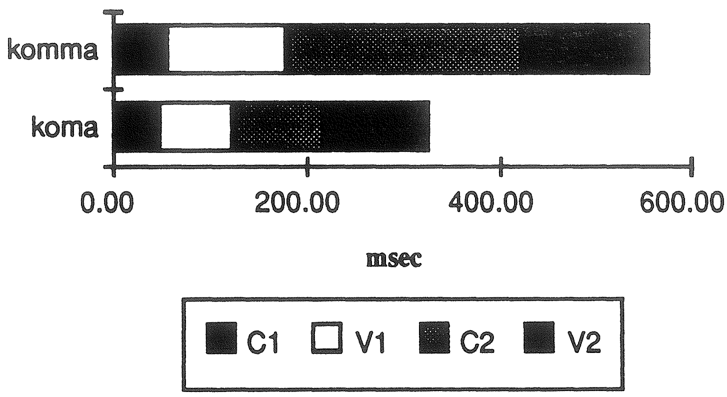
ANU2 Average (2)



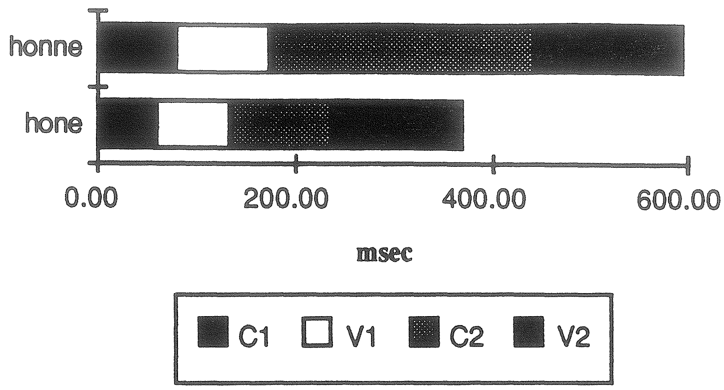
ANU2 Average (3)



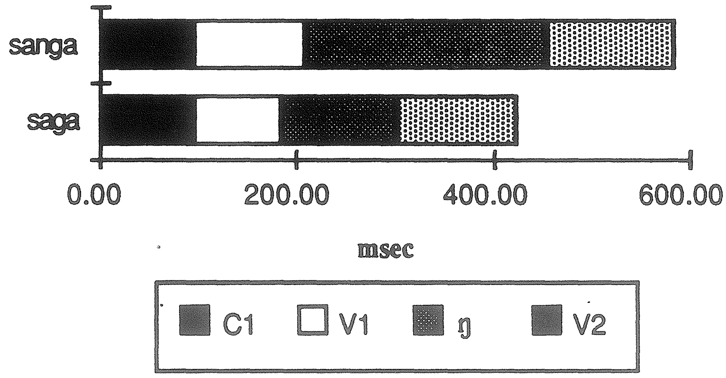
ANU2 Average (4)



ANU2 Average (5)



ANU2 Average (6)



ANU2 Average (7)

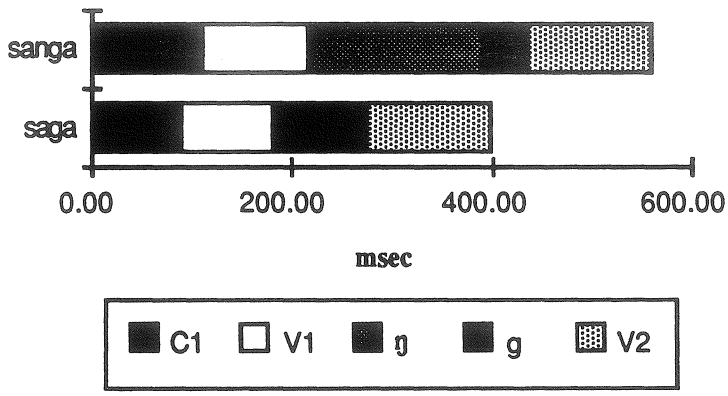


Fig.7.2 ANU2 average (1-7)

7.4.1 Word-duration

SERIES	RATIO
kate/katte/katee	100:123:152
rika/rikka/rikaa	100:130:160
iso/isso/isoo	100:134:170
koma/komma	100:170
hone/honne	100:161
saŋa/saŋŋa	100:129
saga/sanga	100:141

Table 7.3 Word duration (ratio)

Average: ANU2

In ANU1, it was found that underdifferentiation of word duration between two- and three-mora expressions was a common problem for beginning learners. The ANU2 data obtained after the pronunciation practice demonstrate the following tendencies:

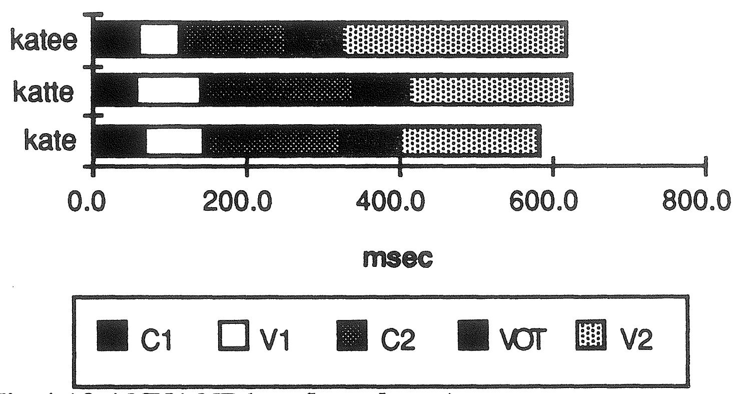
1. The problem of underdifferentiation has diminished. This is particularly so for words with a nasal C2. The ratios of differentiation for the KOMA, HONE, SAGA (saŋa/saŋŋa) and (saga/sanga) series were 100:122, 100:119, 100:98 and 100:99 respectively in ANU1. In ANU2, they have become 100:170, 100:161, 100:129 and 100:141 (Table 7.3). From the viewpoint of clear differentiation, ANU2 data may be regarded as showing improvement. Interestingly, when these results were compared to NS results, it was found that the durational contrasts were overexaggerated. The corresponding ratios of NS are: 100:140 (KOMA), 100:150 (HONE) and 100:129 (SAGA). This means that when the learners try to improve their timing control, they may overshoot the NS target and overexaggerate the durational contrast.

2. With regard to words with an obstruent C2, it was found that the degree of differentiation was often different between CVCCV and CVCVV, being greater in CVCVV than CVCCV. This implies that a greater degree of overdifferentiation is associated with the word-final V2 (CVCVV) than the intervocalic C2 (CVCCV) (see below concerning long/short contrast).

The pronunciation lesson had a clear effect upon the degree of differentiation in word duration. It can be said that overall ANU2 data demonstrate some improvements over ANU1 data. It was found, however, that the degree of differentiation after the pronunciation practice varied depending on the segmental structure of the words.

Individuals: ANU2

This section presents some individual subject's data concerning word duration. The following bar graphs show examples of improvement in word duration in ANU2 (Appendix 7.7), demonstrating the changes from ANU1. (The bar graphs of ANU1 results are presented before those of ANU2 results for comparison.) For example, in NP1, the word duration of [kate] was too long in comparison with that of the corresponding three-mora expressions. Thus it was impossible to distinguish two-mora expressions from three-mora expressions on the basis of word duration. In NP2, on the other hand, the word duration of [kate] has become significantly shorter than that of the corresponding three-mora expressions, making the durational contrast between two- and three-mora expressions similar to that of NS. The rest of the examples also show a similar tendency:



(Fig.4.18 ANU1 NP kate/katte/katee)

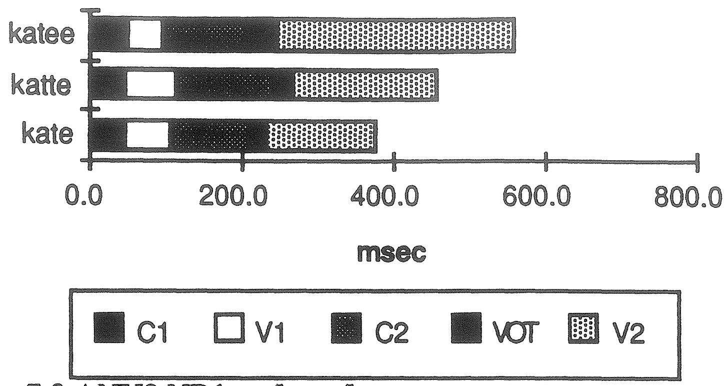
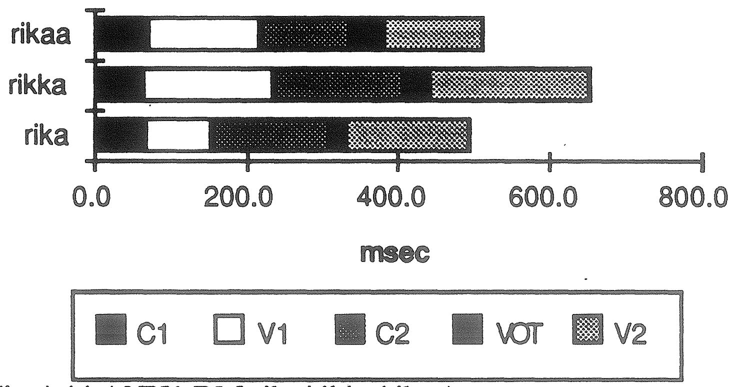


Fig.7.3 ANU2 NP kate/katte/katee



(Fig.4.44 ANU1 BM rika/rikka/rikaa)

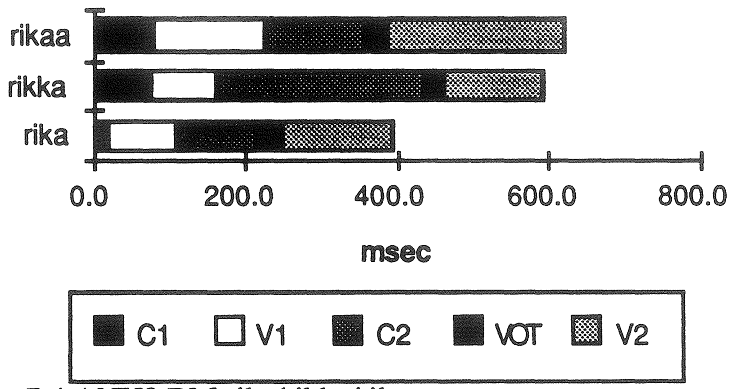
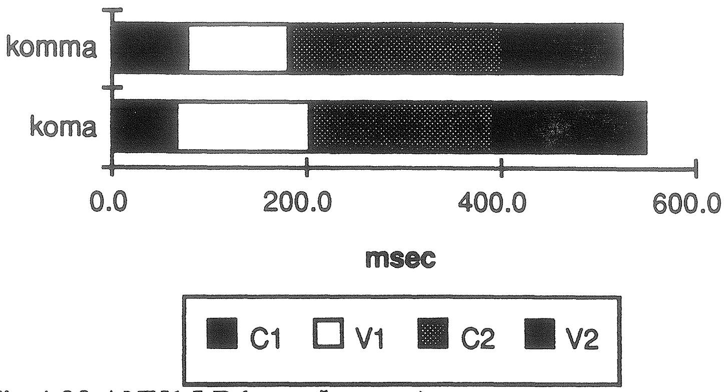


Fig.7.4 ANU2 BM rika/rikka/rikaa



(Fig.4.20 ANU1 LB koma/komma)

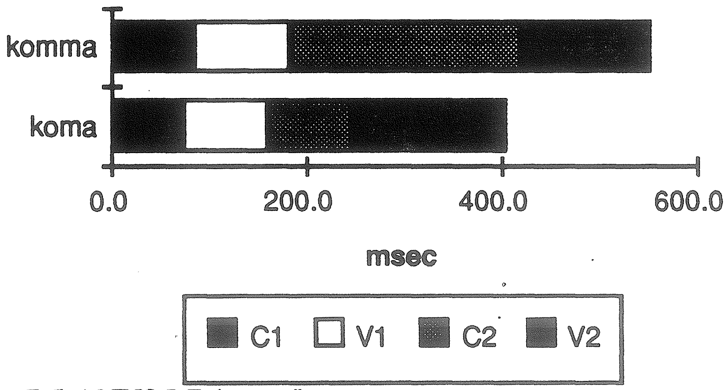
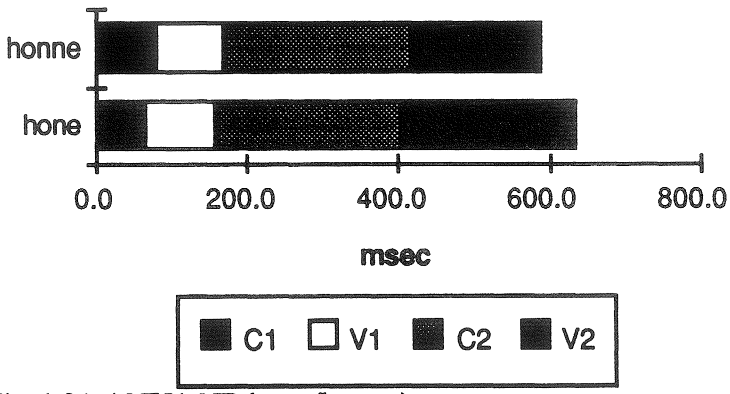


Fig.7.5 ANU2 LB koma/komma



(Fig.4.21 ANU1 NP hone/honne)

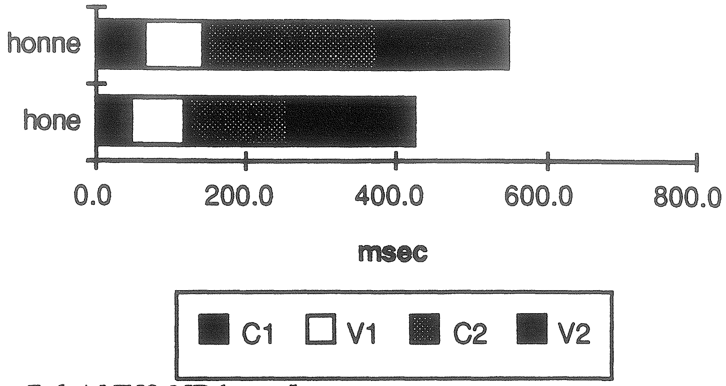
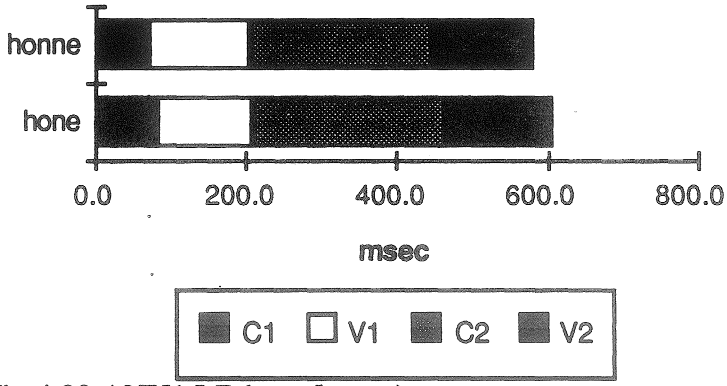


Fig.7.6 ANU2 NP hone/honne



(Fig.4.22 ANU1 LB hone/honne)

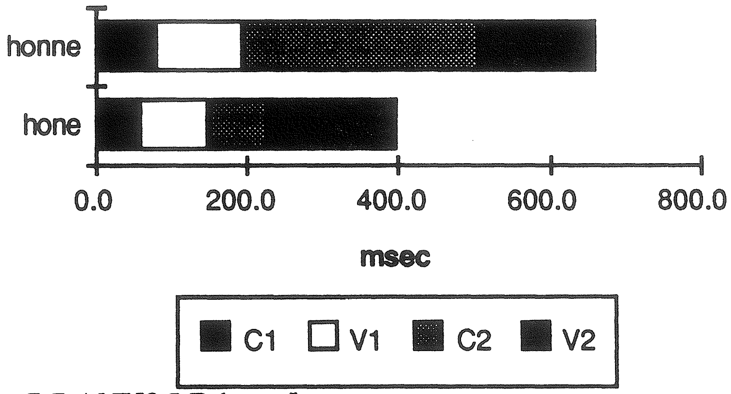
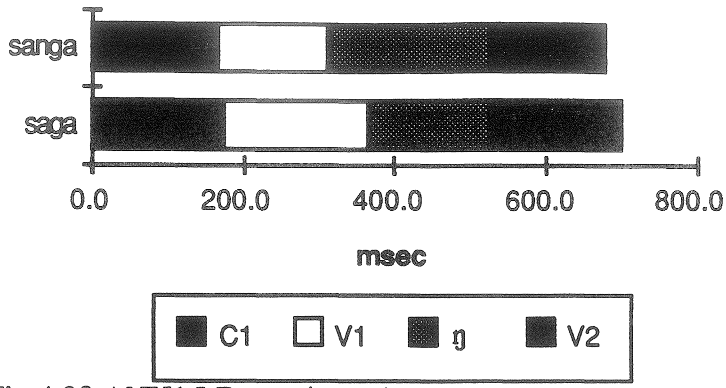


Fig.7.7 ANU2 LB hone/honne



(Fig.4.23 ANU1 LB saga/sanga)

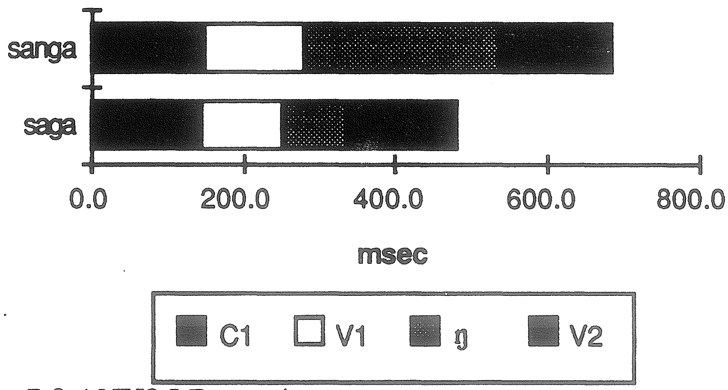
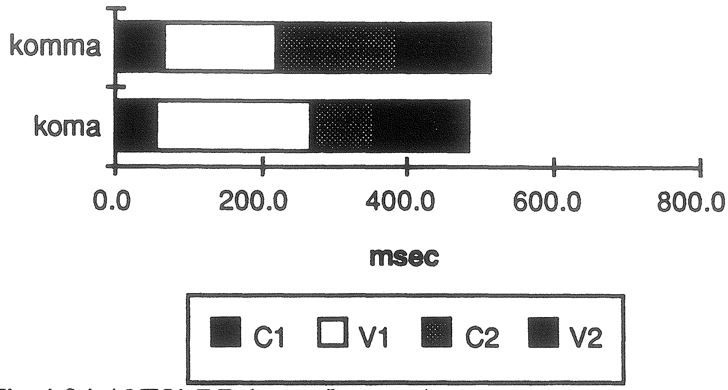


Fig.7.8 ANU2 LB saga/sanga

The word-level durational contrast between two- and three-mora expressions is overexaggerated by some speakers after the lesson. For example, none of the three-mora expressions with a nasal C2 was significantly longer than the corresponding two-mora expressions in BD1. In BD2, however, the duration of the three-mora expressions is overexaggerated and the difference between the two- and three-mora expressions is much greater than NS. PF 1/2 also show a similar tendency. Interestingly, following the lesson, the duration of the first syllable in two-mora expressions for some students tended to be shorter than that of the first syllable in three-mora expressions. To my ear, some speakers sounded as if they were trying to produce two-mora expressions 'faster' than the corresponding three-mora expressions. This may be one strategy that learners applied in order to make word-level durational contrasts.



(Fig.4.24 ANU1 BD koma/komma)

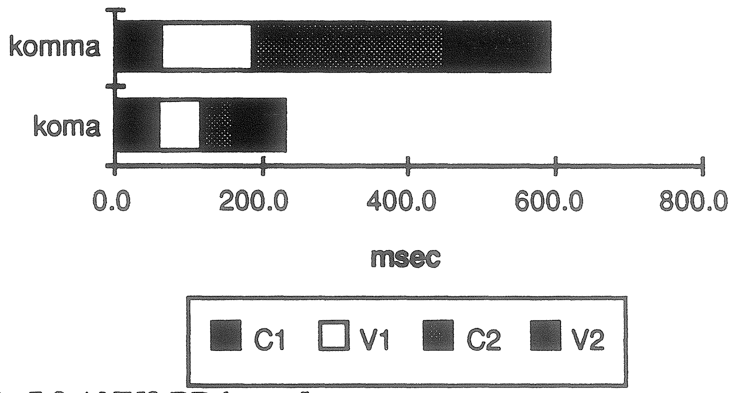
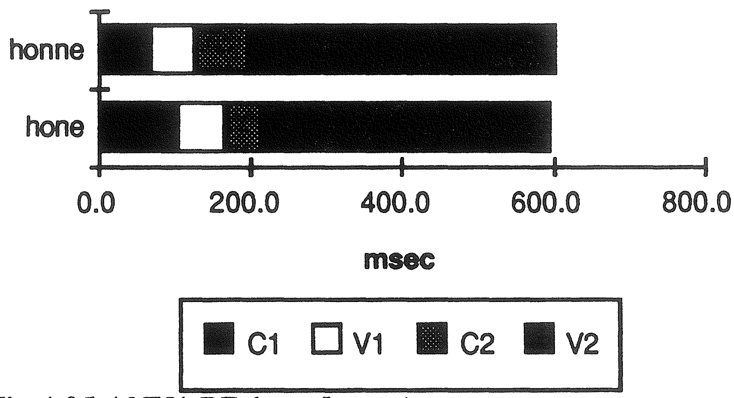


Fig.7.9 ANU2 BD koma/komma



(Fig.4.25 ANU1 BD hone/honne)

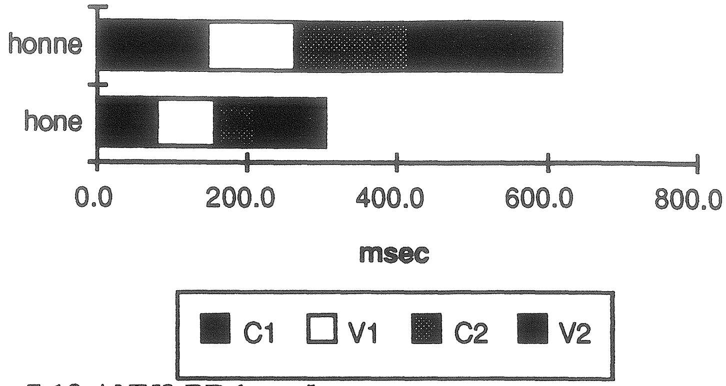
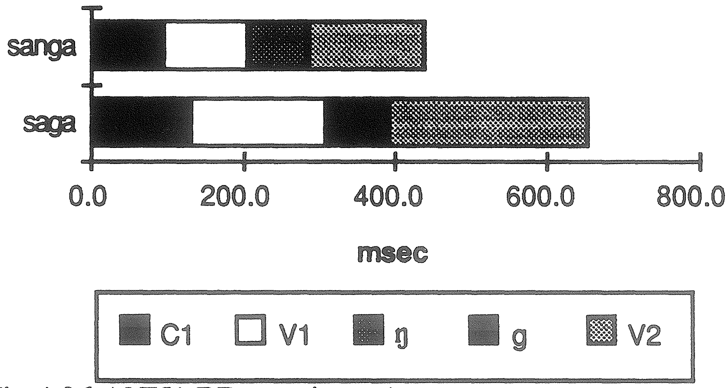


Fig.7.10 ANU2 BD hone/honne



(Fig.4.26 ANU1 BD saga/sanga)

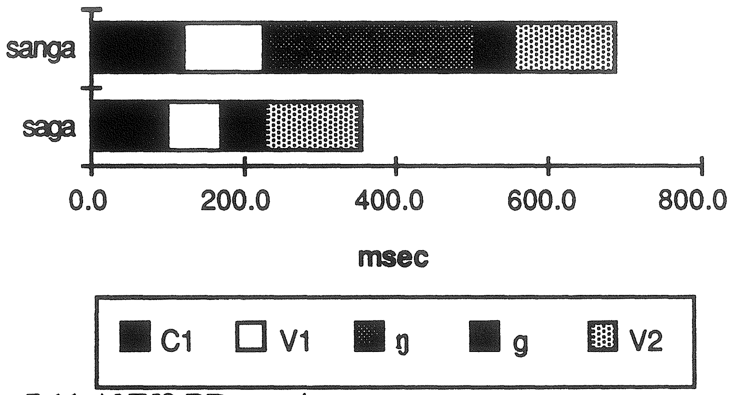
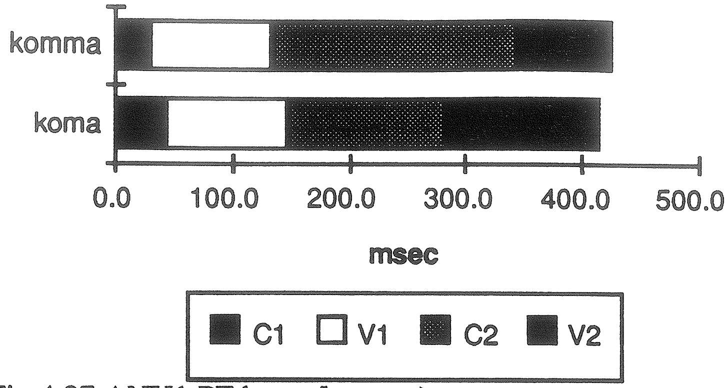


Fig.7.11 ANU2 BD saga/sanga



(Fig.4.27 ANU1 PF koma/komma)

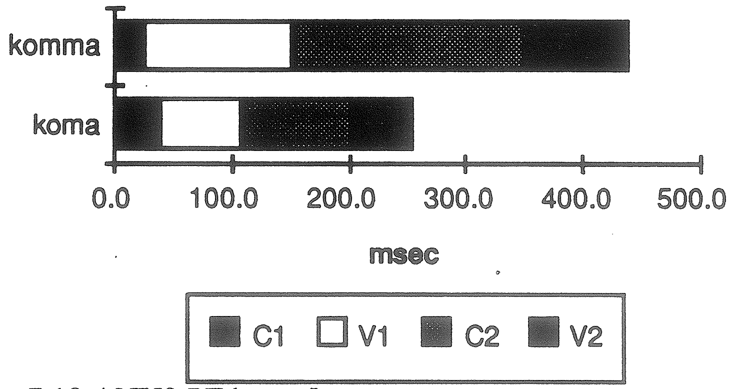
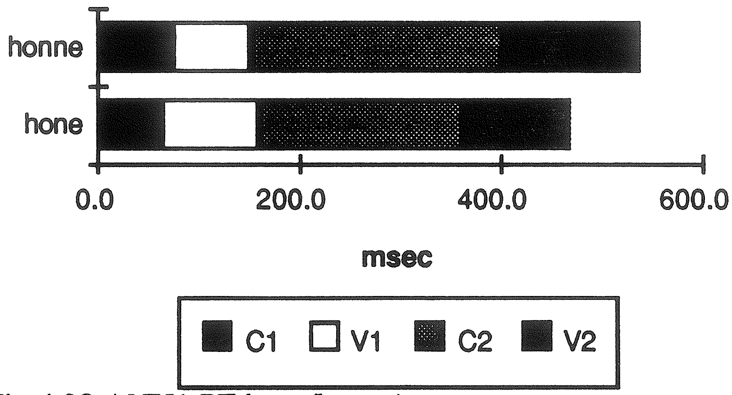


Fig.7.12 ANU2 PF koma/komma



(Fig.4.28 ANU1 PF hone/honne)

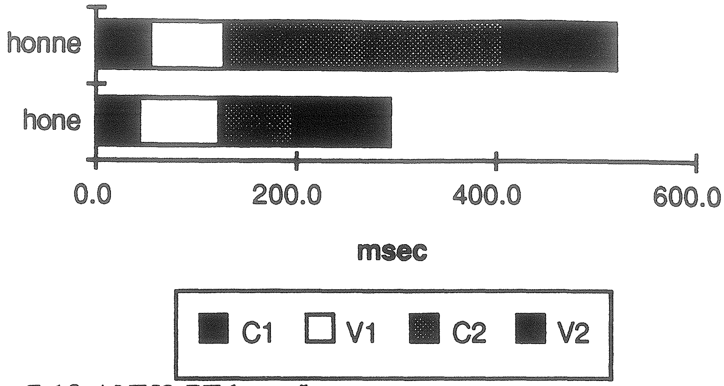
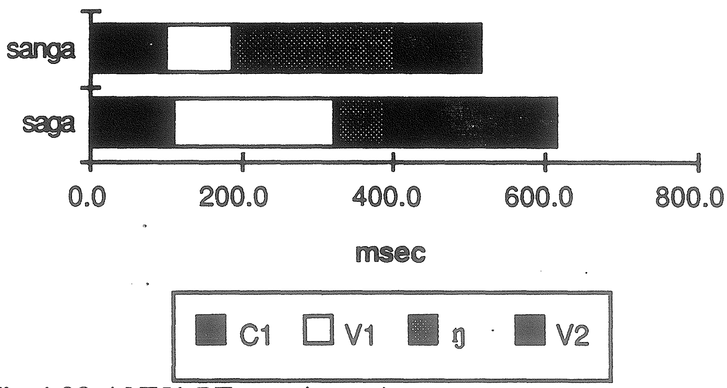


Fig.7.13 ANU2 PF hone/honne



(Fig.4.29 ANU1 PF saga/sanga)

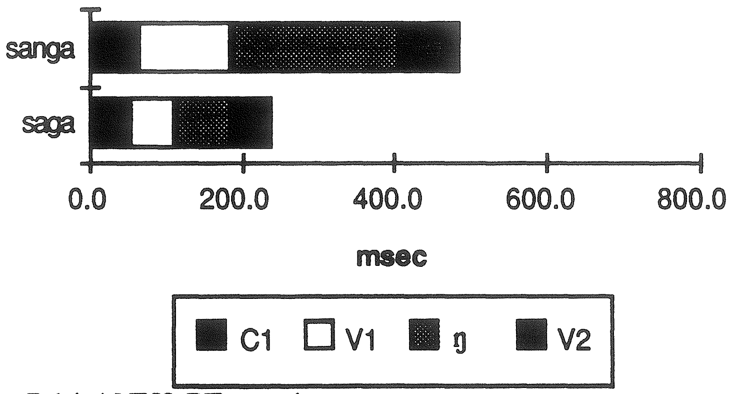
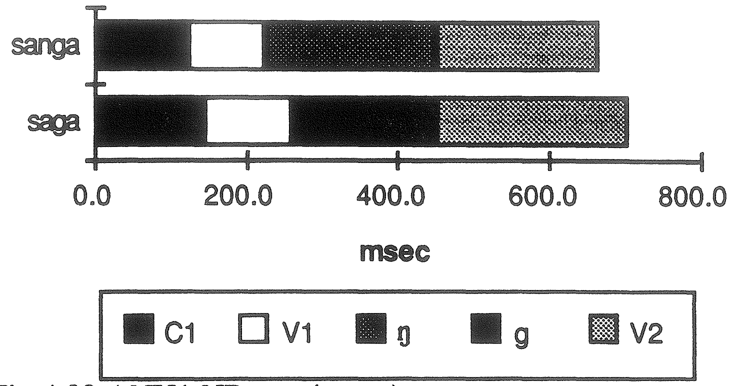


Fig.7.14 ANU2 PF saga/sanga

In Chapter 4, it was found that some learners used different sounds, [g] and [ŋ], in order to distinguish the two- and three-mora expressions in the SAGA series rather than manipulating the segmental duration to produce the appropriate durational contrasts with the same sound. After the lesson, some improvements were observed in the durational contrasts of these learners' speech as shown below.



(Fig.4.39 ANU1 NP saga/sanga)

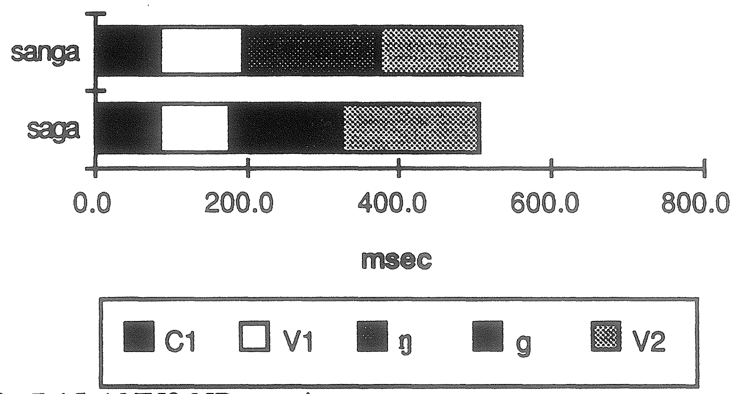
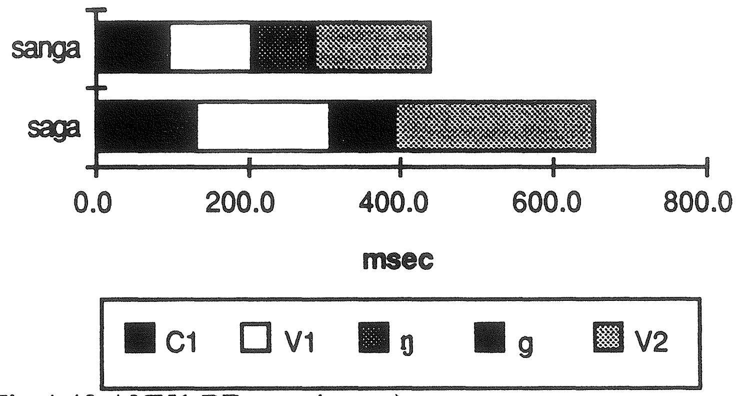


Fig.7.15 ANU2 NP saga/sanga



(Fig.4.40 ANU1 BD saga/sanga)

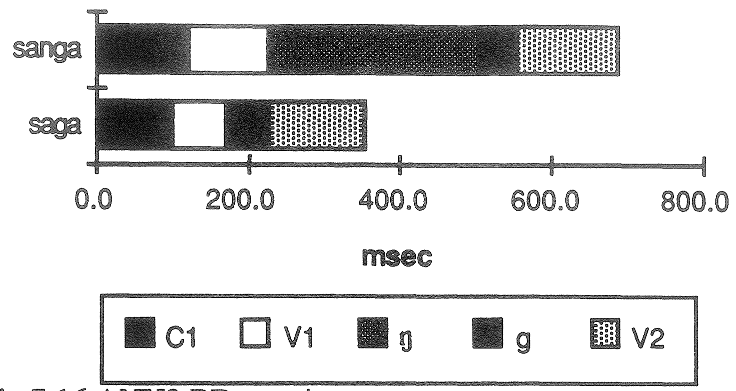


Fig.7.16 ANU2 BD saga/sanga

BD's pronunciation has changed from BD1 [saga/saŋa] to BD2 [saga/saŋga]. These examples also show that learners are able to distinguish the word duration without using different sounds.

7.4.2 Long/short contrast

Average: ANU2

The ANU2 results demonstrate that the degree of differentiation is greater than for ANU1 in which long/short contrast was often underdifferentiated.

The ratios for the KATE, RIKA and ISO series have changed from 69:100:63, 62:100:72 and 72:100:83 (ANU1) to 53:100:61, 54:100:59 and 60:100:72 (ANU2) respectively (see Table 7.4 below). Thus the degree of differentiation is greater in all examples of ANU2 and this can be regarded as an improvement. In comparison with NS (42:100:39, 39:100:37 and 48:100:49 respectively), however, the long/short contrast for the obstruent C2 is still underdifferentiated.

It was found in ANU1 that the learners tended to differentiate the duration of word-final V2 more clearly than that of intervocalic C2. This tendency is even more exaggerated in ANU2. As a result, the long/short contrast for V2 in ANU2 has become more distinctive than in NS (i.e. overexaggeration).

The average ratio of differentiation of a stop C2 for ANU2 is 1.8:1, which is between the results of Han's (1992) American advanced-level subjects (2:1) and those of ANU1 (1.5:1).

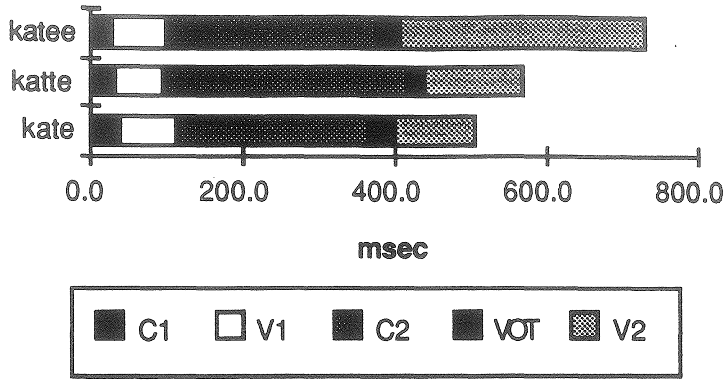
	C2	V2
1 kate	53	45
2 katte	100	58
3 katee	61	100
4 rika	54	41
5 rikka	100	41
6 rikaa	59	100
7 iso	60	39
8 isso	100	40
9 isoo	72	100
10 koma	39	
11 komma	100	
12 hone	37	
13 honne	100	
14 saga	55	
15 sanga	100	
16 saŋa	52	
17 saŋŋa	100	

Table 7.4 Long/short contrast (ratio)

Individuals: ANU2

Individual subjects' ratios for long/short contrasts are presented in Appendix 7.8.

In Chapter 4, it was found that the main causes of underdifferentiation were shorter duration of geminate C2, and longer duration of single C2. In ANU2, the former is no longer a problem; the latter also showed some improvement in the following examples (the auditory impression of these examples has also improved):



(Fig.4.31 ANU1 LD kate/katte/katee)

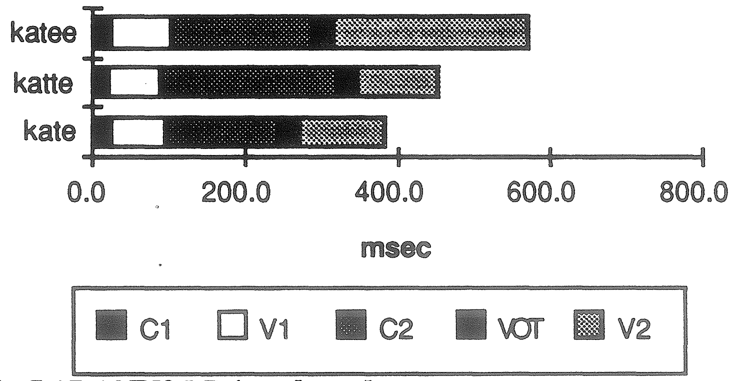
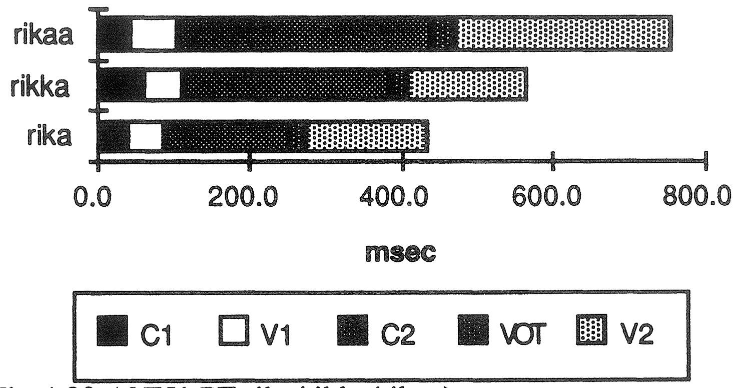


Fig.7.17 ANU2 LD kate/katte/katee



(Fig.4.32 ANU1 PF rika/rikka/rikaa)

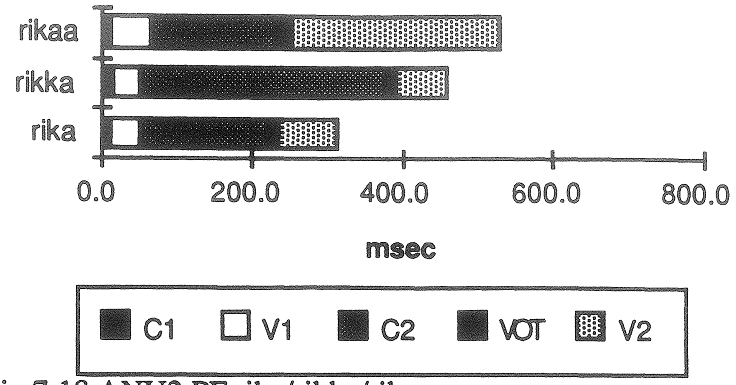
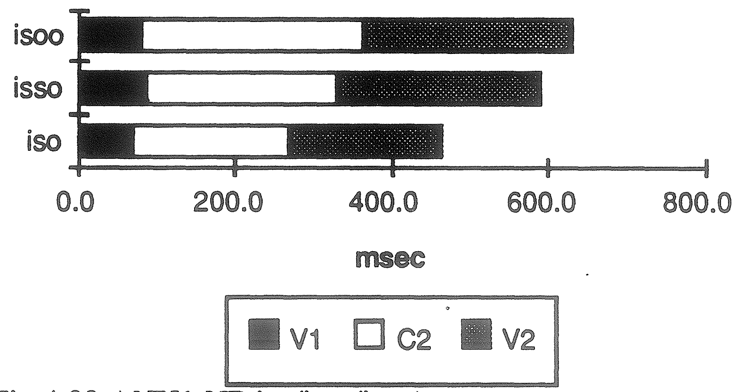


Fig.7.18 ANU2 PF rika/rikka/rikaa



(Fig.4.33 ANU1 NP iso/isso/isoo)

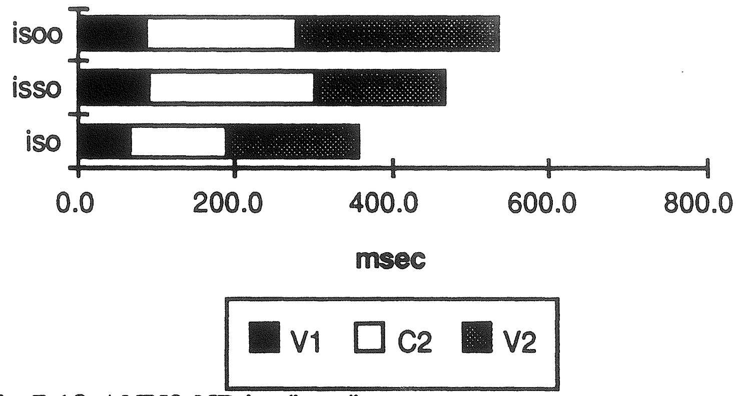


Fig.7.19 ANU2 NP iso/isso/isoo

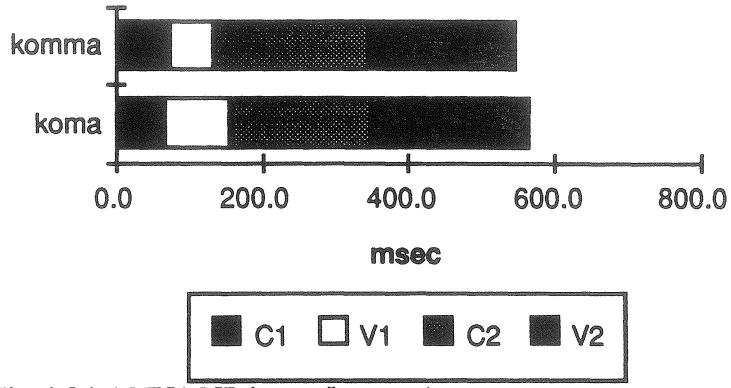
It is important to note that the problem associated with longer duration of single C2 appeared to be more difficult to overcome than that with shorter duration of geminate C2. The problem is still evident in LD2 (RIKA and ISO series) and LB2 (RIKA series) after the lesson. Therefore it can be said that greater difficulty is involved in improving the timing control of the single C2 in CVCV/CVCVV sequences than that of the geminate C2 in CVCCV sequences.

It is not appropriate, however, to attribute this result simply to the difficulty associated with the speech production of the single C2. It seems more likely that the mental representations of the sequences are related to the learners' previous language experience (possibly L1), and that their speech production is also influenced by the way they categorize the given sequence. A supporting piece of evidence for this claim is that the learners showed greater threshold values in their perceptual targets than NS (Chapter 3). Some single C2 pronounced by learners have greater values than the corresponding values of NS. The results obtained in the present study may indicate that once these perceptual categories are established they are difficult to modify.

On the other hand, a CVCCV sequence which is new to the learners may be produced by overexaggerating the duration of geminate C2. This strategy is effective in terms of making a clear durational contrast (although the learners' output is still not the same as the NS norm). The problem involved in the timing organization of CVCV is that because their L1 has a superficially equivalent CVCV sequence (although it is acoustically different from L2), it may be categorized as a part of their L1. Consequently, the timing organization of CVCV sequence may be more difficult to acquire than that of CVCCV.

A new tendency which appeared after the lesson was overexaggeration of the duration of geminate C2. For example, the differentiation of the long/short contrast in BD2 (25/100/26) in the KATE series and KL2 (24/100/30) in the RIKA series is more distinctive than in NS (42/100/39 in the KATE series and 39/100/37 in the RIKA series). These ratios show an improvement over those obtained in ANU1 (BD1 109/100/35 and KL1 66/100/65 in the corresponding examples), and learners' pronunciation in ANU2 gave a better auditory impression to an NS.

As mentioned earlier, ANU1 learners often had problems with the production of long/short contrasts in words with a nasal C2, but most of the problem areas found in the earlier data showed improvement as illustrated below:



(Fig.4.34 ANU1 NP koma/komma)

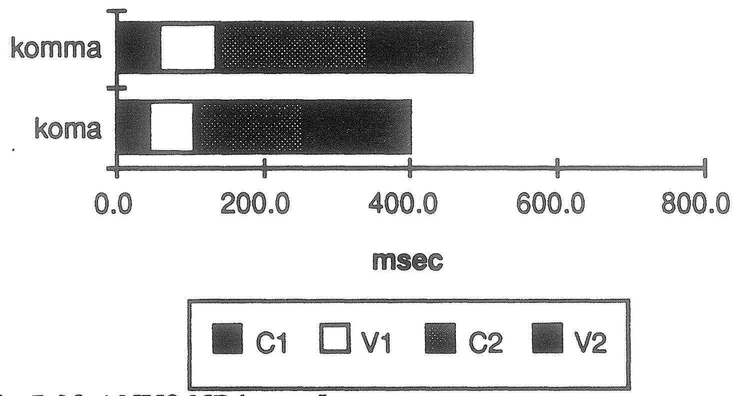
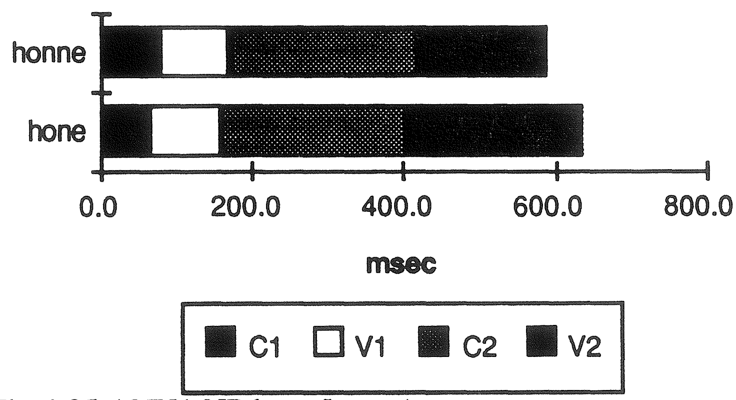


Fig.7.20 ANU2 NP koma/komma



(Fig.4.35 ANU1 NP hone/honne)

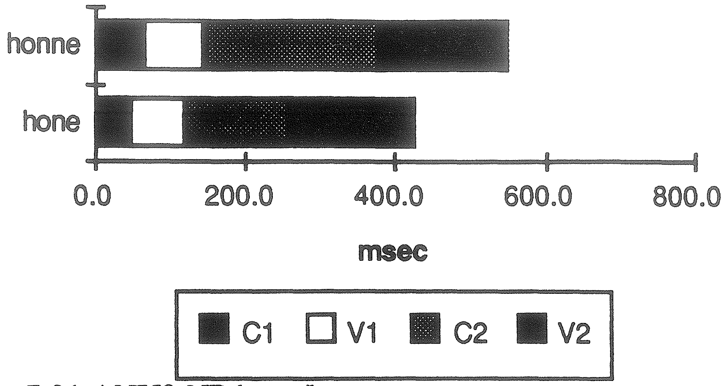
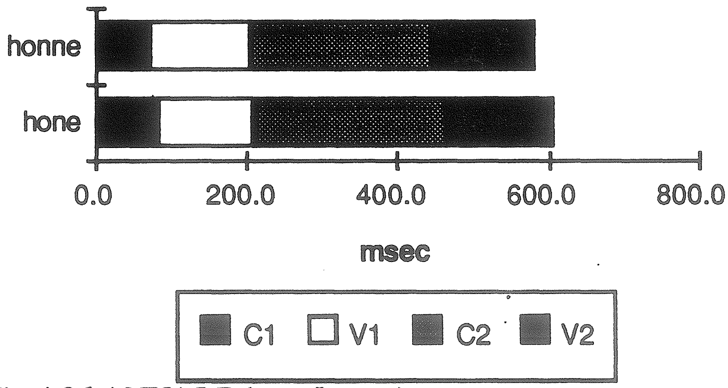


Fig.7.21 ANU2 NP hone/honne



(Fig.4.36 ANU1 LB hone/honne)

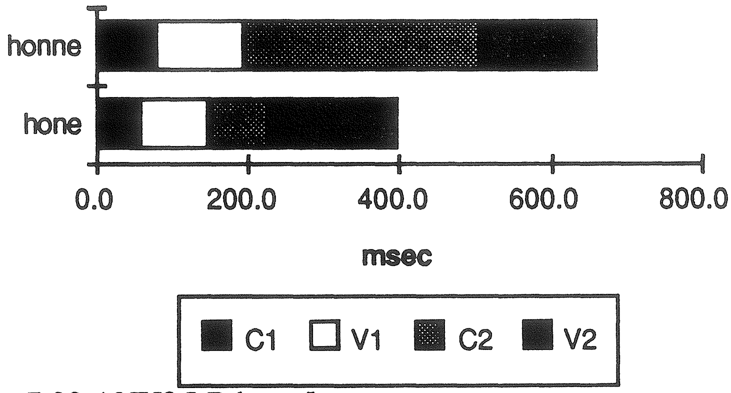
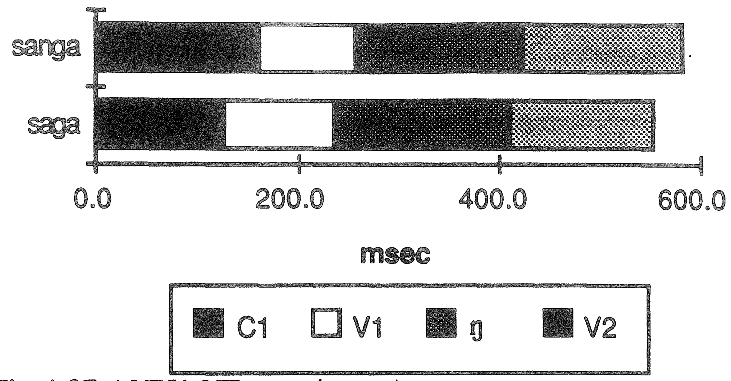


Fig.7.22 ANU2 LB hone/honne



(Fig.4.37 ANU1 NB saga/sanga)

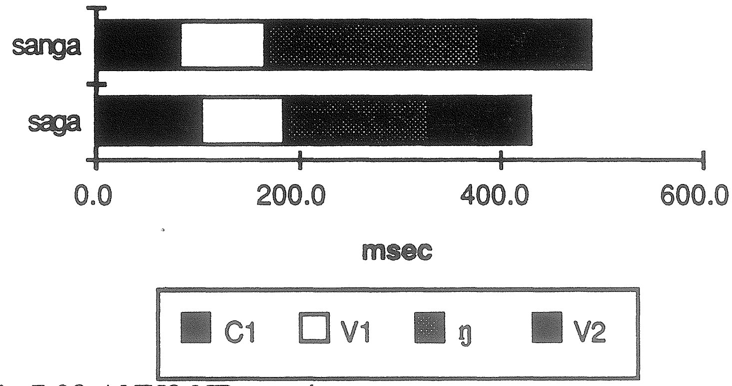


Fig.7.23 ANU2 NB saga/sanga

Vowels

Average: ANU2

The average ratios of V2 were generally more clearly differentiated in ANU2 (45:58:100, 41:41:100 and 39:40:100) than those of C2 in ANU1 (48:58:100, 50:53:100 and 49:55:100) (KATE, RIKA and ISO series respectively).

As a result, the rate of differentiation in the V2 ratios of the RIKA and ISO series has become greater than for NS (47:44:100, 49:46:100 and 44:42:100).

Individuals: ANU2

Common types of underdifferentiation in ANU1 were:

1. The duration of long V2 short. This tendency was often found in words with an HL pitch pattern, namely, the RIKA series. For example, the durational contrast in the V2 was underdifferentiated for this reason in LB1 (92:96:100) and NP1 (91:91:100). The ratios of these speakers changed to LB2 (37:36:100) and NP2 (63:70:100) after the lesson, thus showing greater differentiation.

2. The duration of short V2 was too long. This tendency was common in examples with an LH pitch pattern, i.e. in the KATE and ISO series. For example, the durational ratios of V2 were (73:64:100) for NB1 and (73:96:100) for NP1 in the ISO series. These ratios changed to (38:52:100) in NB2 and (66:64:100) in NP2, and the degree of differentiation thus became greater after the lesson.

In the ANU2 results, overexaggeration of the durational contrast was also observed in V2. For example, PF2 (29:25:100) and LD2 (20:26:100) in the RIKA series, and KL2 (27:25:100) in the ISO series demonstrate greater differentiation than NS in the RIKA series (49:46:100) and in the ISO series (44:42:100) respectively.

7.4.3 C2/V1 ratios

In ANU2, there is a tendency for the C2/V1 ratios in words with a single C2 to become smaller and for those with a geminate C2 to become greater in comparison with those obtained in ANU1. For example, the average C2/V1 ratios for the RIKA series were 2.46:3.06:2.20 in the ANU1 data and 1.94:3.82:1.86 in ANU2 (see Table 7.5 below). Consequently the differences in the ratios between single and geminate C2 are greater. The overall ratios, however, are still greater than the corresponding NS data (1.31:2.82:1.13). This means that the duration of the C2 is still longer in relation to that of the V1 in learners' speech than in NS speech.

In ANU1, it was found that the duration of the short V2 was too long in relation to that of the V1 (especially in words with an LH pitch pattern). For example, the average V2/V1 ratios for the ISO series were 2.16:1.95:3.94 in the ANU1 data and

1.86:1.48:4.37 in ANU2. Therefore it can be said that after the lesson, the learners' V2/V1 ratios became more clearly differentiated in a similar way to those of NS (1.72:1.33:3.75).

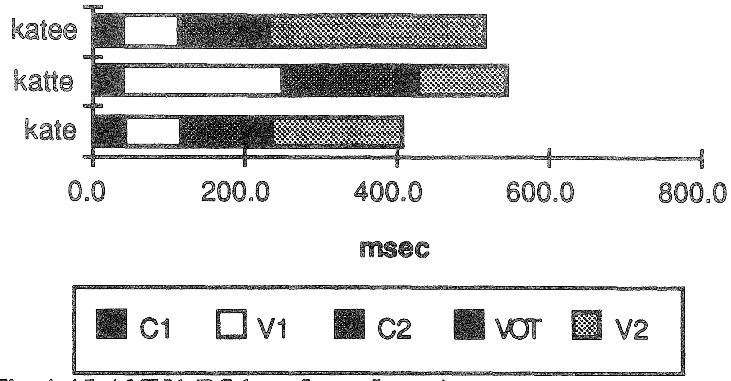
Average: ANU2

	C2/V1	V2/V1
1 kate	1.56	1.90
2 katte	2.97	1.73
3 katee	1.91	4.49
4 rika	1.94	1.89
5 rikka	3.82	2.00
6 rikaa	1.86	3.99
7 iso	1.94	1.86
8 isso	2.51	1.48
9 isoo	2.13	4.37
10 koma	1.26	1.41
11 komma	1.96	1.03
12 hone	1.32	1.74
13 honne	2.76	1.57
14 saga	1.27	1.36
15 sanga	1.86	1.15
16 saŋa	0.94	1.33
17 saŋŋa	1.62	1.19

Table 7.5 C2/V1 and V2/V1 ratios

Individuals: ANU2

Appendix 7.9 demonstrates individual ratios of C2/V1. The following figures illustrate individual improvements.



(Fig.4.45 ANU1 BS kate/katte/katee)

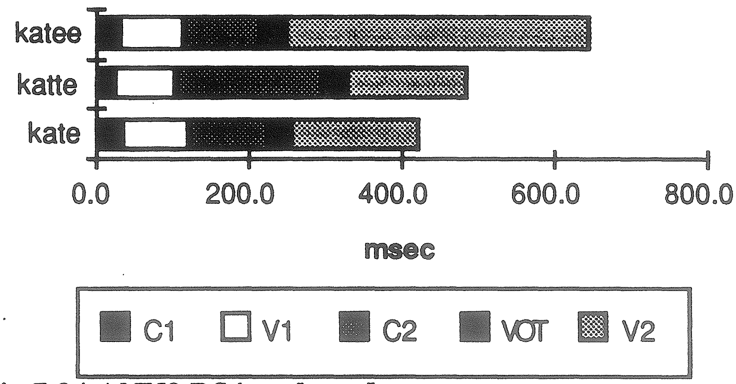
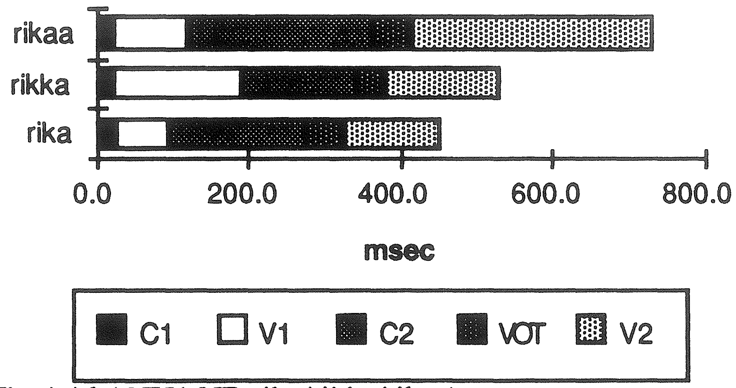


Fig.7.24 ANU2 BS kate/katte/katee



(Fig.4.46 ANU1 NB rika/rikka/rikaa)

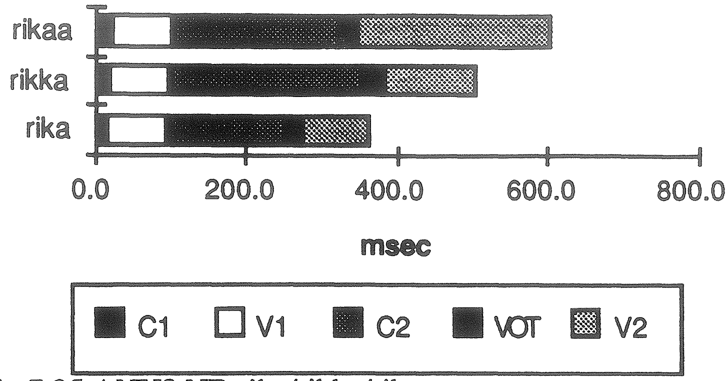


Fig.7.25 ANU2 NB rika/rikka/rikaa

(Fig.4.47 ANU1 PF iso/isso/isoo)

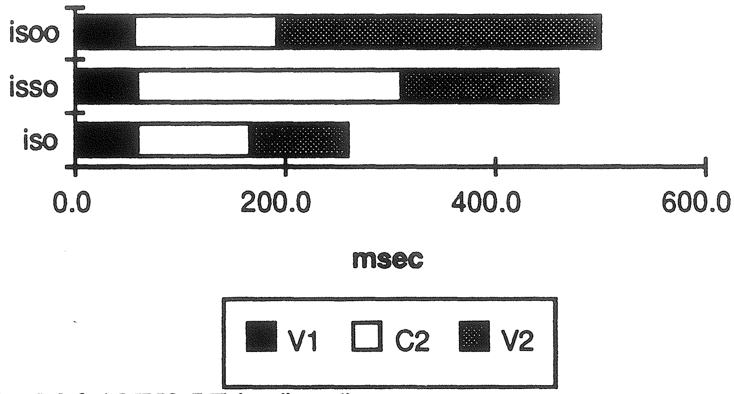
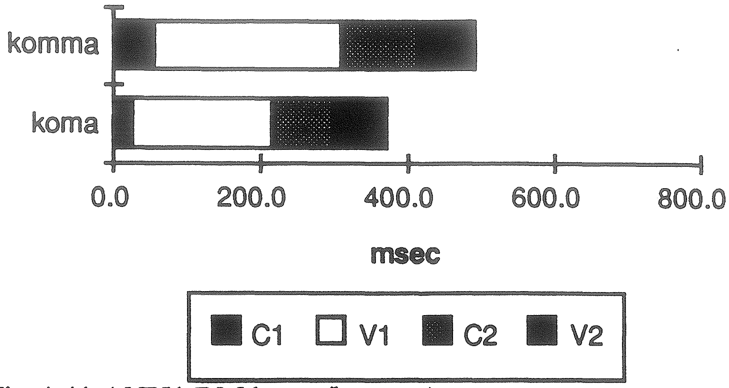


Fig.7.26 ANU2 PF iso/isso/isoo



(Fig.4.41 ANU1 BM koma/komma)

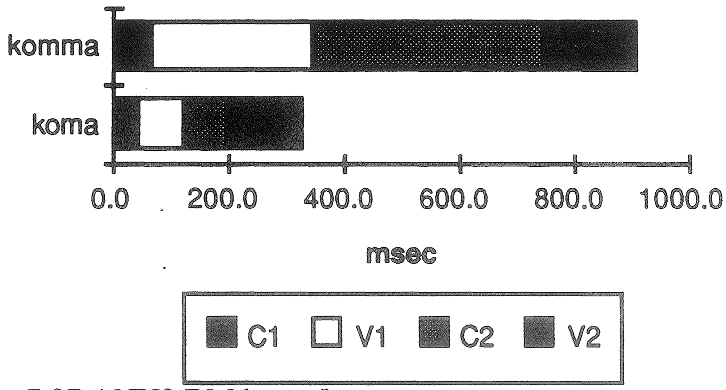


Fig.7.27 ANU2 BM koma/komma

The problem caused by the longer duration of V1 before a geminate C2 is eased after the pronunciation exercises as shown in Fig.7.24-27: BS1 pronounced a long V1 in [katte], NB1 in [rikka], PF1 in [isso] and BM1 in [komma] before the lesson, but all showed some improvement after the lesson. We can thus say that longer duration of V1 before a geminate C2 is a common problem for beginning-level learners, which can be improved by introducing some pronunciation lessons which focus on timing control.

Subjects who did not clearly differentiate the C2/V1 ratios of RIKA series in ANU1 such as KL1 (2.90:3.45:2.26) and PF1 (3.13:5.29:5.21) showed some improvement in the corresponding ratios in ANU2: KL2 (1.65:6.86:2.11) and PF2 (3.81:8.05:2.89). Clear differentiation can now be observed in the ratios after the lesson for those speakers. On the other hand, LD2 still has some problems in C2/V1 ratios due to greater values for the single C2.

7.4.4 Discussion

After the lesson, the problem of underdifferentiation of durational contrasts (both word duration and long/short contrast in C2 and V2) diminished. Instead, many learners applied a strategy of overexaggeration in order to make the durational contrast. As a result, the durational ratios of ANU2 were greater than those of NS in some examples.

The other interesting finding was the learners' interpretation of 'shorter' duration. They were aware of the concept of durational contrast, and tried to achieve the distinction by changing the speech tempo; some learners produced two-mora expressions 'faster' than the corresponding three-mora expressions in order to achieve the durational contrast.

Additionally, the degree of differentiation depended on the segmental structure of the utterances. It was greater in CVCVV than CVCCV. This implies that the acquisition of the durational contrast in C2 involves greater difficulty than that of V2.

7.5 Identification tests

Identification tests refer to the tests in which learners are asked to identify which words they had attempted to pronounce while listening to the recording of their speech production test (Chapter 4).

7.5.1 Aims

In the previous section, the acoustic-phonetic details of learners' speech production were compared before and after the lesson, and it was claimed that various aspects of their speech production more closely approximated the NS norm.

The purpose of this chapter is to examine learners' perception of their own speech production. whether learners could identify which words they had attempted to pronounce for the experiment in Chapter 4. It was found that various learner strategies were operative for making durational contrasts, and that the phonetic realization was not necessarily the same as the NS norm. If learners are able to discriminate the acoustic-phonetic characteristics of the phonetic realization for the durational contrast in the self-judgment of their own speech production, they should be able to identify which words they pronounced. In order to test this hypothesis, three learners who demonstrated certain characteristics in their speech production (see 7.5.2) were asked to listen to their recording and identify which words they were trying to produce about one month after the recording. The results are compared before and after the lesson.

The other purpose of this test is to investigate the extent to which the changes in the learners' speech production after the pronunciation lesson improved NS auditory judgment of learners' pronunciation. It is not appropriate to conclude that the learners' speech has improved after the lesson, unless we can demonstrate that the learners' speech is now better understood by NS. For this reason, six NS were asked to listen to recordings of the learners' pronunciation recorded before and after the lesson.

7.5.2 Subjects

Three informants (BM, BD and LD) were asked to listen to their own recordings and identify which words they had attempted to pronounce in the experiment conducted in Chapter 4. The main characteristics of their pronunciation were as follows: longer duration of V1 before a geminate C2 in the KATE, RIKA and KOMA series (BM); shorter duration of a geminate C2 in the ISO series (BD); and longer duration of a single C2 (LD).

Six NS subjects are from the same group of subjects who participated in the experiments discussed in Chapter 3 and 4.

7.5.3 Procedure

The words recorded in production test 1 (Appendix 4.2) were randomised and recorded onto a cassette tape. The learners were asked to listen to their own recording and circle one of the given words on the answer sheet, which they thought they had attempted to pronounce. 'Not sure' was always given at the end of the options as well as all the words in the given series (e.g. [kate/katte/katee/not sure] are given as options for [kate], and [koma/komma/not sure] for [komma]). There were fifteen words in the list ([kate/katte/katee], [rika/rikka/rikaa], [iso/isso/isoo], [koma/komma], [hone/honne] and [saga/sanga]) each of which was repeated three times in production test 1. Therefore forty-five utterances were randomised and tested in the experiment.

Similarly, the NS subjects were also asked to listen to learners' recording and circle one of the given words which they thought learners had pronounced.

The types of misjudgment by learners and NS are divided into the following five categories:

1. Short -> Long (C2)	e.g. [rika]	->	[rikka]
2. Long -> Short (C2)	e.g. [rikka]	->	[rika]
3. Short -> Long (V2)	e.g. [rika]	->	[rikaa]
4. Long -> Short (V2)	e.g. [rikaa]	->	[rika]
5. Not sure			

For example, when a learner circles [rikka] for the word which she had attempted to pronounce as [rika] in the recording session, this error is considered as misjudgment of short C2 as long C2 (category 1). When a NS listens to learners' recordings, he/she might mark [katee] for the word in which the speaker was trying to say [kate]. This case is considered as category 3, where the NS thought there was a long V2 instead of a short V2.

The errors related to the duration of consonants (1 and 2) were plotted on the plus side of the X axis, while those related to the duration of vowels (3 and 4) was plotted on the minus side of the X axis. Additionally, 'not sure' is plotted onto the plus side of the X axis. The errors were plotted separately when two types of errors occur in one example (e.g. two errors in [rikaa] -> [rikka] were plotted separately as category 1 ([k] -> [kk]) and category 3 ([a] -> [aa])).

7.5.4 Identification tests before the lesson

The following graphs (Fig. 7.28-30) show the results of the three learners' self-judgment and the NS-judgment concerning the pronunciation of the three learners recorded before the lesson.

Learners' self-judgment

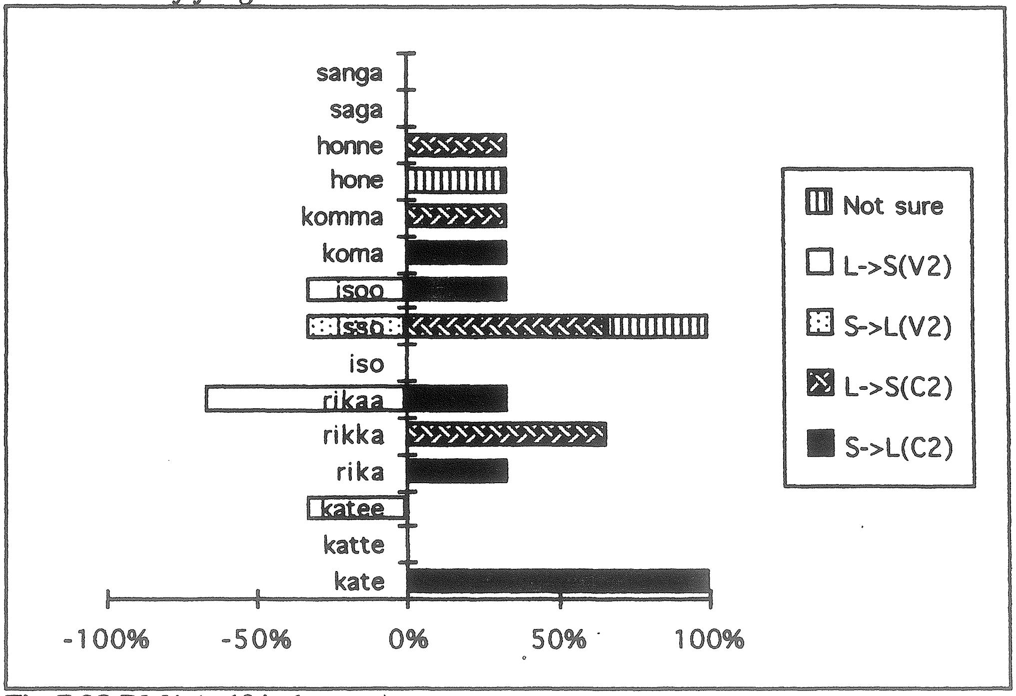


Fig.7.28 BM1 (self-judgment)

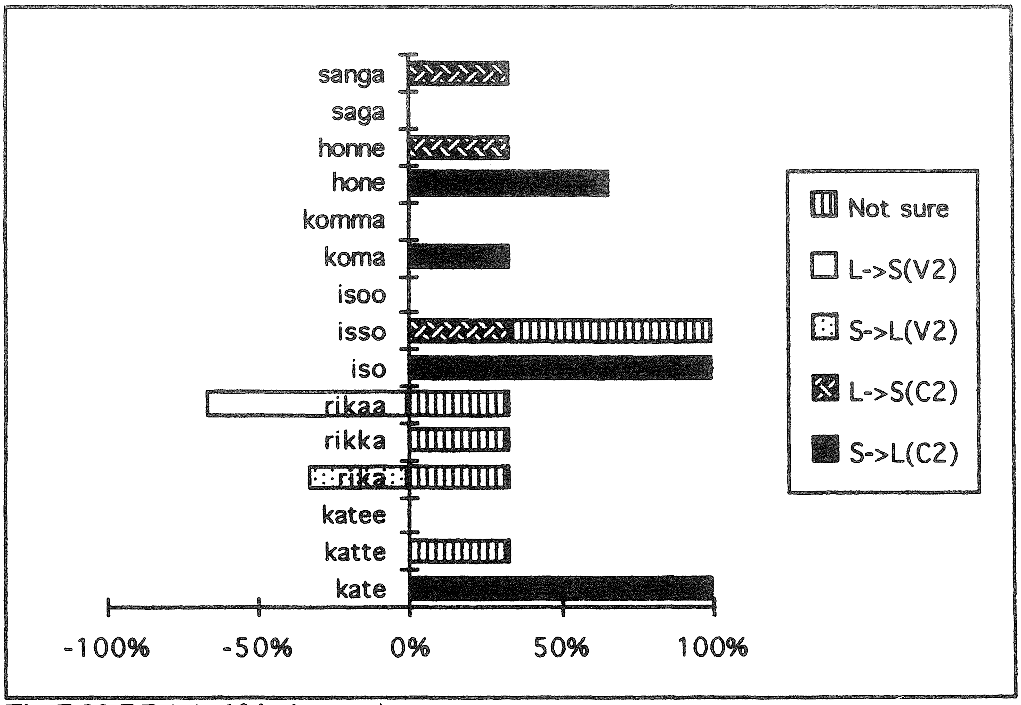


Fig.7.29 BD1 (self-judgment)

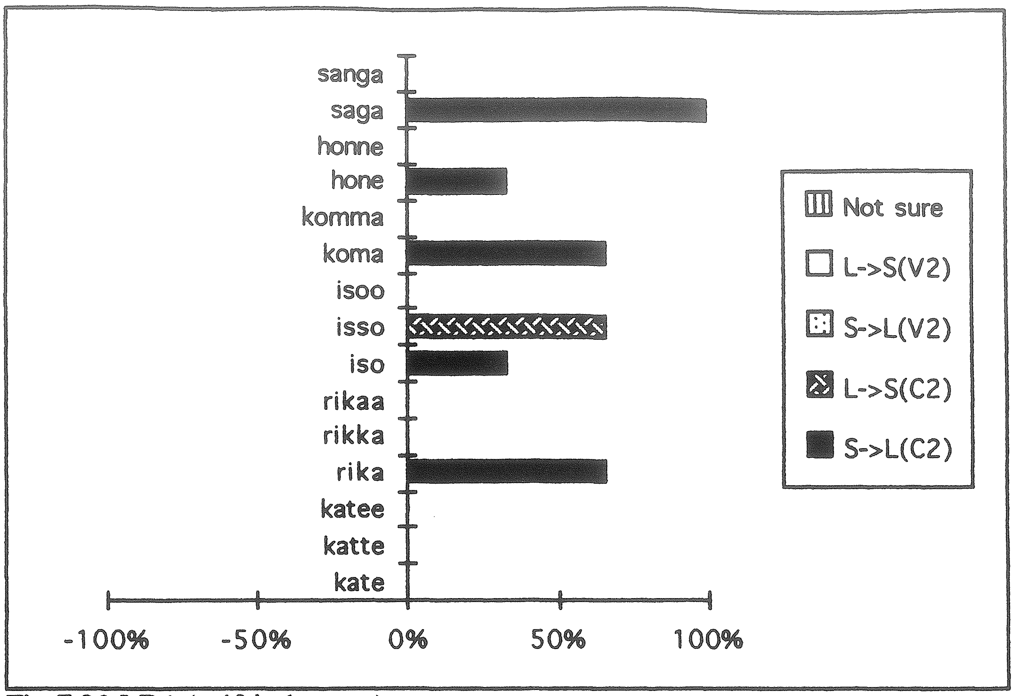


Fig.7.30 LD1 (self-judgment)

Firstly, the results of BM1 and BD1 are similar in that their errors are bidimensional: misjudgment occur in both directions of Long -> Short (in CVCCV) and Short -> Long (in CVCV) in these subjects' results. Also, the misjudgment seems to occur in both vowels and consonants. In both BM1 and BD1, however, the number of errors related to the duration of consonants (e.g. [kate] -> [katte], [isso] -> [iso]) is greater than those related to the duration of vowels (e.g. [rikaa] -> [rika]).

On the other hand, LD's results are rather unidimensional: All her errors are related to Short -> Long in C2 except for [isso]. There is no error related to the duration of V2. For these reasons, it appears that LD1's results are more consistent, and she was able to identify her own pronunciation more accurately than BM1 and BD1 could identify theirs.

In Chapter 4, BM1's speech was described as having a long V1 before the geminate stop C2 in [katte] and [rikka]. Also, it was observed that in [honne], the duration of the V2 was longer. It was claimed that realization of the durational contrast of the nasal C2 tended to be conditioned by pitch patterns: when the pitch pattern was

HL (e.g. KOMA series), it was realized in terms of the duration of the V1, and when the pitch pattern was LH (e.g. HONE series), in terms of the duration of the V2.

From the results obtained in the BM's identification test, however, it is impossible to say that BM1 always used these features as phonetic cues to identify which words she had attempted to pronounce. With regard to words with an obstruent C2, she could not correctly identify examples of CVCCV (e.g. [rikka] -> [rika], [isso] -> [iso]), which means that she did not interpret a longer duration of V1 as a cue for a geminate C2 (see Fig.7.28).

On the other hand, the correct rate of BM1's identification of words with a nasal C2 is 66%. This is comparatively high and is a much better result than NS-judgment of BM1 (see below for NS judgment). With regard to the nasal C2, there is only a two-way contrast (e.g. koma/komma), while there is a three-way contrast in words with an obstruent C2 (e.g. kate/katte/katee). Therefore, the self-identification of words with a nasal C2 may have been easier than that of words with an obstruent C2. For example, BM's pronunciation of the HONE series was [hone/honee] (instead of [hone/honne]), and yet was identified correctly by this speaker. Another reason for the high identification rate of the SAGA series was because she distinguished these words using [g] and [ŋ].

In Chapter 4, BD1 produced a shorter duration of the geminate C2 in the ISO series. BD1's self-judgment of the ISO series in this experiment was not accurate: none of the examples of [iso] and [isso] was correctly identified. Also, a number of errors are related to Short -> Long C2 in CVCV sequences (e.g. [kate], [iso] and [hone]) in BD1's results. Additionally, BD1 had several 'not sure' responses in her answers, and so it is not possible to say that her self-identification was successful.

LD1 was said to have had a longer duration of the single C2 in Chapter 4. In LD1's self-judgment, Short -> Long in the C2 was the most common misjudgment, especially for words with a nasal C2; the same problem can be observed in the words with an obstruent C2 (e.g. [rika]), but to a lesser extent. Two-thirds of the examples with an obstruent C2 were correctly identified. It is interesting that LD1 was able to recognize the longer duration of a single obstruent C2 as a single C2 with an accuracy of 70% although she could not distinguish the durational contrast in words with a nasal C2.

NS-judgment

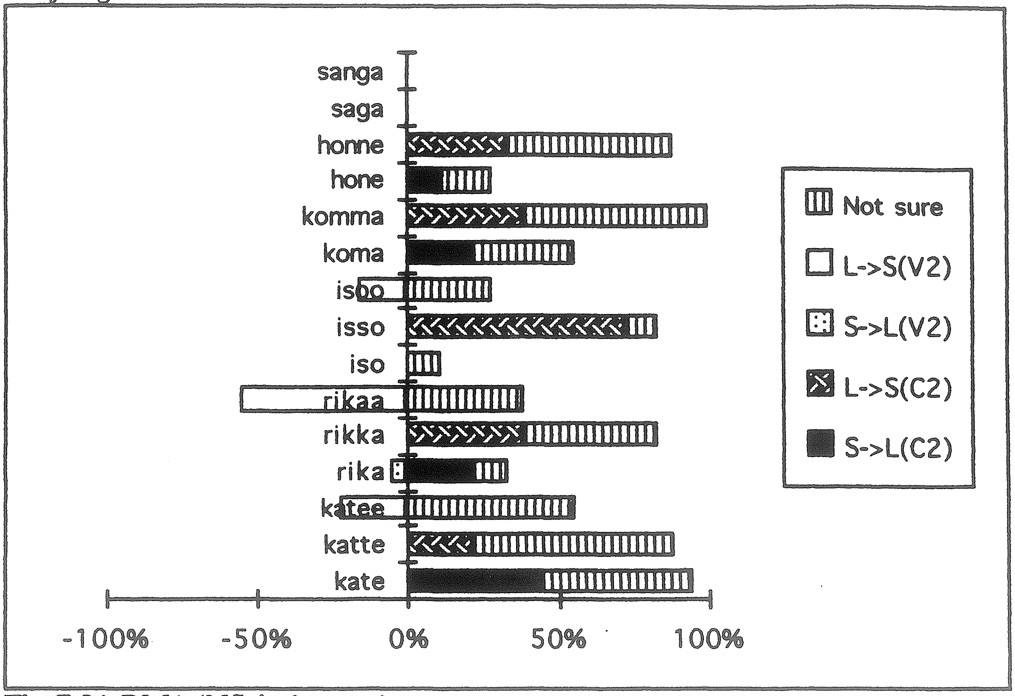


Fig.7.31 BM1 (NS-judgment)

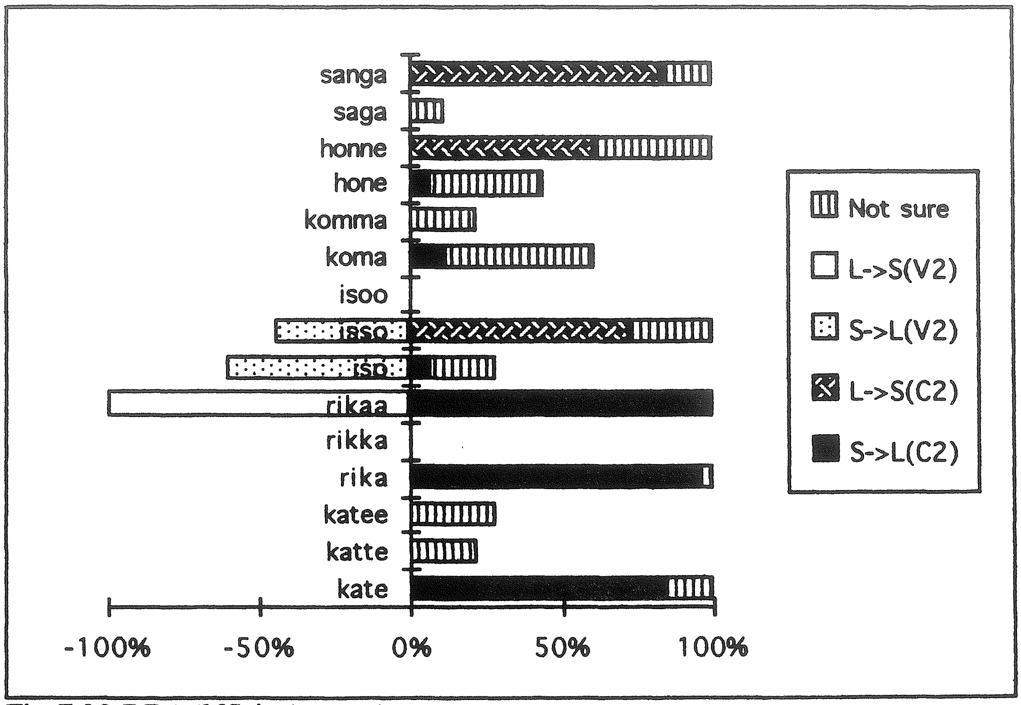


Fig.7.32 BD1 (NS-judgment)

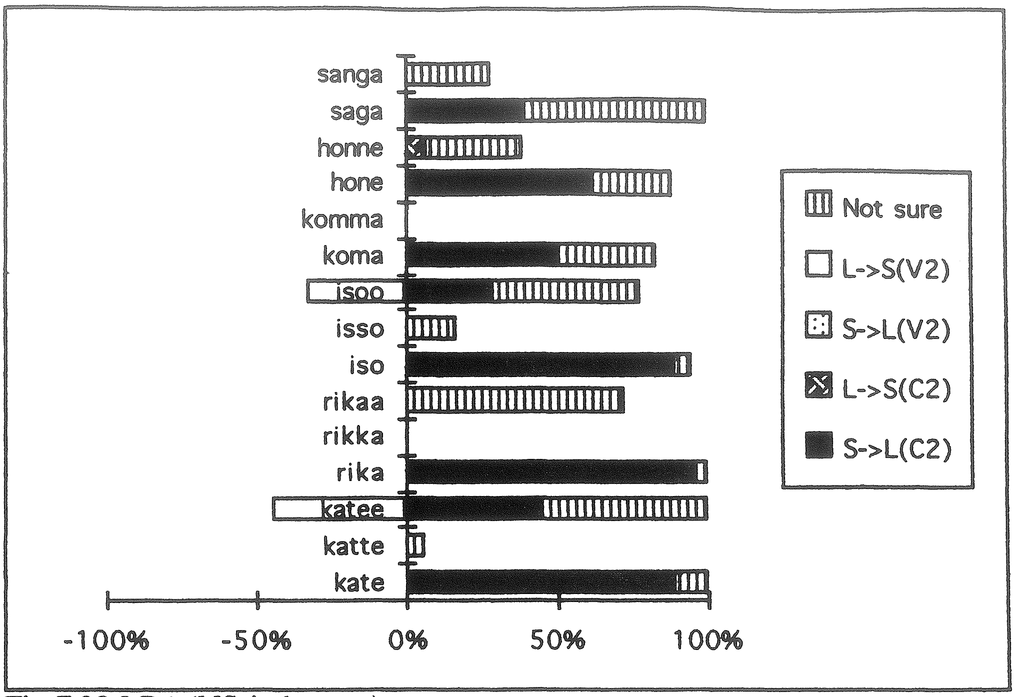


Fig.7.33 LD1 (NS-judgment)

The number of misjudgments concerning learners' pronunciation by NS was much greater than that by the learners. The results of NS-judgment also contained a number of 'not sure' responses, probably due to the fact that the learners' pronunciation frequently did not fit in any of CVCV, CVCVV and CVCCV (e.g. [rikaa] pronounced as [rikkaa] etc).

It is interesting to note that the number of NS misjudgments concerning LD1 was much greater than LD1's own self-judgment. LD1's misjudgment was quite consistent in that all errors were related to consonant duration, amongst which Short -> Long in C2 was the most common. In NS judgments, there were a number of 'not sure' responses. Also, some misjudgments were related to the duration of V2 in LD1. These results show that NS listeners did not identify LD1's pronunciation as well as LD1 identified it herself.

Additionally, some discrepancies between the judgments of NS and learners were observed. For example, all examples in the KATE series were identified correctly by LD1, but NS treated a single C2 as long (e.g. [kate] -> [katte]). The average long/short ratios were 42:100:39 for NS while they were 81:100:87 for LD1. Thus all

C2 in LD's KATE series were too long to be regarded as a single C2 by NS since they were also beyond NS perceptual boundary for a single C2 and therefore were identified as a geminate C2. Interestingly, however, LD1 correctly identified single and geminate C2 using her own perceptual boundary. There may also be some other linguistic cues besides the duration of C2 which played a role in the discrimination of long/short contrasts by this speaker.

Similarly, in NS judgments of BD1, both [rika] and [rikaa] were perceived as having a long C2. Interestingly, BD1 did not consider them to be long in her judgment. BD1's ratios for the long/short contrast of C2 are 81:100:72, while they are 39:100:37 for NS. This provides another example of how differences in the perceptual boundary affect speech production of the learners as well as the perceptual judgment of both NS and learners in the present study.

7.5.5 Identification tests after the lesson

The following figures show the results of the identification tests after the lesson for learners and NS.

Learner's self-judgment

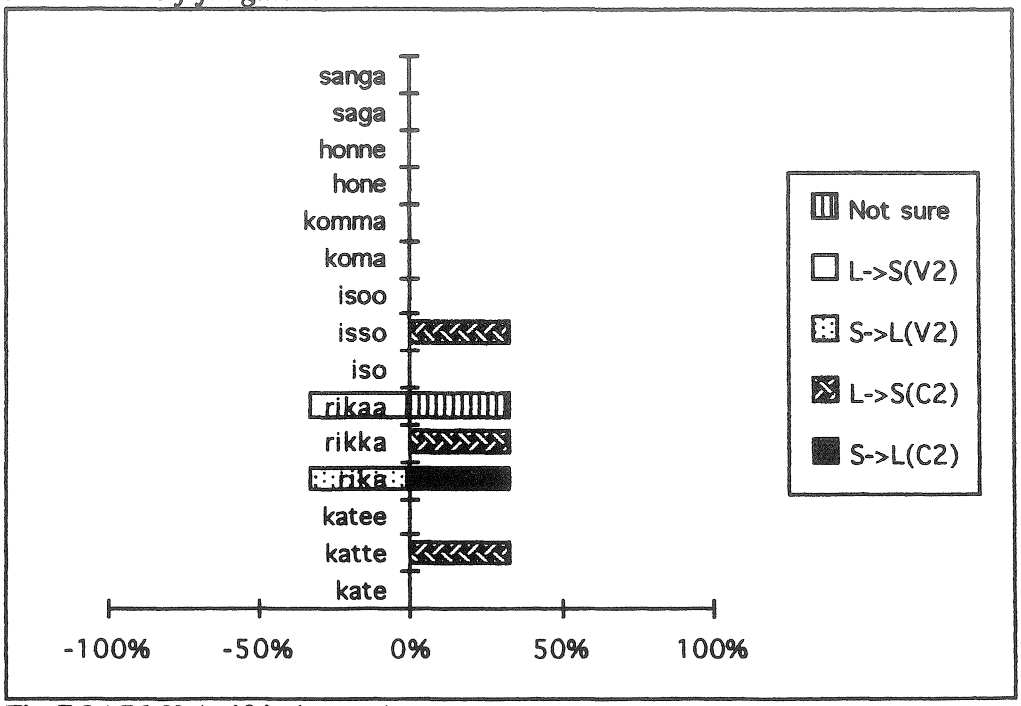


Fig.7.34 BM2 (self-judgment)

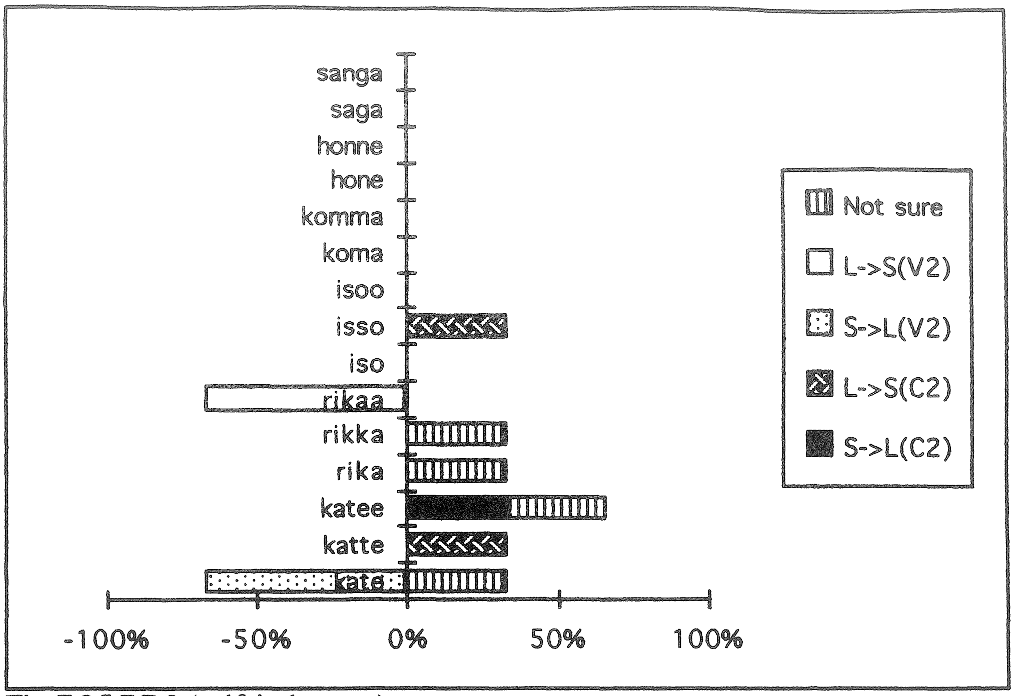


Fig.7.35 BD2 (self-judgment)

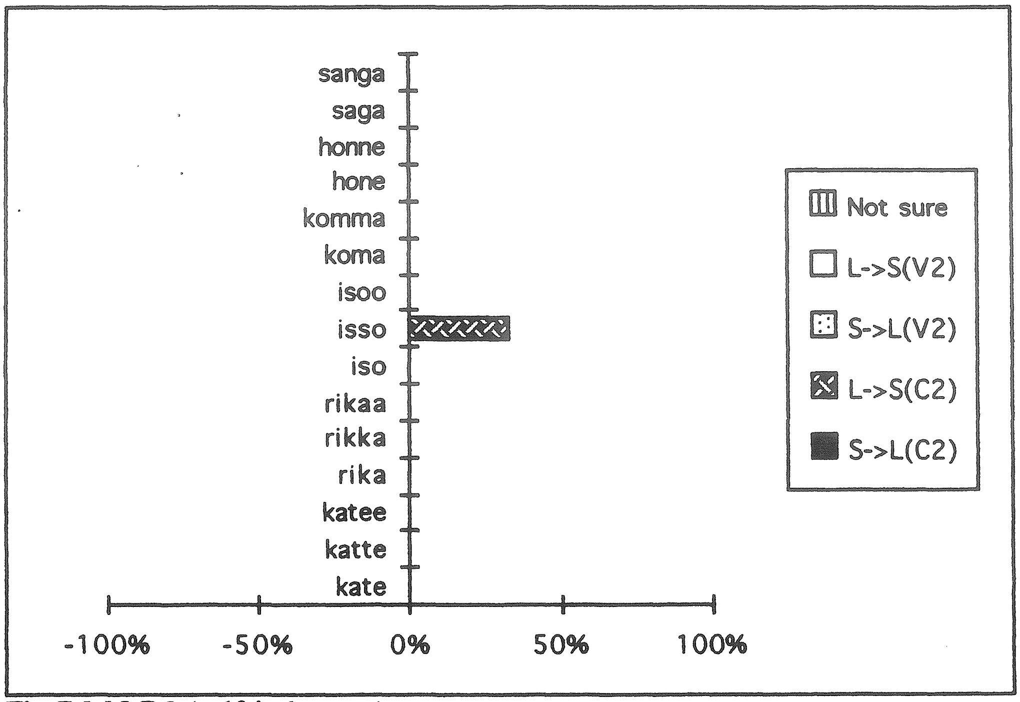


Fig.7.36 LD2 (self-judgment)

The total degree of misjudgment decreased compared to the results obtained before the lesson, especially the detection rate of words with a nasal C2, which reached 100% for all learners. For LD2, the only error was [isso] -> [iso] and the rest of the examples were correctly identified. Therefore, it can be said that the learners' self-judgment was improved after the pronunciation lesson.

NS-judgment

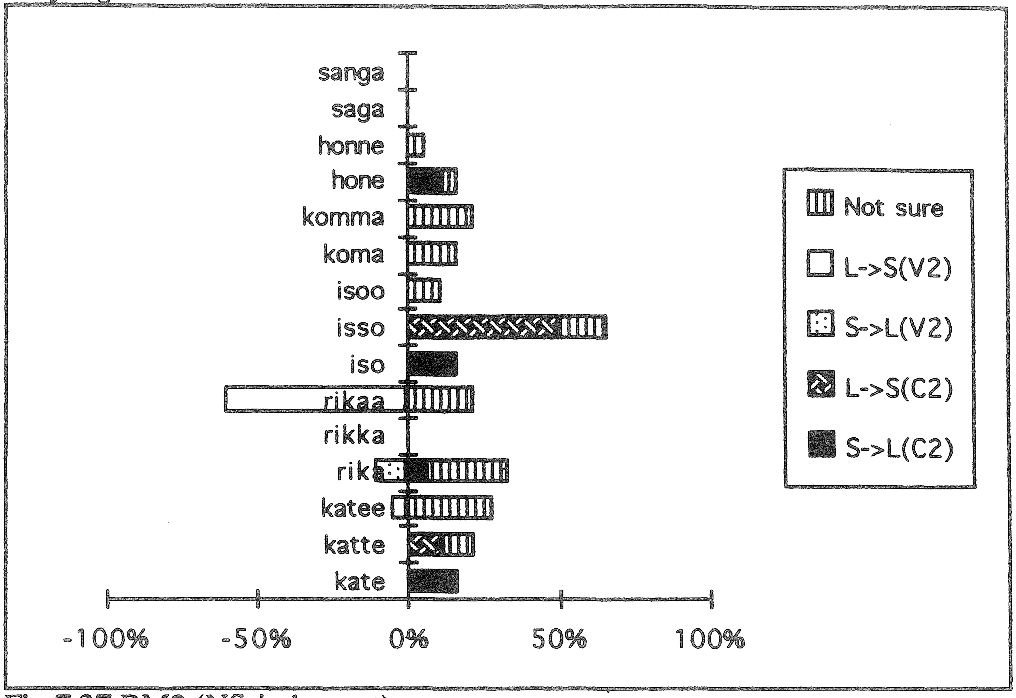


Fig.7.37 BM2 (NS-judgment)

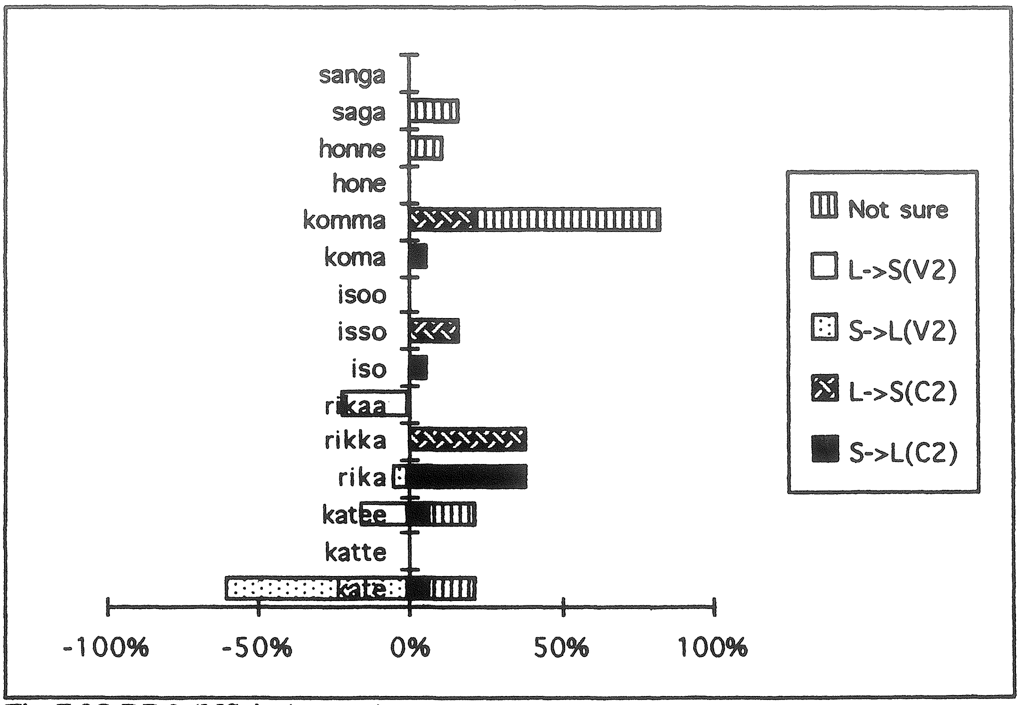


Fig.7.38 BD2 (NS-judgment)

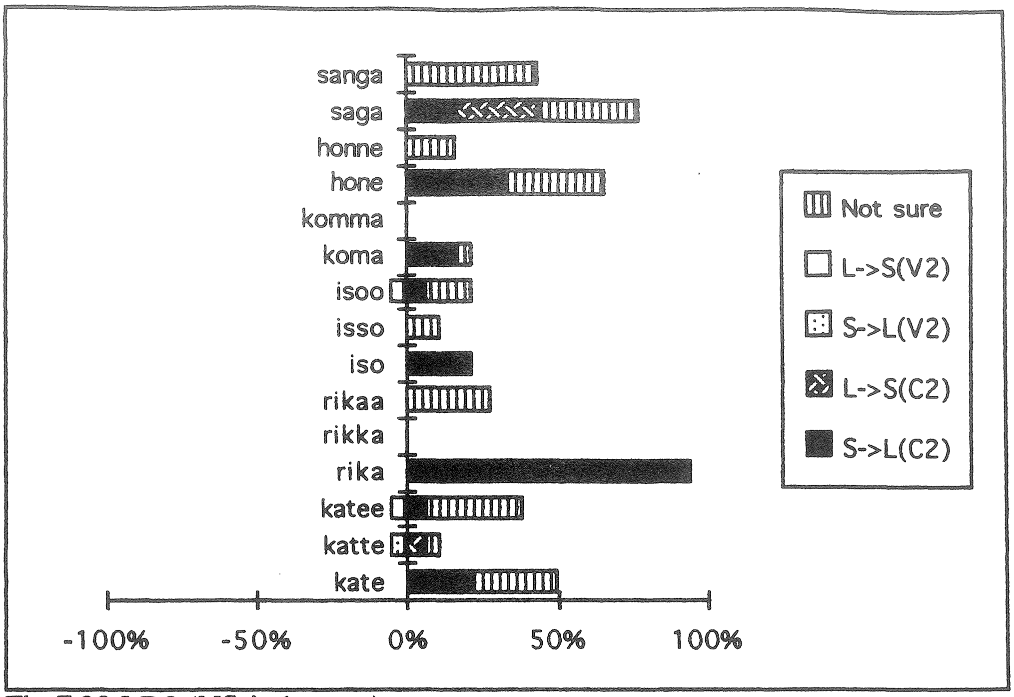


Fig.7.39 LD2 (NS-judgment)

A decrease in the total degree of misjudgment of learners' speech production was also observed in NS results, and from this point of view, it is possible to say that the learners' speech production was better understood by NS after the lesson. The degree of misjudgment on the part of NS, however, is still greater than that of the learners' themselves. For example, although LD2 could identify all examples except [isso] herself, NS made a number of misjudgments related to Short -> Long (C2). Also, there were a number of 'not sure' responses in the NS-judgment. Although learners could identify which words they had attempted to pronounce, NS could not identify them in the same way as the learners could. Thus some discrepancies were still found between NS and the learners' judgment.

As discussed above, one of the major causes of these discrepancies may be due to differences in the perceptual boundaries between NS and learners, since learners' perceptual boundaries have not been modified after the lesson (see 7.3).

7.5.6 Discussion

In the previous section, the results of judgment tests for both learners and NS were presented and the discussion followed on the results both before and after the pronunciation lesson.

Firstly, it was found that the learners were not able to fully identify which words they had attempted to pronounce before the lesson. This result provided little evidence for the assumption that the learners might be able to use characteristic features in their pronunciation (which were described in Chapter 4) as phonetic cues for the identification tests described in this chapter. In comparison with the NS-judgment results, however, the learners' self-judgment showed a much higher rate of correct identification. This was because learners could identify their own pronunciation as long as the examples were clearly distinguished (e.g. [hone] vs [honee]) while NS did not recognize an utterance such as [honee] as a word and therefore expressed uncertainty by circling 'not sure'.

Secondly, it was found that, after the lesson, both learners' and NS-judgment showed improvement. It was pointed out, however, that there were clear differences between learners' and NS judgments and that the rate of correct identification by NS was still less than that of the learners.

Finally, and most significantly, a clear correlation between differences in the perceptual boundaries of NS and learners (cf. Chapter 3) and differences in their judgment was observed. The greater values of the learners' perceptual boundary allowed them to correctly identify their single C2 (even with a longer duration), while NS did not successfully identify them as single C2 since the values fell into the perceptual category of geminate C2 for NS.

In summary, the teaching strategy applied in the lesson can be considered as effective because both NS and learners' judgment showed improvement after the lesson. The results obtained in this section, however, demonstrate that NS could not identify the learners' pronunciation as accurately as learners themselves after the lesson. This aspect of the results is not satisfactory because the lesson can only be considered as fully effective when the learners' pronunciation is understood without any difficulty by NS. Perhaps the time spent for this lesson was not sufficient and more exercises need to be conducted over a longer period of time to achieve better results.

As mentioned in 7.3.4, the optional number of learner contact hours to develop pronunciation which NS can understand without difficulty is a matter for further investigation.

7.6 Conclusion

The pronunciation lesson was effective enough to increase the degree of categorization of this group of beginners to a similar level as the learners who were about to complete the one-year intensive language course (see Chapter 5).

The perceptual boundary, on the other hand, was still different from that of NS after the lesson, and this implies that the mechanisms involved in learners' perceptual judgment were not similar to those of NS. In fact, learners' perceptual boundaries did not approximate those of advanced learners.

As far as the results obtained in this study are concerned, it can be said that the acquisition of a greater degree of categorization precedes the acquisition of NS-like perceptual boundaries. This agrees with the findings obtained in the longitudinal study (Chapter 5). Further study is needed to find the stage at which the latter aspect of learners' perception is improved, and a teaching method which may expedite the learning.

In terms of speech production, some improvement was observed in terms of making clear durational contrasts. Some learners' strategies to achieve this goal were interesting: for example, some learners applied the strategy of speaking fast (i.e. changing the speech tempo) in two-mora expressions so that they could make a clear contrast in word duration between two- and three-mora expressions. Overexaggeration of the duration of long phonemes was another common strategy.

Finally, learners' improvement after the lesson can also be seen in the judgment test. The results were better than those obtained before the lesson, but NS judgments of learners' speech production were not as good as learners' self-judgment. The effectiveness of the lesson, therefore, was not satisfactory in this respect.

Chapter 8 Conclusion

This study has investigated various aspects of the acquisition of Japanese timing organization from the viewpoint of English-Japanese IL phonology.

In order to examine the processes which are involved in the formation of English-Japanese IL phonology, acoustically-based investigations of both perception and production were conducted with learners at various proficiency levels. The following processes were found to be operative:

1. Overexaggeration

One of the most interesting findings was that overexaggeration was a common strategy for both geminate stop C2 and long V2 in order to make the durational contrast, not only in the speech production of beginning learners, but also in that of advanced learners. It was found that the cause of underdifferentiation was often due to the longer duration of a single C2 rather than the shorter duration of a geminate C2, and as a result, the duration of the geminated C2 had to be overexaggerated in order to make the durational contrast (Chapter 3). Therefore it can be said that the common assumption of underdifferentiation due to shorter duration of geminate stops (Han 1992) does not give the full picture in respect of learners' problems. In one speaker's speech production in the longitudinal study (Chapter 5), it was found that the tendency for overexaggeration gradually become more pronounced during the year's observation.

The implication of this result for the teaching situation is that it is important for instructors to emphasise the appropriate timing control for learners rather than just focussing on making a clear differentiation of the phonological durational contrast. If instructors conduct pronunciation practice in which the focus is merely on making a clear differentiation, they may be replacing the problem of learners' overexaggeration of the durational contrast with another, that of underdifferentiation. Overexaggeration will also make learners' speech dissimilar to NS norms and therefore sound unnatural.

2. Syllable modification strategies

A unique syllable modification strategy was identified in both cross-sectional and longitudinal observations: it was found that the duration of V1 was prolonged only before a geminate C2. This finding is particularly interesting from the viewpoint of IL syllable structure; it seems that learners have L1 interference from the CVC-syllable structure of English, and that their implementation strategies for achieving the phonological contrasts are quite different from those of NS. This example demonstrated that learners were successful in making the distinction between single and geminate C2 in their own way, although their phonetic realization differed from the NS norm. Similarly, the interpretation of CVCCV sequences as CVC#CV was also found as a learner strategy, which supports the claim that unit of segmentation is largely determined by L1 (Mehler et al. 1981; Cutler et al. 1983, 1986; Otake et al. 1993). In this case, learners who applied this strategy successfully produced long duration for geminate C2 although they were using a segmentation unit which was different from that of NS.

3. L1 transfer

L1 transfer was operative in the learners' IL syllable structure and it was also found that a longer duration of VOT was always associated with the speech production of beginning learners. These results provide support for the weak version of CAH, which claims that learners' L1 influences the acquisition of the phonology of their L2. It was found that this problem was less prominent in advanced learners' speech production and thus supports the claim that L1 transfer is more common in beginning learners than in intermediate or advanced learners (Major 1987).

4. Phonetic approximation

Although a longer duration of VOT was associated with the speech production of beginning learners as a result of L1 interference, it was found that this situation was different in the VOT of advanced learners' speech production. In fact, their VOT was intermediate between English and Japanese, and this provided a good example of phonetic approximation. In both cross-sectional and longitudinal data, it was found that VOT values became smaller as the proficiency level of the learners increased, but not to the same extent as with NS. This study also found that phonetic approximation existed at the level of speech perception as well as at the level of speech production. It is interesting to note that the characteristics of learners' speech perception were at an intermediate stage between NS and beginners.

5. Universal principles

NS results showed the universal tendency for the relative duration of VOT to be conditioned by the place of articulation. This tendency, however, was not observed in beginning learners' speech production. The results of the present study imply that what has been claimed as a 'phonetic universal' does not necessarily apply to the speech production of beginning learners, who have not fully acquired the phonology of the L2.

Interestingly, advanced learners' results showed that these universal principles were operative in their speech production, which seems to suggest that this occurs only when learners' proficiency levels increase and their speech becomes closer to the phonetic norm of NS. Therefore, the results obtained in this dissertation revealed theoretical insufficiency of what has been claimed as genetically innate Universal Grammar (Chomsky and Halle 1968). Consequently, we can say that this supposedly universal principle is a rather shallow notion and it cannot account for the reasons why relative duration of VOT does show up with advanced learners' speech production, but not with that of beginners. In other words, it is developmental, but not universal. What is found to be universal, instead, is the fact that both the speech production and perception of L2 learners are developmental and systematic. There has been very little previous study concerning the interrelation between the degrees of acquisition of the phonology of a L2 and universal phonetics at different proficiency levels. This is an area certainly worth investigating for its likely contribution to the theory of universal phonetics and SLA.

The above processes were found to be operative in the shaping of IL phonology through cross-sectional and longitudinal investigations of the speech production of learners at various proficiency levels. For IL researchers who are interested in variations in learner strategies and the acquisition processes, the following findings may be useful: Throughout the longitudinal study, subjects demonstrated different approaches: there were 1) those who modified their speech towards the NS norm (observed in CVCV sequences) and overexaggerated the phonological contrast (CVCCV sequences); 2) those who maintained their timing control, which was largely influenced by L1 transfer; and 3) those who constructed their own timing patterns, which were different from both L1 and the target language.

For language instructors, the above findings suggest that it is important for them to introduce pronunciation practice which focuses on the appropriate durational contrast very early in their students' Japanese learning experience. Once learners construct their own patterns, they are difficult to modify at a later stage.

Furthermore, cognitive scientists may find the results obtained from the examinations of learners' perceptual mechanisms of particular interest (Chapter 3). Here, it was found that the beginning learners had inaccurate perceptual targets which were influenced by their L1 experience. Through investigation of longitudinal data, it was found that the learners' perceptual judgment gradually became more accurate in the context of their inaccurate targets (Chapter 5). We argued that the learners' cognitive mechanism may have categorized the CVCV sequence as a part of the already existing L1 category, so that the required duration for the perception of geminate stop C2 and word-final V2 was therefore longer than that of NS. This study has shown that the hypothesis of equivalence classification (Flege 1987d), which was originally applied to categorization of the level of phonetic types, may be extended to larger units such as a CVCV sequence or a word. The learners' inaccurate categorization of the perceptual targets of the CVCV sequence seems to be related to the problem of a longer duration for a single stop C2 in learners' speech production. It seems that the learners' inaccurate perceptual targets could account for some problems in their speech production.

With regard to the fricative C2, it was found that the reverse applied: both beginning and advanced learners had inaccurate perceptual targets which involved smaller threshold values for the fricative C2, and a tendency to underdifferentiate the durational contrast in their speech production. The results obtained in this study clearly show that even near-native advanced speakers have problems in their perception, while they have fewer problems in their speech production. Attaining the perceptual mechanisms of the NS norm (i.e. hearing like NS) is as challenging as the acquisition of NS norms for speech production (i.e. talking like NS) as far as the durational contrast is concerned.

It is a common assumption that learners have the most difficulty in the perceptual judgment of geminate stops. An unexpected finding of this study was that more problems are involved in the perceptual judgment of geminate fricatives and moraic nasals, than in that of geminate stops and long vowels in advanced learners' data. A similar situation can be observed in speech production and thus this study leads us to question the common assumption that production and perception of a geminate C2 is a major problem.

Perhaps less significant for theoretical linguists, but more useful for Japanese language instructors, the findings of Chapter 7 offer additional insight into the teaching situations. An attempt was made to relate the research findings to the classroom and experiments were conducted to test whether efficient instruction would facilitate natural processes of L2 acquisition. After the pronunciation lesson, some improvement was observed in the learners' speech production and it could be said that instruction facilitated the acquisition process of timing control. Perception, on the other hand, did not show any improvement in terms of the threshold values (i.e. the duration required for the perception of geminate C2 and long V2). The only observed improvement was that the learners' perceptual judgment involved less uncertainty after the pronunciation lesson (i.e. it became more categorical). The difficulties involved in perception which were identified in this study suggest that more listening practice needs to be introduced into the classroom to improve learners' listening comprehension.

Three areas of SLA research not investigated in this dissertation offer directions for future study. The first is further investigation of intermediate to advanced levels. Although this study reported the details of timing organization by beginning and advanced learners, and various changes which occurred from the beginning to the intermediate stage, changes which might occur between the intermediate and advanced stages were not investigated. The present study found that the intermediate-level DFAT3 learners (Chapter 5) still had higher threshold values than NS, while advanced-level learners (Chapter 3) showed similar threshold values to those of NS in respect of the word-final V2 and intervocalic stop C2. The remaining question is to find when this change takes place.

The second area is to investigate the level of acquisition at which universal phonetics is likely to become operative in the speech production of learners as it does in the speech production of NS. At the level of speech production, it was found that the universal tendency for VOT applied in the NS and advanced learners' data, but not in the beginning and intermediate learners' data. Exploring such issues concerning universal phonetics would lead to a more comprehensive picture of the English-Japanese IL phonology.

The third area worthy of further investigation is an examination and assessment of the effectiveness of different teaching methodologies in improving learners' perception. The present study found that both the learners' pronunciation and the evaluation of their performance by NS could be improved by instruction. It was found, however, that certain aspects of perception did not improve to the same degree as speech production. To the best of my knowledge, no previous studies appear to provide any comparative investigations on the effectiveness of various exercises to facilitate the acquisition process of perception in Japanese timing organization.

These areas are certainly worthy of investigation for their likely contribution to theories of language teaching methodology.

List of Appendices

A: Sound Stimuli and Statistics

---

[48] One tone is assigned to each mora in the auditory description of pitch patterns.

[49] Appendix 5.7 and 7.2 are dependent t-tests since the comparison is made within the same group (ie. DFAT learners at the beginning and end of the year, and ANU beginners before and after the pronunciation lesson respectively). All other t-tests are independent t-tests since the comparison is made between different groups.

Appendix 3.1 Order of words presented on cassette tape

NO	DETAILS	NSES	START	END	STATUS
SIDE A	C1V1C2V2				
1	SA07N_A	14	150%	20%	D
2	KA14TE_	14	320%	60%	D
3	RI07K_A	16	360%	60%	D
4	KO14M_A	14	20%	150%	A
5	HO07N_E	14	150%	20%	D
6	KA14T_E	16	360%	60%	D
7	I14S_O	14	320%	60%	D
8	I07SO_	14	320%	60%	D
9	RI14KA_	14	320%	60%	D
10	I14SO_	14	320%	60%	D
11	KO07M_A	14	20%	150%	A
12	SA14N_A	14	20%	150%	A
13	KA07TE_	14	320%	60%	D
14	KA14T_E	16	60%	360%	A
15	RI14K_A	14	320%	60%	D
16	HO07N_E	13	30%	150%	A
17	RI14KA_	16	60%	360%	A
18	I14S_O	14	100%	360%	A
19	KO14M_A	13	140%	20%	D
20	RI07K_A	14	100%	360%	A
21	HO14N_E	13	140%	20%	D
22	RI07KA_	16	60%	360%	A
23	I14SO_	16	60%	360%	A
24	KA07T_E	14	100%	360%	A
25	SA14N_A	13	140%	20%	D
26	KA14TE_	14	100%	360%	A
27	I07S_O	14	320%	60%	D
28	I07SO_	14	100%	360%	A

29	RI14K_A	16	60%	360%	A
30	KA07T_E	14	320%	60%	D
31	RI07KA_	16	360%	60%	D
32	SA07N_A	14	20%	150%	A
33	I07S_O	14	100%	360%	A
34	KO07M_A	14	150%	20%	D
35	KA07TE_	14	100%	360%	A
36	HO14N_E	14	20%	150%	D
SIDE B					
37	RI07K_A	16	60%	360%	A
38	KA14T_E	14	320%	60%	D
39	RI14KA_	14	100%	360%	A
40	SA07N_A	13	140%	20%	D
41	I07SO_	16	360%	60%	D
42	I07S_O	16	360%	60%	D
43	HO14N_E	14	150%	20%	D
44	KO07M_A	13	140%	20%	D
45	KA07TE_	16	60%	360%	A
46	KA14TE_	16	360%	60%	D
47	RI07KA_	14	320%	60%	D
48	I14SO_	14	100%	360%	A
49	I14S_O	16	360%	60%	D
50	HO07N_E	14	20%	150%	A
51	SA14N_A	14	150%	20%	D
52	KA14T_E	14	100%	360%	A
53	RI14K_A	16	360%	60%	D
54	KO14M_A	14	150%	20%	D
55	I07SO_	16	60%	360%	A
56	KA14TE_	16	60%	360%	A
57	RI14K_A	14	100%	360%	A
58	KO14M_A	13	30%	150%	A
59	I14S_O	16	60%	360%	A
60	SA14N_A	13	30%	150%	A
61	RI07KA_	14	100%	360%	A
62	HO14N_E	13	30%	150%	A
63	KA07T_E	16	360%	60%	D

64	KO07M_A	13	30%	150%	A
65	I07S_O	16	60%	360%	A
66	HO07N_E	13	140%	20%	D
67	I14SO_	16	360%	60%	D
68	KA07TE_	16	360%	60%	D
69	KA07T_E	16	60%	360%	A
70	RI14KA_	16	360%	60%	D
71	SA07N_A	13	30%	150%	A
72	RI07K_A	14	320%	60%	D

NSES: Number of stimuli in each series

V1 duration: 07=70%; 14=140%

C1V1C2_V2 (e.g. KA07T_E): words with manipulated C2 duration

C1V1C2V2_ (e.g. KA07TE_): words with manipulated V2 duration

A= ascending series; D= descending series

Appendix 3.2 Ascending and descending series- NS

STIMULI	T-test	P value
KA07TE_	1.406	.168
KA14TE_	1.221	.2297
RI07KA_	-1.84	.0736
RI14KA_	.694	.4921
I07SO_	.922	.3622
I14SO_	2.373	.0228
KA07T_E	1.433	.1601
KA14T_E	1.459	.1528
RI07K_A	.651	.5187
RI14K_A	.218	.8284
I07S_O	.653	.5177
I14S_O	1.221	.2297
HO07N_E	2.829	.0074
HO14N_E	1.981	.0549
KO07M_A	1.334	.1902
KO14M_A	3.455	.0014
SA07N_A	1.811	.0781
SA14N_A	3.182	.0029

Appendix 3.3 V1 duration- NS

STIMULI	F-test	P value
KA07/14TE_	10.018	.0022
RI07/14KA_	4.562	.0359
I07/14SO_	19.088	.0001
KA07/14T_E	3.071	.0837
RI07/14K_A	4.643	.0343
I07/14S_O	.854	.3583
HO07/14N_E	.177	.6752
KO07/14M_A	3.971	.0499
SA07/14N_A	12.238	.0008

Appendix 3.4 Question marks indicated by Beginners

1)	KA14TE_	A	160%	BS
2)	KA07T_E	D	220%	NP
3)	RI14KA_	D	180%	NP
4)	RI07K_A	A	200%	NP
5)	I07SO_	A	180%	NP
6)	I07SO_	D	180%	BS
7)	I14SO_	D	220%	BS
8)	HO14N_E	A	50%	NP
9)	KO07M_A	A	50%	LB
10)	KO07M_A	D	60%	NP
11)	SA07N_A	D	70%	NB

Appendix 3.5 NS/Beginners (perceptual boundary)

STIMULI	T-test	P value
KA07TE_	-3.68	.0004
KA14TE_	-2.486	.015
RI07KA_	-2.728	.0079
RI14KA_	-3.324	.0014
I07SO_	-3.58	.0006
I14SO_	-2.315	.0232
KA07T_E	-2.535	.0132
KA14T_E	.778	.439
RI07K_A	-3.449	.0009
RI14K_A	-1.969	.0525

I07S_O	1.845	.0688
I14S_O	5.057	.0001
HO07N_E	-3.871	.0002
HO14N_E	-2.163	.0336
KO07M_A	-5.514	.0001
KO14M_A	-3.224	.0018
SA07N_A	-2.547	.0128
SA14N_A	3.01	.0035

Appendix 3.6 Ascending and descending series- Beginners

STIMULI	T-test	P value
KA07TE_	-3.783	.0005
KA14TE_	-3.145	.0032
RI07KA_	-2.897	.0062
RI14KA_	-3.074	.0039
I07SO_	-4.391	.0001
I14SO_	-6.042	.0001
KA07T_E	-3.451	.0014
KA14T_E	-2.288	.0278
RI07K_A	-2.667	.0112
RI14K_A	-2.491	.0172
I07S_O	-2.794	.0081
I14S_O	-.991	.3278
HO07N_E	-3.221	.0026
HO14N_E	-3.142	.0032
KO07M_A	-4.333	.0001
KO14M_A	-1.615	.1145
SA07N_A	-2.526	.0158
SA14N_A	.375	.7096

Appendix 3.7 V1 duration- Beginners

STIMULI	F-test	P value
KA07/14TE_	.757	.3869
RI07/14KA_	3.166	.0792
I07/14SO_	2.049	.1564
KA07/14T_E	3.274	.0744

RI07/14K_A	.028	.8672
I07/14S_O	4.935	.0293
HO07/14N_E	1.394	.2414
KO07/14M_A	.274	.6025
SA07/14N_A	5.374	.0231

Appendix 3.8 Question marks indicated by Advanced learners

KA14T_E	D	220%	RA
RI14K_A	A	180%	RA
I14S_O	A	180%	RA
HO14N_E	A	100%	RA
SA07N_A	D	70%	RA

Appendix 3.9 NS/Advanced (perceptual boundary)

STIMULI	T-test	P value
KA07TE_	.136	.8923
KA14TE_	.517	.6065
RI07KA_	-.111	.9119
RI14KA_	-.204	.8389
I07SO_	.075	.9401
I14SO_	1.361	.1778
KA07T_E	-.065	.9482
KA14T_E	-.7	.486
RI07K_A	1.956	.0544
RI14K_A	.709	.4809
I07S_O	2.869	.0054
I14S_O	4.327	.0001
HO07N_E	-3.558	.0007
HO14N_E	-3.555	.0007
KO07M_A	-4.543	.0001
KO14M_A	-2.303	.0243
SA07N_A	-.433	.6666
SA14N_A	1.576	.1196

Appendix 3.10 Ascending and descending series- Advanced

STIMULI	T-test	P value
KA07TE_	-.53	.5999
KA14TE_	.307	.7612
RI07KA_	-.189	.8512
RI14KA_	-1.358	.1847
I07SO_	0	
I14SO_	.784	.439
KA07T_E	-.415	.6809
KA14T_E	.789	.4363
RI07K_A	0	
RI14K_A	-1.472	.1513
I07S_O	-1.327	.1946
I14S_O	2.126	.0418
HO07N_E	.815	.4212
HO14N_E	.799	.4305
KO07M_A	-1.479	.1496
KO14M_A	-.831	.4125
SA07N_A	.611	.5455
SA14N_A	2.122	.0422

Appendix 3.11 V1 duration- Advanced

STIMULI	F-test	P value
KA07/14TE_	2.366	.1293
RI07/14KA_	1.421	.238
I07/14SO_	1.592	.2119
KA07/14T_E	2.929	.0922
RI07/14K_A	8.321	.0054
I07/14S_O	1.644	.2047
HO07/14N_E	.619	.4345
KO07/14M_A	0	
SA07/14N_A	1.741	.1921

Appendix 4.1 Average Values- NS

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	46.38	69.93	117.43	17.76	127.63	379.13
	13.31	25.01	29.63	5.79	26.57	59.62
2 katte	38.94	86.56	278.22	15.27	118.98	537.98
	13.19	22.01	44.52	4.43	22.02	73.61
3 katee	43.27	67.58	109.17	18.78	269.51	508.31
	10.72	23.65	19.61	6.06	57.53	86.33
4 rika	19.55	71.88	94.51	24.71	116.33	326.97
	11.47	13.95	19.40	4.25	24.32	42.07
5 rikka	26.11	86.80	244.62	28.26	107.56	493.34
	13.42	17.64	43.30	7.27	30.71	81.37
6 rikaa	22.56	80.34	91.07	27.12	235.18	456.29
	14.53	15.18	18.95	5.89	57.93	62.15
7 iso	0.00	75.29	134.61	0.00	129.83	339.73
	0.00	23.73	24.58	0.00	27.23	58.06
8 isso	0.00	92.86	279.83	0.00	123.09	495.77
	0.00	27.75	41.18	0.00	28.49	78.55
9 isoo	0.00	78.73	137.72	0.00	294.93	511.38
	0.00	26.79	19.45	0.00	44.87	63.27
10 koma	43.62	77.23	74.68	0.00	118.86	314.40
	8.03	25.23	16.25	0.00	22.30	53.70
11 komma	40.15	90.64	192.80	0.00	117.10	440.70
	9.27	17.31	25.46	0.00	32.17	51.27
12 hone	64.05	77.47	77.79	0.00	132.29	351.60
	24.61	28.36	17.48	0.00	21.94	48.90
13 honne	62.00	99.56	226.81	0.00	139.07	527.44
	21.10	21.00	45.10	0.00	22.75	73.86
14 saga	131.84	98.14	61.51	0.00	110.65	402.13
	33.08	11.40	14.09	0.00	22.45	40.78
15 sanga	126.07	109.30	114.63	67.22	103.25	520.47
	25.91	15.23	22.10	20.58	28.02	45.56

(The average figures in 'msec' (above) and 'standard deviation' (below) are given in each word.)

Appendix 4.2 Auditory impression- Beginners

Obstruents/Vowels

JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kateʔ	katte	kate	katee	kateʔ	kateʔ	kaatheʔ	katteʔ	katteʔ	katte	kateʔ
2 katteʔ	katte	kaate	kaatte	katteʔ	katteʔ	katteʔ	katteʔ	kattee	kattee	kattee
3 kateeʔ	katee	kaatee	kateɨ	katee	kateeʔ	katee	katee	kattee	katee	kattee
4 ɾikaʔ	ɾikkaʔ	ɾikaʔ	ɾikaʔ	ɾika -ɾikkaʔ	ɾikaʔ	ɾikkaʔ	ɾikkaʔ	ɾikkaʔ	ɾikka	ʁikkaʔ
5 ɾikkaʔ	ɾikkaʔ	ɾiika	ɾikkaʔ	ɾikkaʔ	ɾikkaʔ	ɾikkaʔ	ɾikkaʔ	ɾiikaʔ	ɾikka	ʁikkaʔ
6 ɾikaa	ɾikkaʔ	ɾiika	ɾikaaʔ	ɾikaa -ɾikkaa	ɾikaa	ɾikkaʔ -ɾikkaa	ɾikkaa	ɾikkaa	ɾikka	ʁikkaa
7 isoʔ	isoo	isoʔ	isoʔ	issoʔ	isoʔ	issoʔ	issoʔ	isoo	isso	iso
8 issoʔ	isoʔ	iso	issoʔ	issoʔ	issoo	issoʔ	issoʔ	issoo	issoo	issoʔ
9 isooʔ	isoo	iisoo	issoo	isoo	isoo	isoo	issoo	isoo	issoo	isoo

Obstruents/Vowels-pitch assignment

JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 LH	LH	LH	LHH	LH	LH	LLH	LH	LH	LH	HL
2 LH	LH	LH	LLH	LH	LH	LH	LH	LHH	LHH	LHH
3 LHH	LHH	LLHH	LHL	LHH	LHH	LHH	LHH	LHH	LHH	LHH
4 HL	HL	HL	HL	HL	HL	HL	HL	LH	HL	HL
5 HL	HL	HL	HL	HL	HL	HL	HL	LHH	HL	HL
6 HLL	HLL	HHL	HLL	HLL	HLL	HLL	HLL	LHH	HLL	HLL
7 LH	HL	LH	LH	LH	LH	LH	LH	LHH	LH	LH
8 LH	HL	LH	LH	LH	LHH	LH	LH	LHH	LHH	LH
9 LHH	LHH	LLHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH

Nasals

JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 komaʔ	koomaʔ	koomaʔ	komaʔ	komaʔ	komaʔ	kommaʔ	koomaʔ	komaʔ	kommaa	koma
11kommaʔ	kommaʔ	kooma	kommaa	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaa	kommaʔ
12 honeʔ	honeeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honneʔ	honneʔ	honeʔ	honne	honne
13 honneʔ	honee	honee	honne	honneʔ	honneʔ	honeʔ	honneʔ	honneʔ	honne	honneʔ
14 sagaʔ	sagaa	sagaʔ	sagaʔ	saŋaʔ	sagaʔ	saŋaʔ	saŋaʔ	sag/ŋaʔ	sagaa	saagaa
15 saŋgaʔ	saŋaʔ	saŋga	saŋaʔ	saŋaʔ	saŋgaʔ	saŋgaʔ	saŋgaʔ	saŋaʔ	saŋaa	saŋga

Nasals - pitch pattern

JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 HL	HHL	HHL	HLL	HL	HL	HHL	HHL	LH	HHLL	HL
11 HLL	HHL	HHL	HL	HLL	HLL	HLL	HHL	HHL	HHLL	HHL

12 LH	HLL	LH	HL	LH	LH	LLH	LLH	LH	LLH	LLH
13 LHH	HLL	LHH	LLH	LLH	LLH	LH	LLH	LLH	LLH	LLH
14 HL	HLL	HL	HL	HL	HL	HL	HL	LH	HLL	HHLL
15 HLL	HL	HHL	HL	HL	HHL	HLL	HHL	LH	HHLL	HHL

The pitch patterns shown in bold type indicate inappropriate tone (e.g. HL -> LH)

Appendix 4.3 Average Values- Beginners

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	45.11	95.07	176.86	35.59	162.58	515.21
	15.66	30.32	59.51	18.56	41.60	97.52
2 katte	43.00	118.52	255.35	29.88	198.27	645.01
	13.43	59.73	75.05	16.52	70.35	99.24
3 katee	40.96	86.07	160.66	37.34	339.78	664.81
	11.70	45.37	56.93	18.79	47.25	100.77
4 rika	31.70	75.10	184.76	31.29	146.93	469.78
	17.60	19.14	59.77	15.33	27.90	72.90
5 rikka	38.12	97.15	296.86	32.94	157.88	622.95
	20.43	45.65	99.91	12.71	34.28	76.84
6 rikaa	39.59	97.44	214.00	41.53	295.48	688.04
	21.34	41.76	84.34	15.66	137.70	156.23
7 iso	0.00	81.99	173.22	0.00	176.78	431.99
	0.00	24.63	43.20	0.00	53.16	63.79
8 isso	0.00	100.33	242.25	0.00	195.83	538.42
	0.00	33.10	51.46	0.00	80.50	102.48
9 isoo	0.00	90.86	201.40	0.00	358.09	650.35
	0.00	40.75	53.37	0.00	69.76	61.20
10 koma	49.18	126.38	122.24	0.00	127.33	425.12
	13.46	51.23	47.38	0.00	42.91	93.49
11 komma	55.43	122.23	208.25	0.00	130.76	516.67
	17.82	56.78	50.59	0.00	54.68	52.74
12 hone	72.27	92.24	151.51	0.00	168.94	484.96
	20.14	27.93	72.86	0.00	81.28	102.98
13 honne	82.13	93.31	197.58	0.00	202.54	575.55
	20.82	30.06	63.57	0.00	109.23	75.42
14 saNa	142.60	117.38	154.79	0.00	135.87	550.63
	37.41	49.55	42.26	0.00	32.84	107.83
15 saNNa	127.33	100.79	166.38	0.00	145.70	540.20

	30.10	24.67	60.18	0.00	42.40	106.27
14 saga	130.62	148.08	102.49	0.00	186.42	567.61
	22.81	45.41	43.19	0.00	70.28	108.09
15 sanga	107.87	119.92	137.13	53.72	141.35	559.98
	34.02	16.89	37.68	10.28	10.30	69.77

(The average figures in 'msec' (above) and 'standard deviation' (below) are given in each word.)

Appendix 4.4 Word duration ratios- Beginners

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	100	100	100	100	100	100	100	100	100	100	100
2 katte	142	95	138	133	133	136	109	113	152	108	160
3 katee	134	88	151	127	140	145	129	144	136	106	148
4 rika	100	100	100	100	100	100	100	100	100	100	100
5 rikka	151	113	132	196	120	155	143	125	117	121	130
6 rikaa	140	93	104	222	162	222	100	173	162	109	173
7 iso	100	100	100	100	100	100	100	100	100	100	100
8 isso	146	102	123	162	104	177	096	110	121	127	143
9 isoo	151	154	199	155	145	173	126	169	130	136	141

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 koma	100	100	100	100	100	100	100	100	100	100	100
11komma	140	106	132	211	142	162	95	121	109	97	103
12 hone	100	100	100	100	100	100	100	100	100	100	100
13 honne	150	101	206	145	106	139	96	112	126	93	115
14 saga	100	100	100	100	100	100	100	100	100	100	100
15 sanga	129	67	128	86	123	97	97	102	106	94	84

Appendix 4.5 Long/short contrast ratios- Beginners

C2

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	42	109	79	53	59	46	66	81	51	90	67
2 katte	100	100	100	100	100	100	100	100	100	100	100
3 katee	39	35	84	56	53	46	64	87	60	72	77
4 rika	39	81	88	23	66	41	58	69	122	69	59
5 rikka	100	100	100	100	100	100	100	100	100	100	100
6 rikaa	37	72	68	45	65	44	51	89	164	70	121
7 iso	48	86	77	52	83	53	80	85	63	82	60

8 isso	100	100	100	100	100	100	100	100	100	100	100
9 isoo	49	107	83	70	72	67	69	106	73	116	77

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 koma	39	51	76	39	37	32	86	55	66	92	64
11komma	100	100	100	100	100	100	100	100	100	100	100
12 hone	34	64	62	60	65	54	105	78	69	98	81
13 honne	100	100	100	100	100	100	100	100	100	100	100
14 saga	34	100	46	121	56	57	76	100	106	85	30
15 sanga	100	100	100	100	100	100	100	100	100	100	100

V2

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	47	56	53	61	43	40	34	33	52	62	46
2 katte	44	61	71	41	53	38	35	41	75	73	96
3 katee	100	100	100	100	100	100	100	100	100	100	100
4 rika	49	107	124	33	39	25	92	30	40	91	57
5 rikka	46	122	162	34	38	24	96	34	47	91	55
6 rikaa	100	100	100	100	100	100	100	100	100	100	100
7 iso	44	54	43	46	38	39	42	33	73	73	60
8 isso	42	48	48	62	32	82	37	28	64	96	59
9 isoo	100	100	100	100	100	100	100	100	100	100	100

Appendix 4.6 C2/V1 and V2/V1 ratios- Beginners

C2/V1 ratios

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	1.68	2.39	0.93	1.04	1.92	1.47	1.42	3.42	2.32	2.26	2.54
2 katte	3.21	2.51	0.75	0.71	3.04	2.99	2.58	4.70	3.61	2.27	3.10
3 katee	1.62	1.40	0.72	1.08	2.18	1.96	1.99	3.71	2.65	2.55	2.69
4 rika	1.31	3.13	1.73	0.95	2.90	1.36	3.59	3.75	2.62	2.88	3.13
5 rikka	2.82	5.02	1.01	3.80	3.45	2.80	6.54	5.41	0.87	3.28	5.29
6 rikaa	1.13	2.88	0.79	1.47	2.26	0.86	3.10	4.58	2.54	2.45	5.21
7 iso	1.79	2.76	1.97	1.69	2.58	1.20	1.67	3.46	1.94	2.96	2.03
8 isso	3.01	2.63	1.90	2.18	3.13	2.05	3.73	3.44	2.00	2.89	1.63
9 isoo	1.75	4.33	0.71	2.21	3.04	1.66	2.24	3.94	1.70	3.56	2.46

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 koma	0.97	0.39	0.42	1.39	1.15	0.57	1.37	1.06	1.23	2.20	1.28
11komma	2.13	1.06	0.42	2.15	2.74	1.40	2.08	2.92	2.23	3.44	1.97
12 hone	1.00	0.70	1.34	1.34	1.18	0.75	1.98	3.46	1.48	2.53	2.09
13 honne	2.28	1.25	1.50	2.14	2.52	1.25	1.82	3.71	2.05	2.64	3.23
14 saga	0.63	0.44	1.72	0.87	1.31	0.58	0.85	2.05	1.55	1.61	0.31
15 sanga	1.66	0.73	1.99	0.98	2.27	1.20	1.43	4.52	1.65	2.17	2.40

V2/V1 ratios

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	1.82	1.77	1.36	2.29	1.66	1.57	0.96	1.45	2.31	2.33	2.31
2 katte	1.37	2.17	1.16	0.56	1.92	1.43	1.21	1.98	2.59	2.47	3.93
3 katee	3.99	5.71	1.86	3.69	4.94	5.29	4.10	4.37	4.26	5.26	4.62
4 rika	1.62	1.98	1.86	1.23	3.07	1.28	1.66	1.90	1.84	2.83	3.15
5 rikka	1.24	2.90	1.25	1.20	2.36	1.04	1.85	2.19	0.88	2.24	3.03
6 rikaa	2.93	1.89	0.89	3.01	6.34	2.98	1.79	6.12	3.32	2.66	4.50
7 iso	1.72	5.01	2.33	1.40	1.86	1.66	1.14	1.59	3.17	2.88	2.52
8 isso	1.33	3.66	1.97	1.28	1.56	3.15	1.79	1.14	1.80	3.03	1.19
9 isoo	3.75	11.64	1.83	3.00	6.66	4.67	4.26	4.39	3.29	3.36	3.94

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 koma	1.54	0.60	0.39	1.61	1.29	1.16	1.13	0.59	1.25	2.42	1.27
11komma	1.29	0.79	0.29	2.57	0.67	1.23	1.13	1.10	1.26	3.24	0.78
12 hone	1.71	5.72	2.37	1.47	1.33	1.23	1.11	2.11	1.42	2.42	1.10
13 honne	1.40	6.98	4.73	2.13	1.63	1.25	1.00	1.79	1.53	1.83	1.77
14 saga	1.13	1.46	1.02	1.07	1.73	0.95	0.88	1.10	1.29	2.13	1.05
15 sanga	0.94	1.41	1.34	1.56	2.20	1.04	0.99	0.87	1.66	2.06	1.23

Appendix 4.7 Auditory impression- Advanced learners

Obstruents/Vowels

JAP	AA	BL	CI	RA	RW	SB	SF	TN
1 kate?	kate?	kate?	kate?	kate?	kate?	kate?	kate?	kate?
2 katte?	katte?	katte?	katte?	katte?	katte?	katte?	katte?	katte?
3 katee?	katee?	katee?	katee?	katee?	katee?	katee?	katee?	katee
4 rika?	rika?	rika?	rika?	rika?	rika?	rika?	rika?	rika?
5 rikka?	rikka?	rikka?	rikka?	rikka?	rikka?	rikka?	rikka?	rikka?
6 rikaa	rikaa?	rikaa?	rikaa?	rikaa?	rikaa?	rikaa?	rikaa?	rikaa

7 isoʔ	isoʔ	isoʔ	isoʔ	isoʔ	isoʔ	isoʔ	isoʔ	isoʔ
8 issoʔ	isoʔ- issoʔ	issoʔ	issoʔ	issoʔ	issoʔ	issoʔ	issoʔ	issoʔ
9 isooʔ	isooʔ	isooʔ	isooʔ	isooʔ	isooʔ	isooʔ	isooʔ	isooʔ

Obstruents/Vowels-pitch assignment

JAP	AA	BL	CI	RA	RW	SB	SF	TN
1 LH	LH	LH	LH	LH	LH	LH	LH	**HL**
2 LH	LH	LH	LH	LH	LH	LH	LH	**HL**
3 LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH	**LHL**
4 HL	HL	HL	HL	HL	HL	HL	HL	HL
5 HL	HL	HL	HL	HL	HL	HL	HL	HL
6 HLL	HLL	HLL	HLL	HLL	HLL	HLL	HLL	HLL
7 LH	LH	LH	LH	LH	LH	LH	LH	LH
8 LH	LH	LH	LH	LH	LH	LH	LH	LH
9 LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH

Nasals

JAP	AA	BL	CI	RA	RW	SB	SF	TN
10 komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ
11 kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ
12 honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ
13 honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ
14 sagaʔ	sagaʔ	saŋaʔ	sagaʔ	sagaʔ	sagaʔ	sagaʔ	sagaʔ	saŋaʔ
15 saŋgaʔ	saŋŋaʔ	saŋŋaʔ	saŋgaʔ	saŋgaʔ	saŋgaʔ	saŋgaʔ	saŋgaʔ	saŋŋaʔ

Nasals - pitch pattern

JAP	AA	BL	CI	RA	RW	SB	SF	TN
10 HL	HL	HL	HL	HL	HL	HL	HL	HL
11 HLL	HLL	HLL	HLL	HLL	HLL	HLL	HLL	HLL
12 LH	LH	LH	LH	LH	LH	LH	LH	LH
13 LHH	LHH	LHH	LLH	LLH	LLH	LHH	LLH	LHH
14 HL	HL	HL	HL	HL	HL	HL	HL	HL
15 HLL	HLL	HLL	HLL	HHL	HHL	HHH	HLL	HLL

The pitch patterns shown in bold type indicate inappropriate tone (e.g. HL -> LH)

Appendix 4.8 Average Values- Advanced learners

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	49.79 10.12	72.84 17.19	102.75 19.86	30.06 10.33	139.98 29.06	395.42 41.87
2 katte	45.33 13.22	76.99 19.04	287.62 70.08	20.75 7.50	139.95 35.90	570.63 77.36
3 katee	48.50 12.93	62.61 18.55	114.18 17.23	30.42 10.84	285.39 48.05	541.10 48.78
4 rika	15.58 9.25	64.51 16.75	86.76 23.83	32.61 7.51	123.00 25.14	322.46 41.44
5 rikka	20.20 14.49	74.72 23.77	283.03 67.94	28.73 10.65	121.55 22.26	528.23 67.39
6 rikaa	23.30 15.75	69.17 15.73	96.08 22.13	37.52 10.95	258.36 50.30	484.43 59.34
7 iso	0.00 0.00	73.30 24.13	141.26 20.77	0.00 0.00	142.47 27.47	357.02 47.75
8 isso	0.00 0.00	79.94 15.12	258.98 49.06	0.00 0.00	126.10 24.72	465.03 44.28
9 isoo	0.00 0.00	70.61 19.87	163.85 25.67	0.00 0.00	299.80 38.05	534.25 45.01
10 koma	54.61 20.64	68.86 12.64	89.70 21.92	0.00 0.00	114.65 29.79	327.82 40.23
11 komma	61.55 24.07	86.11 26.70	190.43 37.60	0.00 0.00	118.43 36.46	456.52 72.66
12 hone	80.18 23.26	79.86 19.54	84.79 19.51	0.00 0.00	138.50 25.08	383.33 40.46
13 honne	77.77 20.02	89.96 20.49	251.41 33.94	0.00 0.00	139.30 22.02	558.44 54.63
14 saNa	149.40 23.68	123.75 19.76	82.18 9.54	0.00 0.00	131.53 40.23	486.87 75.83
15 saNNa	114.71 13.36	100.16 14.85	199.69 44.89	0.00 0.00	127.99 20.02	542.54 82.12
14 saga	120.88 13.76	109.58 14.49	79.08 29.63	0.00 0.00	116.82 22.65	426.37 44.30
15 sanga	114.27 17.91	107.69 22.60	141.79 24.25	89.32 21.28	120.55 0.00	573.61 9.00

Appendix 5.1 Ascending and descending series- DFAT1

STIMULI	T-test	P value
KA07TE_	-1.116	.2834
KA14TE_	-.913	.3765
RI07KA_	-.95	.358
RI14KA_	-.939	.3638
I07SO_	-1.062	.3063
I14SO_	-1.024	.3231
KA07T_E	-2.575	.022
KA14T_E	-2.538	.0237
RI07K_A	-3.68	.0025
RI14K_A	-2.931	.0109
I07S_O	-1.178	.2583
I14S_O	-1.802	.0931
HO07N_E	-.942	.3622
HO14N_E	-2.411	.0302
KO07M_A	-2.96	.0103
KO14M_A	-.914	.3761
SA07N_A	-.993	.3376
SA14N_A	-.335	.7436

Appendix 5.2 V1 duration- DFAT1

STIMULI	F-test	P value
KA07/14TE_	.169	.684
RI07/14KA_	0	
I07/14SO_	.034	.8556
KA07/14T_E	0	
RI07/14K_A	.149	.7028
I07/14S_O	1.074	.309
HO07/14N_E	1.526	.227
KO07/14M_A	.511	.4805
SA07/14N_A	.264	.6117

Appendix 5.3 Question marks indicated by DFAT2

1)	KA07TE_	A	200%	PK
2)	I07S_O	D	160%	DS

Appendix 5.4 Ascending and descending series- DFAT2

STIMULI	T-test	P value
KA07TE_	-1.949	.0716
KA14TE_	-1.125	.2795
RI07KA_	.434	.6711
RI14KA_	-2.2	.0451
I07SO_	-1.303	.2137
I14SO_	-2.357	.0335
KA07T_E	-1.691	.113
KA14T_E	-1.465	.1649
RI07K_A	-2.658	.0187
RI14K_A	-.487	.6341
I07S_O	0	
I14S_O	.536	.6004
HO07N_E	1.93	.0758
HO14N_E	.17	.8677
KO07M_A	-1.616	.1284
KO14M_A	-1.932	.0739
SA07N_A	1.779	.0969
SA14N_A	.451	.6596

Appendix 5.5 V1 duration- DFAT2

STIMULI	F-test	P value
KA07/14TE_	.51	.481
RI07/14KA_	.055	.8164
I07/14SO_	.287	.5963
KA07/14T_E	1.349	.2553
RI07/14K_A	1.426	.2424
I07/14S_O	1.605	.2156
HO07/14N_E	.585	.4512
KO07/14M_A	.014	.9062
SA07/14N_A	4.684	.0395

Appendix 5.6 Question marks indicated by DFAT3

1)	KA14TE_	A	200%	PK
2)	KA07T_E	D	220%	PK

Appendix 5.7 Perceptual boundary- DFAT1/DFAT3 (dependent t-test)

STIMULI	T-test	P value
KA07TE_	-1.017	.3841
KA14TE_	-1.009	.3871
RI07KA_	-.378	.7306
RI14KA_	.151	.8893
I07SO_	-.305	.7806
I14SO_	-1.203	.3154
KA07T_E	1.711	.1856
KA14T_E	-1.4	.256
RI07K_A	.076	.944
RI14K_A	-1.271	.2933
I07S_O	.617	.5808
I14S_O	-1.842	.1627
HO07N_E	-.19	.8618
HO14N_E	-.271	.8043
KO07M_A	1.399	.2562
KO14M_A	.426	.6987
SA07N_A	-3.082	.0541
SA14N_A	.502	.65

Appendix 5.8 Ascending and descending series- DFAT3

STIMULI	T-test	P value
KA07TE_	.522	.6097
KA14TE_	-.411	.6875
RI07KA_	-1.288	.2186
RI14KA_	-.134	.8955
I07SO_	1.23	.2389
I14SO_	-.832	.4193
KA07T_E	-.55	.5911
KA14T_E	-.884	.3917
RI07K_A	-.228	.8232
RI14K_A	-.497	.6271
I07S_O	-.753	.4637
I14S_O	.129	.8995
HO07N_E	-.661	.5191
HO14N_E	.687	.5034

KO07M_A	-.332	.7448
KO14M_A	0	
SA07N_A	.12	.9063
SA14N_A	0	

Appendix 5.9 V1 duration- DFAT3

STIMULI	F-test	P value
KA07/14TE_	.864	.3606
RI07/14KA_	.095	.7598
I07/14SO_	.382	.5416
KA07/14T_E	1.76	.1953
RI07/14K_A	1.014	.3226
I07/14S_O	.202	.6567
HO07/14N_E	1.351	.2549
KO07/14M_A	.669	.4204
SA07/14N_A	1.128	.2973

Appendix 6.1 Auditory impression- DFAT1

	AP	DS	PK	PV
1 kate	[katteʔ] LH	[kattee] LHH	[katteʔ] LH	[katte] LH
2 katte	[katteʔ] LH	[kattee] LHH	[katteeʔ] LHH	[katte] LH
3 katee	[katteeʔ] LHH	[kattee] LHH	[kattee] LHH	[kattee] LHH
4 rika	[ɾikkaʔ] HL	[ɾikka] HL	[ɾikka] HL	[ʀikka] HL
5 rikka	[ɾikkaʔ] HL	[ɾikka] HL	[ɾiika] HHL	[ʀikka] HL
6 rikaa	[ɾikkaa] HLL	[ɾikkaa] HLL	[ɾikkaa] HLL	[ʀikka] HL
7 iso	[issoʔ] LH	[issoo] LHH	[issooʔ] LHH	[isso] LH
8 isso	[issoʔ] LH	[issoo] LHH	[issoo] LHH	[isso] LH
9 isoo	[issooʔ] LHH	[issoo] LHH	[issoʔ] LH	[issoo] LHH

10 koma	[komaʔ] HL	[kooma] HHL	[kooma] HHL	[koma] HL
11 komma	[kommaʔ] HLL	[komma] HLL	[komma] HHL	[komma] HLL
12 hone	[honeʔ] LH	[honee] LHH	[honee] LHH	[honeʔ] LH
13 honne	[honnee] LLHH	[honee] LHH	[honee] LHH	[honne] LLH
14 saga	[saŋaʔ] HL	[saŋa] HL	[saŋŋa] HHL	[saaŋaʔ] HHL
15 sanga	[saŋŋaʔ] HLL	[saŋŋa] HLL	[saŋŋa] HHL	[saŋgaʔ] HLL

Appendix 6.2 Average values- DFAT1

AP

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	63.6 13.0	129.5 22.1	218.7 38.8	48.6 12.1	207.9 4.2	668.3 27.9
2 katte	51.2 7.0	155.4 9.7	310.8 43.6	32.4 2.8	224.0 32.5	773.8 26.2
3 katee	42.0 7.3	136.4 12.3	305.8 13.9	62.8 12.6	350.9 11.7	898.0 19.9
4 rika	21.1 3.5	107.4 5.4	247.3 6.4	42.6 4.7	146.9 17.0	565.3 22.9
5 rikka	16.3 0.9	115.7 17.1	378.3 31.7	32.3 0.8	156.1 10.4	698.8 11.8
6 rikaa	41.1 17.5	91.3 13.8	284.8 19.9	39.9 6.6	421.2 41.3	878.3 74.3
7 iso	0.0 0.0	154.1 16.8	269.6 21.7	0.0 0.0	203.0 39.2	626.7 46.5
8 isso	0.0 0.0	160.4 10.2	306.1 8.0	0.0 0.0	197.5 35.6	663.9 37.5
9 isoo	0.0 0.0	150.4 20.0	274.0 10.1	0.0 0.0	377.6 39.6	802.0 32.2
10 koma	71.8 10.1	151.1 11.5	99.3 13.2	0.0 0.0	227.7 15.7	550.0 6.9
11 komma	86.9	177.1	183.7	93.9	206.5	654.3

	3.4	23.4	13.6	38.8	27.4	10.7
12 hone	64.5	148.5	128.9	0.0	230.8	572.7
	10.8	8.0	3.3	0.0	23.1	30.4
13 honne	84.5	181.0	237.5	0.0	287.7	790.8
	13.4	21.7	47.6	0.0	71.5	45.5
14 saNa	155.3	177.8	95.6	0.0	220.7	649.4
	25.3	11.3	6.3	0.0	14.8	21.2
15 saNNa	160.4	196.3	160.3	0.0	223.8	740.7
	26.2	12.7	34.3	0.0	20.9	30.6

DS

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	41.9	83.8	175.0	35.7	242.9	579.4
	6.6	9.9	18.6	9.9	12.4	26.9
2 katte	41.9	90.4	254.6	41.9	274.8	703.6
	10.2	2.6	8.8	8.8	6.3	11.2
3 katee	43.3	81.2	264.4	44.0	259.4	692.4
	17.0	16.4	14.5	9.8	1.3	11.6
4 rika	37.7	113.3	196.8	61.2	169.3	578.2
	13.6	6.4	23.1	17.1	33.2	42.4
5 rikka	50.4	157.7	293.0	55.9	192.4	749.5
	10.4	15.8	13.1	27.3	13.3	14.5
6 rikaa	37.1	126.3	293.6	55.6	192.2	704.7
	5.0	9.6	13.6	27.3	19.3	36.9
7 iso	0.0	105.7	268.8	0.0	321.3	695.8
	0.0	12.5	22.6	0.0	34.7	11.3
8 isso	0.0	107.7	227.3	0.0	242.0	577.0
	0.0	22.4	48.7	0.0	19.5	39.1
9 isoo	0.0	145.9	250.7	0.0	315.1	711.7
	0.0	10.6	18.8	0.0	3.5	13.6
10 koma	45.8	164.1	100.0	0.0	150.7	460.5
	14.9	11.4	17.3	0.0	5.7	19.1
11 komma	52.8	106.7	159.5	0.0	111.1	430.1
	1.9	10.7	17.7	0.0	6.1	33.3
12 hone	95.4	115.2	175.3	0.0	292.6	678.5
	24.7	22.0	27.7	0.0	22.2	30.1
13 honne	112.6	101.4	172.9	0.0	292.4	679.2
	25.7	2.5	7.4	0.0	16.7	19.8

14 saNa	165.6 30.6	135.5 16.1	173.6 21.8	0.0 0.0	181.6 16.0	656.3 42.8
15 saNNa	166.6 18.1	118.3 14.2	180.9 1.1	0.0 0.0	208.5 6.1	674.3 25.2

PK

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	59.1 9.9	135.9 9.0	303.6 9.6	54.0 18.0	222.9 24.9	775.5 23.8
2 katte	90.1 13.9	139.0 18.5	295.3 39.0	42.0 5.8	296.0 74.5	862.4 137.4
3 katee	67.2 1.3	133.7 15.2	329.1 28.3	56.5 4.8	373.9 62.6	960.4 60.1
4 rika	16.8 2.3	192.6 31.8	175.0 55.7	74.6 14.9	207.5 1.7	666.5 30.6
5 rikka	54.3 12.6	249.7 30.5	137.9 23.6	55.0 14.3	199.2 18.1	696.0 19.1
6 rikaa	12.8 4.3	170.5 24.5	164.8 32.8	54.1 10.4	293.2 24.4	695.3 20.0
7 iso	0.0 0.0	189.1 34.5	331.6 49.3	0.0 0.0	272.1 59.9	792.8 78.4
8 isso	0.0 0.0	301.1 35.0	233.1 48.7	0.0 0.0	354.6 14.3	888.8 35.9
9 isoo	0.0 0.0	254.8 30.1	277.7 53.7	0.0 0.0	294.1 19.8	826.6 33.8
10 koma	78.8 4.3	237.1 10.8	157.0 15.1	0.0 0.0	177.9 11.3	650.8 9.6
11 komma	76.2 11.6	100.4 32.0	348.1 18.1	0.0 0.0	167.7 10.0	692.3 14.3
12 hone	134.3 48.5	202.8 19.2	193.4 17.5	0.0 0.0	281.3 3.0	811.8 50.6
13 honne	110.6 33.3	209.1 25.6	163.4 43.5	0.0 0.0	302.4 15.1	785.5 30.9
14 saNa	149.9 10.8	116.2 3.9	264.2 7.9	0.0 0.0	194.3 20.1	724.7 37.5
15 saNNa	152.6 12.7	122.2 9.6	278.7 8.0	0.0 0.0	199.7 11.5	753.2 4.6

PV

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	45.9 5.2	86.2 18.3	215.3 7.9	29.3 2.5	200.9 6.2	577.6 17.7
2 katte	49.1 1.2	67.8 20.1	270.7 30.8	24.7 3.3	214.0 4.7	626.3 7.2
3 katee	52.1 6.9	94.8 8.3	277.3 20.6	47.3 3.5	389.9 7.4	861.4 17.8
4 rika	16.0 6.0	75.4 9.4	251.5 14.9	40.9 0.7	236.6 5.7	620.4 31.2
5 rikka	55.3 8.9	121.9 5.5	564.0 74.2	34.6 1.8	275.9 8.4	1051.7 74.1
6 rikaa	36.2 18.5	98.5 15.9	231.0 43.4	45.7 9.2	253.5 6.9	665.0 43.0
7 iso	0.0 0.0	87.5 4.1	257.7 21.2	0.0 0.0	207.8 14.4	553.0 11.6
8 isso	0.0 0.0	94.0 2.5	299.1 14.4	0.0 0.0	217.4 8.0	610.5 14.7
9 isoo	0.0 0.0	89.7 4.7	274.3 20.4	0.0 0.0	379.3 9.6	743.4 30.3
10 koma	62.2 11.5	114.1 10.1	86.8 3.4	0.0 0.0	138.7 27.4	401.9 28.0
11 komma	62.9 10.2	126.2 8.9	171.5 27.1	0.0 0.0	256.1 18.4	616.7 6.5
12 hone	83.6 22.1	80.5 16.8	92.8 27.0	0.0 0.0	192.9 67.9	449.9 61.4
13 honne	74.8 32.4	86.3 8.2	239.8 6.5	0.0 0.0	324.0 14.4	724.8 39.5
14 saNa	162.3 8.8	255.8 19.8	434.3 52.5	0.0 0.0	149.1 19.2	880.5 42.9
15 saNga	128.5 15.5	96.0 10.6	230.4 0.9	180.6 16.0	133.9 5.8	769.4 12.9

Appendix 6.3 Word duration ratios- DFAT1

	AP	DS	PK	PV
1 kate	100	100	100	100
2 katte	116	121	111	108
3 katee	134	120	124	149
4 rika	100	100	100	100
5 rikka	124	130	104	170
6 rikaa	155	122	104	107
7 iso	100	100	100	100
8 isso	106	83	112	110
9 isoo	128	102	104	134
10 koma	100	100	100	100
11 komma	119	93	106	153
12 kone	100	100	100	100
13 konne	138	100	97	161
14 saga	100	100	100	100
15 sanga	114	103	104	87

Appendix 6.4 Long/short contrast ratios- DFAT1

C2

	AP	DS	PK	PV
1 kate	70	69	103	80
2 katte	100	100	100	100
3 katee	98	104	111	102
4 rika	65	67	127	45
5 rikka	100	100	100	100
6 rikaa	75	100	120	41
7 iso	88	118	142	86
8 isso	100	100	100	100
9 isoo	90	110	119	92
10 koma	54	63	45	51
11 komma	100	100	100	100
12 kone	54	101	118	39
13 konne	100	100	100	100
14 saga	60	96	95	106
15 sanga	100	100	100	100

V2

	AP	DS	PK	PV
1 kate	59	94	60	52
2 katte	64	106	79	55
3 katee	100	100	100	100
4 rika	35	88	71	93
5 rikka	37	100	68	109
6 rikaa	100	100	100	100
7 iso	54	102	93	55
8 isso	52	77	121	57
9 isoo	100	100	100	100

Appendix 6.5 C2/V1 and V2/V1 ratios- DFAT1

C2/V1 ratios

	AP	DS	PK	PV
1 kate	1.69	2.09	2.23	2.50
2 katte	2.00	2.81	2.12	3.99
3 katee	2.24	3.26	2.46	2.92
4 rika	2.30	1.74	0.91	3.34
5 rikka	3.27	1.86	0.55	4.63
6 rikaa	3.12	2.32	0.97	2.34
7 iso	1.75	2.54	1.75	2.95
8 isso	1.91	2.11	0.77	3.18
9 isoo	1.82	1.72	1.09	3.06
10 koma	0.66	0.61	0.66	0.76
11 komma	1.04	1.49	3.47	1.36
12 hone	0.87	1.52	0.95	1.15
13 honne	1.31	1.70	0.78	2.78
14 saga	0.54	1.28	2.27	1.70
15 sanga	0.82	1.53	2.28	2.40

V2/V1 ratios

	AP	DS	PK	PV
1 kate	1.61	2.90	1.64	2.33
2 katte	1.44	3.04	2.13	3.16
3 katee	2.57	3.19	2.80	4.11
4 rika	1.37	1.49	1.08	3.14
5 rikka	1.35	1.22	0.80	2.26

6 rikaa	4.61	1.52	1.72	2.57
7 iso	1.32	3.04	1.44	2.37
8 isso	1.23	2.25	1.18	2.31
9 isoo	2.51	2.16	1.15	4.23
10 koma	1.51	0.92	0.75	1.22
11 komma	1.17	1.04	1.67	2.03
12 hone	1.55	2.54	1.39	2.40
13 honne	1.59	2.88	1.45	3.75
14 saga	1.24	1.34	1.67	0.58
15 sanga	1.14	1.76	1.63	1.39

Appendix 6.6 Auditory impression- DFAT2

	AP	DS	PK	PV
1 kate	[kate] LH	[kattee] LHH	[kattee] LHH	[kattee] LHH
2 katte	[katteʔ] LH	[kattee] LHH	[kattee] LHH	[kattee] LHH
3 katee	[katee] LHH	[kattee] LHH	[kattee] LHH	[kattee] LHH
4 rika	[ɾikka] HL	[ɾikkaʔ] HL	[ɾikka] HL	[ʁikkaa] HLL
5 rikka	[ɾikkaʔ] HL	[ɾikkaʔ] HL	[ɾikka] HL	[ʁikkaa] HLL
6 rikaa	[ɾikkaa] HLL	[ɾikkaa] HLL	[ɾiika(a)] HHL(L)	[ʁikkaa] HLL
7 iso	[isoʔ] LH	[issoo] LHH	[issoʔ] LH	[issoo] LHH
8 isso	[issoo] LHH	[issoo] LHH	[isso] LH	[issoo] LHH
9 isoo	[isoo] LHH	[issoo] LHH	[issoo] LHH	[issoo] LHH
10 koma	[komaʔ] HL	[kommaʔ] HLL	[kommaʔ] HHL	[koomaa] HHLL
11 komma	[kommaʔ] HLL	[kommaʔ] HLL	[kommaʔ] HHL	[koomaa] HHLL
12 hone	[honee] LHH	[honee] LHH	[honeʔ] LH	[honee] LHH

13 honne	[honee] LHH	[honee] LHH	[honnee] LLHH	[honnee] LLHH
14 saga	[saŋaʔ] HL	[saŋŋa] HHL	[saŋŋa] HHL	[saʔŋaa] HLL
15 sanga	[saŋŋaʔ] HLL	[saŋŋa] HLL	[saŋŋaʔ] HHL	[saŋgaa] HHLL

Appendix 6.7 Average values- DFAT2

AP

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	70.2	88.3	133.0	61.1	256.2	608.7
	2.5	4.8	14.5	14.8	9.9	11.0
2 katte	59.1	118.8	307.2	40.1	209.4	734.6
	11.7	9.1	13.0	13.4	6.3	4.0
3 katee	57.2	90.8	122.8	60.3	463.7	794.8
	2.2	7.4	9.4	16.9	9.2	27.7
4 rika	19.3	94.5	185.3	25.5	245.7	570.3
	3.2	6.9	40.8	4.6	10.9	27.0
5 rikka	13.5	95.6	363.9	19.9	192.4	685.3
	0.3	6.2	45.7	5.9	15.6	21.2
6 rikaa	18.7	78.8	225.7	33.2	432.2	788.6
	2.8	8.2	12.9	4.4	3.5	13.9
7 iso	0.0	92.1	159.1	0.0	252.4	503.6
	0.0	4.8	15.2	0.0	26.9	15.2
8 isso	0.0	131.0	327.0	0.0	257.2	715.2
	0.0	9.1	20.1	0.0	11.9	17.5
9 isoo	0.0	102.1	196.9	0.0	477.2	776.2
	0.0	5.8	5.2	0.0	30.3	36.1
10 koma	50.4	107.4	93.1	0.0	181.0	432.0
	5.1	11.8	2.6	0.0	7.2	10.8
11 komma	56.4	156.6	186.6	0.0	236.8	636.4
	11.8	17.8	16.2	0.0	24.9	10.0
12 hone	57.5	111.9	115.3	0.0	259.6	544.3
	17.4	4.6	5.1	0.0	16.3	6.5
13 honne	73.5	133.4	135.7	0.0	495.1	837.8
	8.4	3.8	9.1	0.0	48.9	62.1
14 saNa	143.9	144.4	89.8	0.0	212.5	590.5

	16.0	9.0	8.5	0.0	24.7	32.2
15 saNNa	159.7	201.5	172.2	0.0	242.9	776.3
	26.5	7.6	9.0	0.0	16.4	26.3

DS

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	51.2	112.3	170.3	31.4	306.2	671.5
	16.8	12.3	12.0	11.5	15.8	31.6
2 katte	34.2	126.4	286.1	25.1	351.5	823.2
	3.5	15.9	35.2	8.6	19.1	43.8
3 katee	53.1	119.2	229.3	43.0	394.6	839.3
	16.4	25.9	11.4	8.1	49.1	58.5
4 rika	32.4	108.2	222.1	28.5	159.0	550.2
	13.7	28.8	6.1	9.7	0.9	20.1
5 rikka	28.0	106.7	296.4	39.1	192.5	662.7
	10.5	21.1	8.8	12.8	14.1	34.4
6 rikaa	34.9	160.1	227.0	29.6	418.1	869.7
	8.7	14.4	49.0	1.5	25.5	67.1
7 iso	0.0	161.4	229.9	0.0	346.4	737.7
	0.0	7.9	24.8	0.0	38.0	35.6
8 isso	0.0	137.4	242.6	0.0	375.0	755.0
	0.0	23.4	14.7	0.0	27.2	56.7
9 isoo	0.0	143.2	223.0	0.0	331.1	697.2
	0.0	20.0	7.9	0.0	22.0	34.6
10 koma	43.8	128.8	173.7	0.0	116.9	463.2
	5.9	20.5	22.0	0.0	19.3	39.3
11 komma	56.5	118.8	205.6	0.0	145.9	526.8
	9.4	10.6	4.6	0.0	12.3	14.9
12 hone	97.5	118.9	176.2	0.0	237.9	630.5
	4.4	11.6	17.1	0.0	10.4	16.1
13 honne	115.6	161.4	157.8	0.0	322.4	757.2
	24.6	12.6	15.7	0.0	24.8	24.1
14 saNa	138.5	121.2	187.0	0.0	142.1	588.9
	5.6	8.8	8.6	0.0	5.5	10.8
15 saNNa	139.6	131.1	237.5	0.0	213.6	721.7
	16.8	22.7	30.6	0.0	35.3	72.7

PK

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	66.7	131.7	277.7	49.0	302.7	827.7
	11.7	33.1	19.2	33.5	72.0	41.0
2 katte	58.7	172.0	262.7	23.3	293.3	810.0
	10.4	42.8	27.2	6.7	18.2	20.1
3 katee	67.7	87.3	271.0	37.7	380.7	844.3
	4.9	3.8	17.3	10.4	52.9	47.4
4 rika	17.0	119.0	188.3	44.0	249.3	617.7
	2.6	12.3	10.7	3.6	50.8	60.3
5 rikka	24.7	124.7	224.3	54.3	167.0	595.0
	4.5	26.5	23.4	15.6	24.1	45.1
6 rikaa	35.0	197.0	112.3	50.0	261.7	656.0
	14.0	6.6	24.6	5.3	58.8	80.5
7 iso	0.0	112.3	273.0	0.0	200.7	586.0
	0.0	5.7	21.8	0.0	35.4	54.7
8 isso	0.0	203.0	272.7	0.0	218.0	693.7
	0.0	29.5	12.1	0.0	25.1	49.1
9 isoo	0.0	125.0	260.7	0.0	299.7	685.3
	0.0	11.5	5.5	0.0	39.5	40.2
10 koma	72.3	133.3	184.0	0.0	146.0	535.7
	4.6	31.8	32.2	0.0	31.0	39.4
11 komma	68.3	91.3	297.3	0.0	167.0	624.0
	16.9	3.8	10.0	0.0	19.0	10.8
12 hone	53.7	127.0	195.7	0.0	241.0	617.3
	14.8	17.7	8.4	0.0	13.7	28.3
13 honne	83.3	137.7	248.0	0.0	330.7	799.7
	13.4	17.9	32.0	0.0	2.5	37.9
14 saNa	156.0	172.3	210.3	0.0	187.3	726.0
	22.1	15.6	45.5	0.0	11.5	53.4
15 saNNa	109.0	123.0	239.3	0.0	151.0	622.3
	10.4	33.6	23.6	0.0	10.4	11.9

PV

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	64.7	102.0	233.3	51.3	467.0	918.3
	0.6	6.2	26.8	6.4	28.9	58.6
2 katte	66.0	98.0	309.7	57.0	461.3	992.0
	3.5	20.0	35.0	19.3	57.8	20.0

3 katee	56.0 11.3	99.3 5.9	266.0 21.5	60.0 9.0	439.0 53.7	920.3 43.0
4 rika	36.3 10.7	107.3 6.5	238.7 4.0	40.7 4.9	373.3 45.1	796.3 46.4
5 rikka	58.3 15.0	87.7 9.3	287.0 60.6	44.3 14.3	309.0 18.5	786.3 21.5
6 rikaa	42.3 25.6	106.3 30.6	248.0 27.1	65.7 11.0	473.7 57.1	936.0 75.7
7 iso	0.0 0.0	97.7 8.4	277.7 25.1	0.0 0.0	428.7 22.3	804.0 53.0
8 isso	0.0 0.0	144.0 16.5	253.3 12.1	0.0 0.0	455.0 12.5	852.3 10.2
9 isoo	0.0 0.0	158.3 5.1	272.0 27.5	0.0 0.0	437.7 34.6	868.0 44.5
10 koma	94.7 14.4	237.0 2.6	110.0 7.9	0.0 0.0	386.7 28.6	828.3 29.0
11 komma	92.0 10.6	220.7 4.0	123.3 6.7	0.0 0.0	386.3 53.3	822.3 36.9
12 hone	113.0 7.5	119.7 5.5	179.3 23.2	0.0 0.0	422.0 9.8	834.0 21.3
13 honne	121.3 26.6	116.7 10.1	212.3 28.6	0.0 0.0	445.0 29.5	895.3 17.2
14 saga	189.0 24.4	133.0 32.8	257.0 31.3	0.0 0.0	360.3 17.0	939.3 47.4
15 saNga	173.3 26.4	131.7 30.0	155.7 11.1	111.3 28.0	370.3 38.9	942.3 78.3

Appendix 6.8 Word duration ratios- DFAT2

	AP	DS	PK	PV
1 kate	100	100	100	100
2 katte	121	123	98	108
3 katee	131	125	102	100
4 rika	100	100	100	100
5 rikka	120	120	96	99
6 rikaa	138	158	106	118
7 iso	100	100	100	100
8 isso	142	102	118	106

9 isoo	154	95	117	108
10 koma	100	100	100	100
11 komma	147	114	116	99
12 hone	100	100	100	100
13 honne	154	120	130	107
14 saga	100	100	100	100
15 sanga	131	123	86	100

Appendix 6.9 Long/short contrast ratios- DFAT2

Long/short contrast (C2)

	AP	DS	PK	PV
1 kate	43	60	106	75
2 katte	100	100	100	100
3 katee	40	80	103	86
4 rika	51	75	84	83
5 rikka	100	100	100	100
6 rikaa	62	77	50	86
7 iso	49	95	100	110
8 isso	100	100	100	100
9 isoo	60	92	96	107
10 koma	50	84	62	89
11 komma	100	100	100	100
12 hone	85	112	79	84
13 honne	100	100	100	100
14 saga	52	79	88	96
15 sanga	100	100	100	100

Long/short contrast (V2)

	AP	DS	PK	PV
1 kate	55	78	80	106
2 katte	45	89	77	105
3 katee	100	100	100	100
4 rika	57	38	95	79
5 rikka	45	46	64	65
6 rikaa	100	100	100	100
7 iso	53	105	67	98
8 isso	54	113	73	104
9 isoo	100	100	100	100

Appendix 6.10 C2/V1 and V2/V1 ratios- DFAT2

C2/V1 ratios

	AP	DS	PK	PV
1 kate	1.51	1.52	2.11	2.29
2 katte	2.59	2.26	1.53	3.16
3 katee	1.35	1.92	3.10	2.68
4 rika	1.96	2.05	1.58	2.22
5 rikka	3.81	2.78	1.80	3.27
6 rikaa	2.87	1.42	0.57	2.33
7 iso	1.73	1.42	2.43	2.84
8 isso	2.50	1.77	1.34	1.76
9 isoo	1.93	1.56	2.09	1.72
10 koma	0.87	1.35	1.38	0.46
11 komma	1.19	1.73	3.26	0.56
12 hone	1.03	1.48	1.54	1.50
13 honne	1.02	0.98	1.80	1.82
14 saga	0.62	1.54	1.22	1.93
15 sanga	0.85	1.81	1.95	2.03

V2/V1 ratios

	AP	DS	PK	PV
1 kate	2.90	2.73	2.30	4.58
2 katte	1.76	2.78	1.71	4.71
3 katee	5.11	3.31	4.36	4.42
4 rika	2.60	1.47	2.09	3.48
5 rikka	2.01	1.80	1.34	3.52
6 rikaa	5.48	2.61	1.33	4.46
7 iso	2.74	2.15	1.79	4.39
8 isso	1.96	2.73	1.07	3.16
9 isoo	4.67	2.31	2.40	2.77
10 koma	1.69	0.91	1.10	1.63
11 komma	1.51	1.23	1.83	1.75
12 hone	2.32	2.00	1.90	3.53
13 honne	3.71	2.00	2.40	3.81
14 saga	1.47	1.17	1.09	2.71
15 sanga	1.21	1.63	1.23	2.81

Appendix 6.11 Auditory impression- DFAT3

	AP	DS	PK	PV
1 kate	[kateʔ] LH	[kattee] LHH	[kattee] LHH	[kattee] LHH
2 katte	[katteʔ] LH	[kattee] LHH	[kattee] LHH	[kattee] HLL
3 katee	[kateeʔ] LHH	[kattee] LHH	[kattee] LHH	[kattee] LHH
4 rika	[ɾikaʔ] HL	[ɾikka] HL	[ɾiikaʔ] HHL	[ʁikkaa] HLL
5 rikka	[ɾikkaʔ] HL	[ɾikkaa] HLL	[ɾiikaʔ] HHL	[ʁikkaa] HLL
6 rikaa	[ɾikaaʔ] HLL	[ɾikkaa] HLL	[ɾiikaa] HHLL	[ʁikkaa] HLL
7 iso	[isoʔ] LH	[issoo] LHH	[issoʔ] LH	[issoo] LHH
8 isso	[isoʔ] LHH	[issoo] LHH	[issoʔ] LH	[issoo] LHH
9 isoo	[isooʔ] LHH	[issoo] LHH	[issooʔ] LHH	[issoo] LHH
10 koma	[komaʔ] HL	[kommaʔ] HHL	[kommaʔ] HHL	[komaa] HLL
11 komma	[kommaʔ] HLL	[kommaʔ] HLL	[kommaʔ] HHL	[kommaa] HHLL
12 hone	[honeʔ] LH	[honee] LHH	[honnee] LLHH	[honee] LHH
13 honne	[honneʔ] LHH	[honee] LHH	[honneʔ] LLH	[honnee] LLHH
14 saga	[saŋaʔ] HL	[saŋŋaa] HHLL	[saŋŋa] HHL	[saŋaa] HLL
15 sanga	[saŋŋaʔ] HLL	[saŋŋaʔ] HLL	[saŋŋaʔ] HHL	[saŋŋaa] HHLL

Appendix 6.12 Average values- DFAT3

AP

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	49.2	85.4	130.2	43.1	207.8	515.7
	8.3	9.3	8.9	9.3	28.5	23.9
2 katte	47.4	80.8	368.6	17.6	210.1	724.5
	6.6	11.4	18.2	0.7	19.8	22.9
3 katee	56.1	68.2	157.5	32.9	379.3	694.0
	7.4	8.9	11.1	8.8	37.5	38.9
4 rika	10.1	65.3	113.8	20.4	154.5	364.1
	4.3	6.2	8.3	6.7	17.0	25.4
5 rikka	9.6	76.4	349.8	17.9	154.7	608.4
	1.9	4.4	19.7	2.6	17.0	34.2
6 rikaa	8.4	66.1	152.1	32.7	369.9	629.3
	1.5	14.9	27.0	2.5	46.5	29.7
7 iso	0.0	87.7	141.9	0.0	175.7	405.3
	0.0	10.4	6.1	0.0	6.1	9.6
8 isso	0.0	73.5	194.8	0.0	177.1	445.4
	0.0	18.1	19.9	0.0	16.6	29.5
9 isoo	0.0	66.8	171.7	0.0	432.1	670.6
	0.0	16.7	16.3	0.0	39.3	34.4
10 koma	59.1	110.6	90.7	0.0	191.9	452.2
	5.3	4.1	1.9	0.0	3.8	8.8
11 komma	49.3	147.8	213.6	0.0	182.0	592.7
	7.9	8.9	11.6	0.0	11.0	5.9
12 hone	50.7	119.1	116.0	0.0	222.9	508.8
	3.3	2.6	11.4	0.0	14.8	7.1
13 honne	79.0	136.6	264.8	0.0	267.4	747.9
	4.1	16.8	46.4	0.0	24.6	76.6
14 saNa	163.3	124.4	88.0	0.0	142.5	518.1
	13.8	8.3	2.2	0.0	13.1	27.8
15 saNNa	140.9	161.8	229.8	0.0	200.2	732.7
	20.2	13.4	22.7	0.0	5.5	31.9

DS

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	38.2	77.2	153.0	30.7	222.7	521.8
	7.0	11.5	8.5	5.2	9.0	18.8
2 katte	34.9	55.2	223.5	37.4	335.3	686.3

	20.8	26.3	76.1	8.2	11.8	27.8
3 katee	39.7	81.7	204.7	28.9	350.1	705.1
	10.5	13.9	4.5	4.5	16.7	14.8
4 rika	18.6	89.2	160.3	27.0	135.6	430.7
	2.0	19.2	37.6	3.3	12.4	32.5
5 rikka	19.2	86.4	280.1	18.8	223.8	628.3
	4.1	12.3	30.8	2.2	19.0	54.4
6 rikaa	24.9	79.3	211.1	30.7	289.6	635.7
	6.3	17.2	44.8	4.2	20.8	53.8
7 iso	0.0	71.4	190.4	0.0	223.2	485.0
	0.0	6.4	6.0	0.0	21.0	18.1
8 isso	0.0	51.8	200.3	0.0	300.6	552.7
	0.0	18.8	11.1	0.0	20.5	10.4
9 isoo	0.0	90.2	200.8	0.0	422.7	713.7
	0.0	8.2	6.2	0.0	21.0	20.4
10 koma	45.3	74.9	166.3	0.0	143.5	430.0
	9.6	10.0	1.3	0.0	20.7	33.8
11 komma	32.8	84.0	147.5	0.0	131.5	395.8
	8.0	9.0	11.3	0.0	20.3	11.4
12 hone	51.8	122.1	95.9	0.0	265.9	535.7
	20.1	14.7	6.4	0.0	18.1	28.6
13 honne	56.9	106.9	131.4	0.0	348.1	643.3
	20.5	9.9	12.2	0.0	30.9	44.1
14 saNa	153.9	141.5	218.7	0.0	229.3	743.4
	34.1	18.9	21.1	0.0	36.9	38.8
15 saNNa	124.8	86.0	216.1	0.0	165.9	592.8
	7.1	28.2	10.4	0.0	8.5	36.7

PK

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	66.9	95.9	249.1	36.7	333.9	782.7
	16.8	33.1	29.8	4.3	21.6	40.4
2 katte	87.0	69.5	325.5	37.3	274.5	793.9
	12.0	4.9	29.4	12.7	22.0	38.0
3 katee	70.9	100.8	310.6	45.0	359.1	886.4
	5.8	10.3	1.5	7.4	28.4	22.6
4 rika	25.8	194.4	107.2	50.1	113.5	491.0
	12.2	20.2	5.2	8.8	16.7	29.0

5 rikka	38.5 15.3	251.8 31.8	120.6 26.4	69.8 5.2	110.6 18.4	591.2 32.5
6 rikaa	24.5 0.2	247.8 20.5	104.7 29.4	73.5 13.9	294.8 25.1	745.3 52.7
7 iso	0.0 0.0	79.0 8.3	293.4 21.9	0.0 0.0	161.2 20.5	533.6 35.4
8 isso	0.0 0.0	213.0 32.5	265.6 22.8	0.0 0.0	184.4 3.2	663.0 10.2
9 isoo	0.0 0.0	86.0 13.6	375.2 28.5	0.0 0.0	339.6 20.7	800.8 62.2
10 koma	85.0 13.9	151.6 16.8	227.1 46.0	0.0 0.0	163.0 32.4	626.7 64.1
11 komma	83.3 1.8	95.6 23.2	248.5 4.6	0.0 0.0	120.9 6.3	548.3 20.5
12 hone	117.1 5.0	145.1 20.9	267.5 19.2	0.0 0.0	356.5 32.8	886.3 64.8
13 honne	100.4 9.4	94.3 13.2	289.8 34.7	0.0 0.0	138.2 30.1	622.7 41.5
14 saNa	206.4 34.4	126.0 15.6	181.1 48.5	0.0 0.0	120.7 18.5	634.3 13.4
15 saNNa	171.7 11.3	140.2 8.5	338.2 17.9	0.0 0.0	152.7 6.0	802.8 9.8

PV

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	49.6 6.3	50.0 2.0	183.3 8.8	61.2 8.2	350.3 30.6	694.3 44.8
2 katte	38.3 4.1	45.8 2.3	261.3 22.3	38.2 5.8	361.5 11.3	745.0 26.2
3 katee	52.2 2.3	50.9 1.0	223.3 2.7	45.6 6.9	374.9 10.8	746.9 10.0
4 rika	28.4 11.4	70.0 7.1	224.0 8.5	45.8 11.0	342.7 16.2	710.9 16.8
5 rikka	40.2 21.1	62.1 5.6	237.1 38.6	44.0 8.5	304.9 13.5	688.3 25.2
6 rikaa	34.5 13.0	63.5 6.9	236.7 28.9	46.8 9.2	365.8 11.3	747.3 26.6
7 iso	0.0	64.9	280.2	0.0	358.8	703.8

	0.0	12.6	17.0	0.0	52.7	49.7
8 isso	0.0	65.1	255.5	0.0	410.3	731.0
	0.0	11.4	26.8	0.0	10.8	23.7
9 isoo	0.0	96.4	259.1	0.0	364.1	719.5
	0.0	11.9	15.9	0.0	8.5	35.5
10 koma	70.6	105.3	165.2	0.0	289.2	630.3
	2.2	1.9	4.9	0.0	23.6	24.5
11 komma	62.3	67.9	198.9	0.0	204.8	533.9
	7.6	24.3	10.8	0.0	30.6	42.6
12 hone	96.0	119.8	172.1	0.0	376.6	764.6
	27.6	4.1	11.3	0.0	56.9	89.8
13 honne	82.4	69.1	192.8	0.0	405.7	750.0
	16.0	10.1	9.6	0.0	33.8	35.8
14 saNa	158.8	113.6	214.4	0.0	225.1	711.9
	18.4	12.9	18.2	0.0	29.4	42.3
15 saNNa	163.2	90.1	182.6	0.0	262.1	698.0
	20.6	5.4	28.0	0.0	15.8	52.7

Appendix 6.13 Word duration ratios- DFAT3

	AP	DS	PK	PV
1 kate	100	100	100	100
2 katte	140	132	101	107
3 katee	135	135	113	108
4 rika	100	100	100	100
5 rikka	167	146	120	97
6 rikaa	173	148	152	105
7 iso	100	100	100	100
8 isso	110	114	124	104
9 isoo	165	147	150	102
10 koma	100	100	100	100
11 komma	131	92	87	85
12 hone	100	100	100	100
13 honne	147	120	70	98
14 saga	100	100	100	100
15 sanga	141	80	127	98

Appendix 6.14 Long/short contrast ratios- DFAT3

Long/short contrast (C2)

	AP	DS	PK	PV
1 kate	35	68	77	70
2 katte	100	100	100	100
3 katee	43	92	95	85
4 rika	33	57	89	95
5 rikka	100	100	100	100
6 rikaa	43	75	87	100
7 iso	73	95	110	110
8 isso	100	100	100	100
9 isoo	88	100	141	101
10 koma	42	113	91	83
11 komma	100	100	100	100
12 hone	44	73	92	89
13 honne	100	100	100	100
14 saga	38	101	54	117
15 sanga	100	100	100	100

Long/short contrast (V2)

	AP	DS	PK	PV
1 kate	55	64	93	93
2 katte	55	96	76	96
3 katee	100	100	100	100
4 rika	42	47	39	94
5 rikka	42	77	38	83
6 rikaa	100	100	100	100
7 iso	41	53	47	99
8 isso	41	71	54	113
9 isoo	100	100	100	100

Appendix 6.15 C2/V1 and V2/V1 ratios- DFAT3

C2/V1 ratios

	AP	DS	PK	PV
1 kate	1.52	1.98	2.60	3.67
2 katte	4.56	4.05	4.68	5.70
3 katee	2.31	2.51	3.08	4.39
4 rika	1.74	1.80	0.55	3.20

5 rikka	4.58	3.24	0.48	3.82
6 rikaa	2.30	2.66	0.42	3.73
7 iso	1.62	2.66	3.71	4.32
8 isso	2.65	3.87	1.25	3.92
9 isoo	2.57	2.23	4.36	2.69
10 koma	0.82	2.22	1.50	1.57
11 komma	1.44	1.76	2.60	2.93
12 hone	0.97	0.78	1.84	1.44
13 honne	1.94	1.23	3.07	2.79
14 saga	0.71	1.55	1.44	1.89
15 sanga	1.42	2.51	2.41	2.03

V2/V1 ratios

	AP	DS	PK	PV
1 kate	2.43	2.88	3.48	7.01
2 katte	2.60	6.07	3.95	7.89
3 katee	5.56	4.29	3.56	7.37
4 rika	2.37	1.52	0.58	4.90
5 rikka	2.02	2.59	0.44	4.91
6 rikaa	5.60	3.65	1.19	5.76
7 iso	2.00	3.13	2.04	5.53
8 isso	2.41	5.80	0.87	6.30
9 isoo	6.47	4.69	3.95	3.78
10 koma	1.74	1.92	1.08	2.75
11 komma	1.23	1.57	1.26	3.02
12 hone	1.87	2.18	2.46	3.14
13 honne	1.96	3.26	1.47	5.87
14 saga	1.15	1.62	0.96	1.98
15 sanga	1.24	1.93	1.09	2.91

Appendix 7.1 Question marks indicated by Beginners 2

1)	KA14TE_	D	220%	BS
2)	RI14KA_	D	160%	NP
3)	RI07K_A	D	220%	BS
4)	RI14K_A	A	140%	NB
5)	I07S_O	A	180%	BS
6)	I07S_O	A	140%	BS

7)	I07S_O	D	200%	BS
8)	HO14N_E	A	80%	NP
9)	KO07M_A	D	60%	NB
10)	KO14M_A	A	70%	LB
11)	SA07N_A	D	60%	NP
12)	SA14N_A	A	70%	LB

Appendix 7.2 Perceptual boundary- ANU1/ANU2 (dependent t-test)

STIMULI	T-test	P value
KA07TE_	.608	.5579
KA14TE_	1.524	.1619
RI07KA_	1.342	.2126
RI14KA_	1.105	.2977
I07SO_	1.893	.0909
I14SO_	-1.383	.1999
KA07T_E	2.775	.0216
KA14T_E	1.099	.3005
RI07K_A	1.614	.1409
RI14K_A	1.57	.1508
I07S_O	2.764	.022
I14S_O	.944	.3697
HO07N_E	.535	.6056
HO14N_E	.374	.7174
KO07M_A	-.879	.4024
KO14M_A	-1.388	.1984
SA07N_A	-.271	.7928
SA14N_A	-1.235	.2479

Appendix 7.3 Ascending and descending series- Beginners 2

STIMULI	T-test	P value
KA07TE_	-2.836	.0073
KA14TE_	-2.985	.0049
RI07KA_	-.578	.5668
RI14KA_	-1.807	.0786
I07SO_	-1.968	.0564
I14SO_	-1.6	.1178

KA07T_E	-2.783	.0083
KA14T_E	-1.914	.0632
RI07K_A	-2.944	.0055
RI14K_A	-3.154	.0031
I07S_O	-2.178	.0357
I14S_O	.474	.6382
HO07N_E	-.09	.9291
HO14N_E	-.213	.8322
KO07M_A	-3.708	.0007
KO14M_A	.544	.5895
SA07N_A	-.595	.5551
SA14N_A	1.453	.1544

Appendix 7.4 V1 duration- Beginners 2

STIMULI	F-test	P value
KA07/14TE_	.582	.448
RI07/14KA_	3.269	.0745
I07/14SO_	2.497	.1182
KA07/14T_E	1.855	.1773
RI07/14K_A	.051	.8218
I07/14S_O	1.384	.2431
HO07/14N_E	.795	.3756
KO07/14M_A	1.499	.2245
SA07/14N_A	1.725	.1931

Appendix 7.5 Auditory impression- Beginners 2

JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate?	katee	kate?	kate?	kate?	kate?	kate?	kate?	kate?	kate	kate?
2 katte?	katte?	katte?	katte?	katte?	katte?	katte?	katte?	katte?	kate?	katte?
3 katee?	katee	katee	katee	katee	katee?	katee	katee?	kattee	katee	katee?
4 rika?	rika?	rika?	rika	rika?	rika?	rika?	rikka?	rika?	rika?	rika?
5 rikka?	rikka?	rikka?	rikka	rikka?	rikka?	rikka?	rikka?	rikka?	rikka	rikaa?
6 rikaa	rikaa?	rikaa	rikaa	rikaa	rikaa	rikkaa?	rikkaa?	rikaa	rikaa	rikaa?
7 iso?	iso?	iso?	iso?	iso?	iso?	iso?	iso?	iso?	iso?	iso?
8 isso?	isso?	iso	isso?	isso?	issoo	isso?	isso?	isso?	isso?	isso?
9 isoo?	isoo	isoo	isoo	isoo	isoo	isoo	isoo	isoo?	isoo	isoo

Obstruents/Vowels-pitch assignment

JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 LH	LH	LH	HL	LH	LH	LH	LH	LH	LH	LH
2 LH	HL	LH	HLL	LH	LH	LH	LH	LH	LH	LH
3 LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH	HLL	LHH
4 HL	HL	HL	HL	HL	HL	HL	HL	HL	HL	HL
5 HL	HL	HL	HL	HL	HL	HL	HL	HL	HL	HL
6 HLL	HLL	HHL	HLL	HLL	HLL	HLL	HLL	HLL	HLL	HLL
7 LH	LH	LH	HL	LH	LH	LH	LH	LH	LH	LH
8 LH	LH	LH	LH	LH	LHH	LH	LH	LH	LH	LH
9 LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH	LHH

Nasals

JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ	komaʔ
11kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ	kommaʔ
12 honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ	honeʔ
13 honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ	honneʔ
14 sagaʔ	sagaʔ	sagaʔ	saga	saŋaʔ	sagaʔ	saŋaʔ	saŋaʔ	saŋaʔ	saŋaʔ	sagaʔ
15 saŋgaʔ	saŋŋaʔ	saŋgaʔ	saŋgaʔ	saŋŋaʔ	saŋŋaʔ	saŋŋaʔ	saŋŋaʔ	saŋŋaʔ	saŋŋaʔ	saŋŋaʔ

Nasals - pitch pattern

JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 HL	HL	HL	HL	HL	HL	HL	HL	HL	HL	HL
11 HLL	HHL	HHL	HLL	HLL	HLL	HLL	HLL	HLL	HLL	HLL
12 LH	LH	LH	LH	LH	LH	LH	LH	LH	LH	LH
13 LHH	LHH	LHH	HLL	LLH	LLH	LLH	LHH	LLH	LLH	LLH
14 HL	HL	HL	HL	HL	HL	HL	HL	HL	HL	HL
15 HLL	HLL	HLL	HLL	HLL	HHL	HLL	HLL	HLL	HLL	HL

The pitch patterns shown in bold type indicate inappropriate tone (e.g. HL -> LH)

Appendix 7.6 Average values- Beginners 2

NO	C1	V1	C2	VOT	V2	TOTAL
1 kate	34.01	79.88	124.82	22.71	151.48	412.90
	8.85	21.32	27.72	10.13	45.39	66.19
2 katte	33.26	79.75	236.65	20.14	138.02	507.81
	8.40	22.35	66.61	8.84	32.30	81.60
3 katee	37.53	75.74	144.81	28.30	339.94	626.33
	10.18	23.04	57.55	13.33	46.43	86.23

4 rika	16.56 4.21	69.17 25.44	134.36 39.46	25.54 8.46	130.88 46.61	376.51 61.52
5 rikka	24.02 18.35	65.13 26.39	248.87 52.51	23.08 6.54	130.11 50.90	491.21 78.16
6 rikaa	27.24 19.74	79.59 37.93	147.97 61.89	30.65 11.71	317.57 65.19	603.02 84.43
7 iso	0.00 0.00	73.20 15.37	141.79 30.60	0.00 0.00	135.98 47.53	350.97 70.17
8 isso	0.00 0.00	93.83 37.28	235.63 25.69	0.00 0.00	139.33 35.16	468.78 63.81
9 isoo	0.00 0.00	79.35 27.81	169.35 43.16	0.00 0.00	347.07 57.01	595.92 48.50
10 koma	48.44 13.12	76.05 15.45	95.89 33.70	0.00 0.00	107.49 35.58	327.86 56.47
11 komma	54.36 19.80	125.67 61.72	246.84 62.82	0.00 0.00	129.38 30.61	556.24 135.27
12 hone	60.58 14.77	76.31 16.05	100.41 33.90	0.00 0.00	132.87 33.40	370.17 53.40
13 honne	78.91 26.31	97.37 27.44	268.42 69.60	0.00 0.00	152.58 36.96	597.27 102.35
14 saNa	95.47 35.67	91.03 13.63	116.02 36.49	0.00 0.00	123.67 39.22	426.19 80.10
15 saNNa	95.46 30.26	112.89 22.22	210.10 61.05	0.00 0.00	130.10 34.01	548.56 95.74
14 saga	87.69 44.43	95.79 28.02	90.05 35.30	0.00 0.00	127.01 44.40	400.54 118.82
15 sanga	110.77 50.20	108.05 33.59	174.53 37.29	41.05 6.20	128.37 36.86	562.77 51.34

Appendix 7.7 Word duration ratios- Beginners 2

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	100	100	100	100	100	100	100	100	100	100	100
2 katte	142	114	147	115	110	158	114	118	125	121	116
3 katee	134	103	180	152	178	154	140	149	177	148	155
4 rika	100	100	100	100	100	100	100	100	100	100	100
5 rikka	151	105	150	111	167	152	113	114	139	121	147
6 rikaa	140	149	157	126	190	190	180	175	167	120	169

7 iso	100	100	100	100	100	100	100	100	100	100	100
8 isso	146	133	129	117	146	160	112	130	127	130	176
9 isoo	151	183	202	156	234	170	124	180	157	149	191

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 koma	100	100	100	100	100	100	100	100	100	100	100
11komma	140	254	276	172	142	150	137	165	150	121	173
12 hone	100	100	100	100	100	100	100	100	100	100	100
13 honne	150	203	226	145	145	179	166	138	128	129	176
14 saga	100	100	100	100	100	100	100	100	100	100	100
15 sanga	129	195	159	97	144	111	143	190	114	111	203

Appendix 7.8 Long/short contrast ratios- Beginners 2

C2

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	42	25	49	53	72	36	60	63	62	83	60
2 katte	100	100	100	100	100	100	100	100	100	100	100
3 katee	39	26	53	52	65	37	83	79	95	84	63
4 rika	39	69	39	76	24	30	67	81	62	65	52
5 rikka	100	100	100	100	100	100	100	100	100	100	100
6 rikaa	37	55	48	60	30	30	108	85	86	56	51
7 iso	48	53	63	75	47	49	66	71	77	59	44
8 isso	100	100	100	100	100	100	100	100	100	100	100
9 isoo	49	48	68	82	82	53	71	97	75	89	55

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 koma	39	17	19	43	43	33	35	52	48	74	46
11komma	100	100	100	100	100	100	100	100	100	100	100
12 hone	34	35	25	46	30	22	26	72	54	60	28
13 honne	100	100	100	100	100	100	100	100	100	100	100
14 saga	34	19	74	71	42	82	32	102	67	82	33
15 sanga	100	100	100	100	100	100	100	100	100	100	100

V2

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	47	96	34	42	30	36	49	45	36	46	44
2 katte	44	34	38	40	30	34	50	41	42	61	37

3 katee	100	100	100	100	100	100	100	100	100	100	100
4 rika	49	48	64	64	34	31	37	20	35	63	29
5 rikka	46	39	57	68	29	26	36	26	47	70	25
6 rikaa	100	100	100	100	100	100	100	100	100	100	100
7 iso	44	27	28	42	27	39	77	30	38	66	30
8 isso	42	28	27	38	25	42	58	33	52	64	49
9 isoo	100	100	100	100	100	100	100	100	100	100	100

Appendix 7.9 C2/V1 and V2/V1 ratios- Beginners 2

C2/V1 ratios

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	1.68	1.27	1.33	1.15	1.90	1.53	1.31	2.00	1.74	1.67	2.10
2 katte	3.21	3.77	2.01	2.38	3.56	3.73	2.60	3.38	3.07	1.85	4.39
3 katee	1.62	1.46	1.23	1.20	2.71	2.00	2.22	2.22	2.40	2.15	1.85
4 rika	1.31	2.08	1.14	1.41	1.65	0.83	2.81	3.44	2.01	2.61	3.81
5 rikka	2.82	2.98	3.13	2.20	6.86	2.65	5.58	5.33	3.03	3.94	8.05
6 rikaa	1.13	1.13	0.88	1.08	2.11	0.70	4.39	4.44	2.70	2.57	2.89
7 iso	1.79	1.79	1.65	2.71	2.34	1.22	1.99	2.97	1.51	1.96	1.88
8 isso	3.01	4.18	2.03	2.89	4.29	1.26	2.16	3.29	1.82	2.42	4.20
9 isoo	1.75	2.29	1.29	2.20	4.91	1.42	1.53	5.14	1.42	2.24	2.42

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 koma	0.97	0.70	0.89	1.79	1.30	0.61	0.95	1.33	2.00	2.38	1.28
11komma	2.13	2.08	1.42	2.49	2.52	1.13	2.44	2.01	3.49	2.65	1.57
12 hone	1.00	0.67	1.80	1.69	1.04	0.86	0.87	2.94	1.38	1.92	0.94
13 honne	2.28	1.28	2.78	2.52	4.55	2.85	2.62	3.16	2.72	2.85	3.60
14 saga	0.63	0.71	0.68	0.96	0.97	1.14	0.74	1.50	1.73	1.52	1.25
15 sanga	1.66	3.05	1.52	1.76	2.21	1.28	1.92	1.05	2.52	1.60	1.82

V2/V1 ratios

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
1 kate	1.82	4.28	1.50	1.89	1.81	1.61	1.37	1.58	1.40	2.31	2.31
2 katte	1.37	1.10	1.22	1.97	2.50	1.32	1.65	1.55	1.77	2.81	2.47
3 katee	3.99	4.95	3.70	4.80	9.60	5.59	3.40	3.09	3.48	6.33	4.39
4 rika	1.62	2.46	1.63	2.21	2.90	1.24	2.24	1.19	1.16	3.34	1.82
5 rikka	1.24	1.98	1.56	2.81	2.54	0.98	2.96	1.89	1.45	3.66	1.74

6 rikaa	2.93	3.44	1.59	3.38	8.74	3.39	5.95	7.26	3.18	6.13	4.90
7 iso	1.72	1.52	1.46	1.85	2.20	1.76	2.91	1.48	1.34	2.61	1.56
8 isso	1.33	1.89	1.10	1.31	1.77	0.98	1.57	1.28	1.68	1.83	2.48
9 isoo	3.75	7.89	3.77	3.25	9.91	4.91	2.71	6.21	3.38	2.96	5.35

WORD	JAP	BD	BM	BS	KL	KP	LB	LD	NB	NP	PF
10 koma	1.54	1.15	1.53	2.28	1.48	1.25	1.81	0.89	1.13	2.19	0.73
11komma	1.29	1.13	0.58	1.86	1.04	0.78	1.33	1.35	0.99	1.79	0.71
12 hone	1.71	1.11	2.87	2.53	1.43	1.66	1.83	1.61	1.52	2.28	1.17
13 honne	1.40	1.66	1.29	2.08	1.42	1.18	1.28	1.82	1.75	2.12	1.46
14 saga	1.13	1.70	1.36	1.60	1.48	1.02	1.32	0.90	1.20	1.89	0.94
15 sanga	0.94	1.27	1.17	1.21	1.58	1.13	1.11	0.82	1.25	1.69	0.66

Appendix 7.10 Misjudgment results

The number of words with misjudgment before the lesson (learners)

	BM1	BD1	LD1
Accuracy (%)	60.0	53.3	75.6
No. of errors	18	21	11
kate	4	4	0
rika	6	6	2
iso	4	6	3
koma	2	1	2
hone	2	3	1
saga	0	1	3

The details of misjudgment before the lesson (learners)

BM1

Types of error	Examples
Short -> Long (C2)	kate (3), koma (1), rika (1)
Long -> Short (C2)	rikka (2), isso (1), komma (1) honne (1)
Short -> Long (V2)	N/A
Long -> Short (V2)	katee (1), rikaa (1)
not sure	hone (1), isso (1)
others	isoo->isso (1), isso->isoo (1) rikaa->rikka (2)

BD1

Types of error	Examples
Short -> Long (C2)	kate (3), iso (3), koma (1) hone (2), sanga (1)
Long -> Short (C2)	isso (1), honne (1)
Short -> Long (V2)	rika (1)
Long -> Short (V2)	rikaa (2)
not sure	katte (1), rika (1), rikka (1) rikaa (1), isso (2)
others	N/A

LD1

Types of error	Examples
Short -> Long (C2)	rika (2), iso (1), koma (2), hone (1) saga (3)
Long -> Short (C2)	isso (2)
Short -> Long (V2)	N/A
Long -> Short (V2)	N/A
not sure	N/A
others	N/A

The details of misjudgment before the lesson (NS)

NS (BM1)

Types of error	Examples
Short -> Long (C2)	kate (8), rika (4), koma (4), hone (2)
Long -> Short (C2)	katte (4), rikka (7), isso (13) komma (7), honne (6)
Short -> Long (V2)	rika (1)
Long -> Short (V2)	katee (4), rikaa (10), isoo (3)
not sure	kate (9), katte (12), katee (10) rika (2), rikka (8), rikaa (7) iso (2), isso (2), isoo (5) koma (6), komma (11) hone (3), honne (10)
others	isso -> isoo (1)

NS (BD1)

Types of error	Examples
Short -> Long (C2)	kate (15), rika (17), iso (1), koma (2), hone (1)

Long -> Short (C2)	isso (5), honne (11), sanga (15)
Short -> Long (V2)	iso (11)
Long -> Short (V2)	N/A
not sure	kate (3), katte (4), katee (5) rika (1), iso (4), isso (5), koma (9) komma (4), hone (7), honne (7) saga (2), sanga (3)
others	rikaa -> rikka (18), isso -> isoo (8)

NS (LD1)

Types of error	Examples
Short -> Long (C2)	kate (16), rika (17), iso (16) koma (9), hone (11), saga (7)
Long -> Short (C2)	honne (1)
Short -> Long (V2)	N/A
Long -> Short (V2)	isoo (1)
not sure	kate (2), katte (1), katee (10) rika (1), rikaa (13) iso (1), isso (3), isoo (9), koma (6), hone (5), honne (6) saga (11), sanga (5)
others	katee -> katte (8), isoo -> iso (5)

The number of words with misjudgment after the lesson (learners)

	BM2	BD2	LD2
Accuracy (%)	86.3	80.4	98.0
kate (3X3)	1	6	0
rika (3X3)	4	3	0
iso (3X3)	2	1	1
koma (2X3)	0	0	0
hone (2X3)	0	0	0
saga (2X3)	0	0	0
not sure	1	4	0

BM2

Short->Long(C2)	rika (1)
Long->Short(C2)	katte (1), rikka (1), isso (2)
Short->Long(V2)	rika (1)
Long->Short(V2)	rikaa (1)
not sure	rikaa (1)

BD2

Short->Long(C2)	N/A
Long->Short(C2)	katte (1), katee (1), isso (1)
Short->Long(V2)	kate (2)
Long->Short(V2)	rikaa (1)
not sure	kate (1), katee (1), rika (1), rikka (1)

LD2

Short->Long(C2)	N/A
Long->Short(C2)	isso (1)
Short->Long(V2)	N/A
Long->Short(V2)	N/A
not sure	N/A

The details of misjudgment after the lesson (NS)

NS (BM2)

Types of error	Examples
Short -> Long (C2)	kate (3), rika (1), iso (3), hone (2)
Long -> Short (C2)	katte (2), isso (9)
Short -> Long (V2)	rika (2)
Long -> Short (V2)	katee (1), rikaa (11)
not sure	katte (2), katee (5), rika (5), rikaa (4) isso (3), isoo (2), koma (3) komma (4), hone (1), honne (1)
others	N/A

NS (BD2)

Types of error	Examples
Short -> Long (C2)	kate (1), rika (7), iso (1), koma (1)
Long -> Short (C2)	rikka (7), isso (3), komma (4)
Short -> Long (V2)	kate (11), rika (1)
Long -> Short (V2)	katee (2), rikaa (4)
not sure	kate (3), katee (3), komma (11) honne (2), saga (3)
others	katee -> katte (1)

NS (LD2)

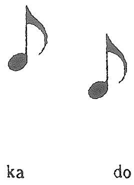
Types of error	Examples
Short -> Long (C2)	kate (4), rika (17), iso (4), koma (3) hone (6), saga (3)
Long -> Short (C2)	sanga (5)

Short -> Long (V2)	N/A
Long -> Short (V2)	N/A
not sure	kate (5), katte (1), katee (6), rikaa (5) isso (2), isoo (3), koma (1), hone (6) honne (3), saga (6), sanga (8)
others	katte -> katee (1), katee -> katte (1) isoo -> iso (1)

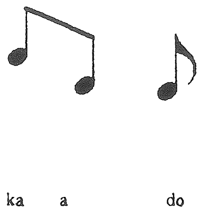
発音練習（長母音）

[a] 1.　かど（角）

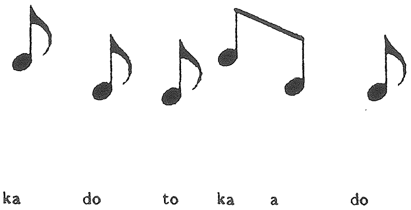
corner



2.　カード（カード）

card

3.　かど（角）と

カード（カード）

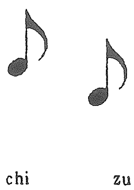


4.　かど（角）に名前を書きました。

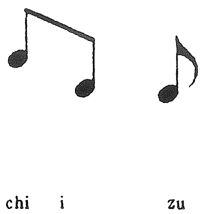
5.　カード（カード）に名前を書きました。

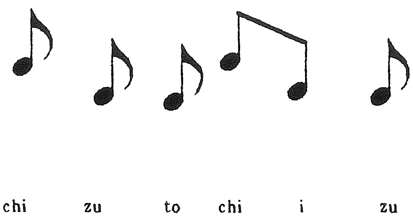
[i] 1.　ちず（地図）

map



2.　チーズ（チーズ）

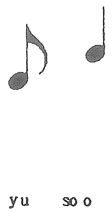
cheese

3.　ちず（地図）と

チーズ（チーズ）

4.　机の上にちず（地図）があります。

5.　机の上にチーズ（チーズ）があります。

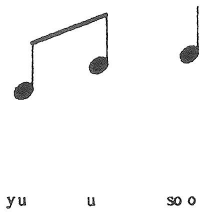
[u] 1.　ゆそう（輸送）

transportaion



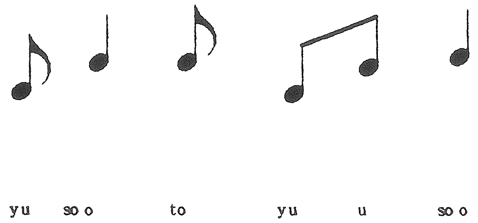
2.　ゆうそう（郵送）

postal



3.　ゆそう（輸送）と

ゆうそう（郵送）

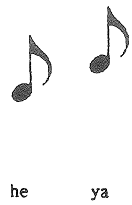


4.　小包をゆそう（輸送）します。

5.　小包をゆうそう（郵送）します。

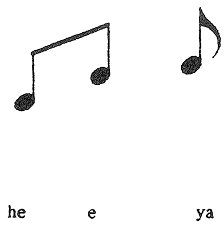
[e] 1.　へや（部屋）

room



2.　へいや（平野）

plain



3.　へや（部屋）と

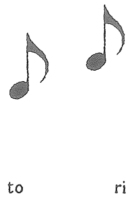
へいや（平野）

4.　広（ひろ）いへや（部屋）ですね。

5.　広いへいや（平野）ですね。

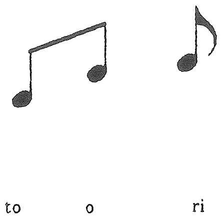
[o] 1.　とり（鳥）

bird



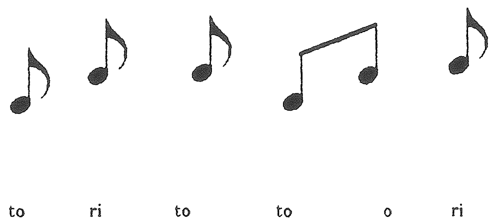
2.　とおり（通り）

avenue



3.　とり（鳥）と

とおり（通り）



4.　とり（鳥）を捕ります。

5.　とおり（通り）を通ります。

文章課題

1.　角の店でカードを買いました。
2.　恵子さんはスキーが好きです。
3.　地図の上にチーズが置いてあります。
4.　姪の目は大きくて可愛いです。
5.　一生一緒に暮らしましょう。
6.　四時に幼児を迎えに行きます。
7.　鳥が通りにたくさんいます。

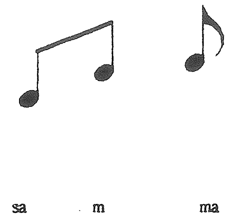
発音練習（長鼻音）

[m] 1.　さま（様）

Mr.Mrs.Ms.

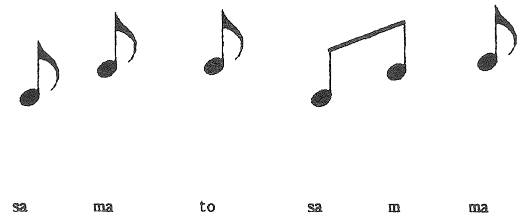
2.　さんま（秋刀魚）

mackerel pike



3.　さま（様）と

さんま（秋刀魚）



4.　さま（様）と言いました。

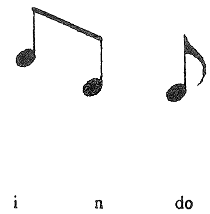
5.　さんま（秋刀魚）と言いました。

[n] 1.　いど（井戸）

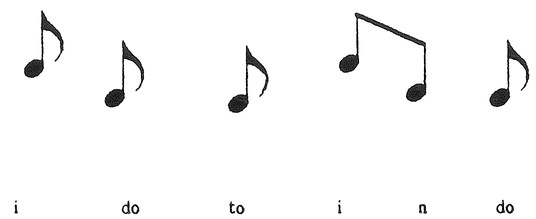
a well

2.　インド

India



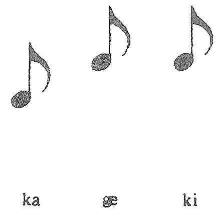
3.　いど（井戸）と

インド

4.　いど（井戸）まで行きます。

5.　インドまで行きます。

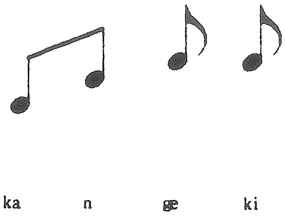
[ŋ] 1.　かげき（過激）

radical



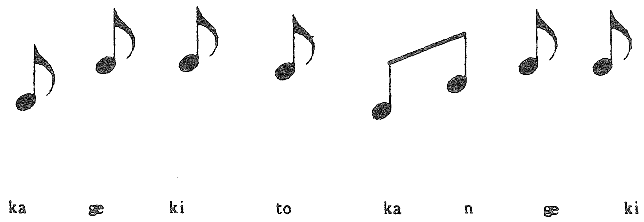
2.　かんげき（感激）

be moved



3.　かげき(過激)と

かんげき(感激)

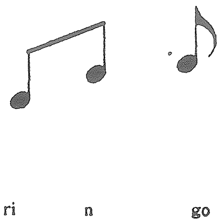


4.　兄(あに)はかげき（過激）でした。

5.　兄(あに)はかんげき（感激）しました。

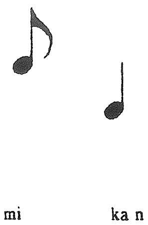
[g] 1.　りんご（林檎）

apple



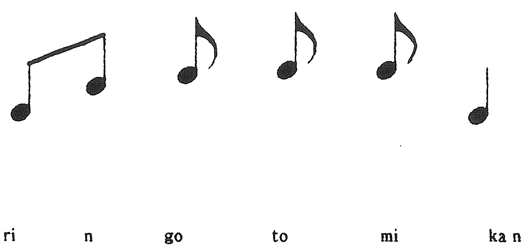
2.　みかん（蜜柑）

orange



3.　りんご（林檎）と

みかん（蜜柑）



4.　りんご（林檎）を食(た)べました。

5.　みかん（蜜柑）を食(た)べました。

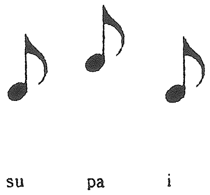
文章課題

1. 晩御飯(ばんごはん)のおかずが秋刀魚(さんま)ではさまにならないよ。
2. インドの緯度(いど)はどれくらいですか。
3. そのニュースには過激派(かげきは)も感激(かんげき)しました。
4. 美香(みか)さんは蜜柑(みかん)を食(た)べました。

発音練習（促音）

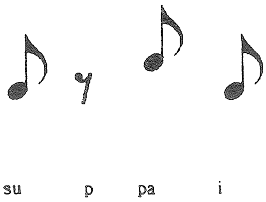
[p] 1. スパイ（スパイ）

spy



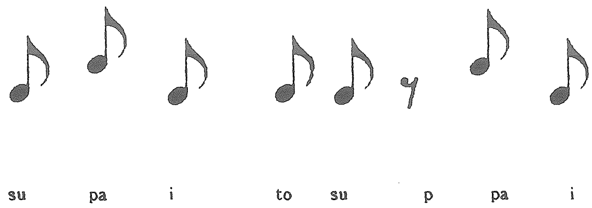
2. すっぱい（酸っぱい）

sour



3. スパイと

すっぱい（酸っぱい）



4. スパイだと思(おも)ったら、マフィアでした。
5. すっぱい（酸っぱい）と思(おも)ったら、レモンでした。

[t] 1. いて（居て）

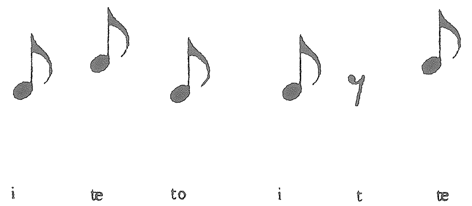
to stay

2. いって（行って）

to go

3.　いて（居て）と

いって（行って）

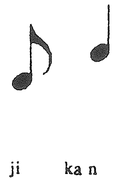


4.　居間にいて（居て）ください。

5.　居間にいって（行って）ください。

[k] 1.　じかん（時間）

time

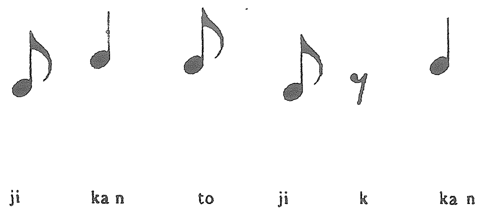


2.　じっかん（実感）

true feeling

3.　じかん（時間）と

じっかん（実感）

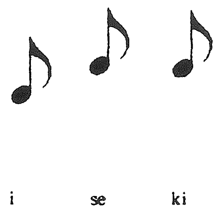


4.　あの時はじかん（時間）がありませんでした。

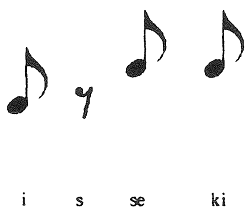
5.　あの時はじっかん（実感）がありませんでした。

[s] 1.　いせき（遺跡）

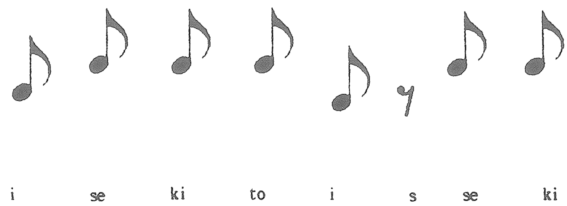
ruins



2.　いっせき（一隻）

one (ship)

3.　いせき(遺跡)と

いっせき(一隻)

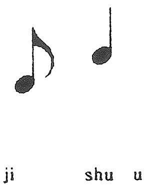


4.　船のいせき(遺跡)があります。

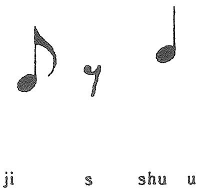
5.　船がいっせき(一隻)あります。

[ʃ] 1.　じしゅう（自習）

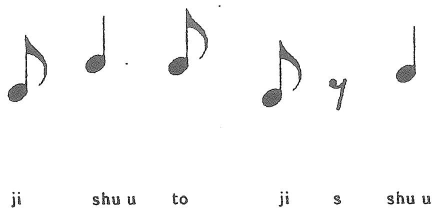
self study



2.　じっしゅう（実習）

practice

3.　じしゅう(自習)と

じっしゅう(実習)



4.　大学でじしゅう（自習）をしました。

5.　大学でじっしゅう（実習）をしました。

文章課題

1.　そのスパイは酸っぱいものが好きです。

2.　今ここに居て、後で居間に行ってください。

3.　そんなに時間が経ったという実感がありません。

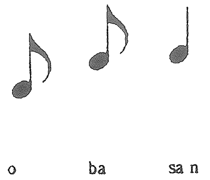
4.　大学の実習の前に自習をしました。

5.　一隻の船で遺跡を回りました。

6.　スーパーマンは都心に突進しました。

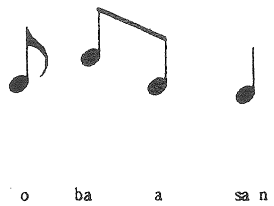
発音練習（まぎらわしいもの）

[a]1.　おばさん（叔母さん）



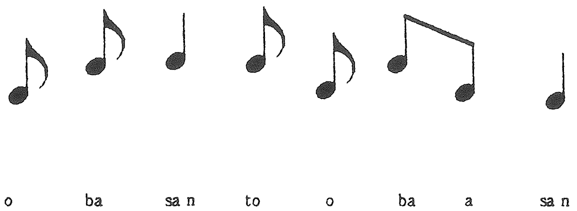
aunt

2.　おばあさん（お祖母さん）



grandmother

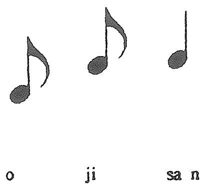
3.　おばさんとおばあさん



4.　おばさん（叔母さん）は元気です。

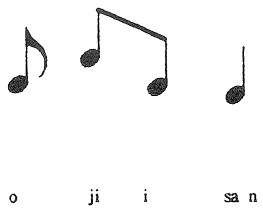
5.　おばあさん（お祖母さん）も元気です。

[i]1.　おじさん（伯父さん）



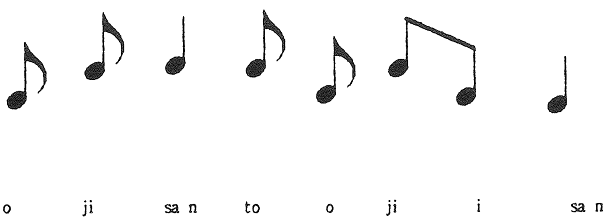
uncle

2.　おじいさん（お祖父さん）



grandfather

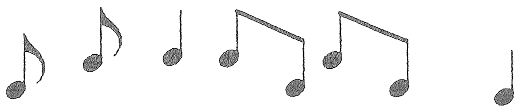
3.　おじさんとおじいさん



4.　おじさん（伯父さん）は元気です。

5.　おじいさん（お祖父さん）も元気です。

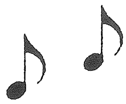
リズムの変形



o ji sa n *to o* ji i sa n

[n] 1. かに（蟹）

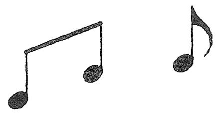
crab



ka ni

2. かんい（簡易）

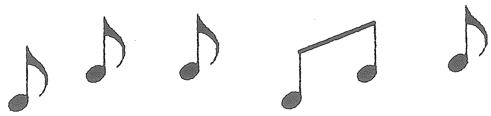
simplicity



ka n' i

3. かに（蟹）と

かんい（簡易）



ka ni to ka n' i

4. かに（蟹）と言(い)いました。

5. かんい（簡易）と言(い)いました。

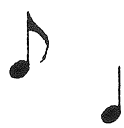
[n] 1. きねん（記念）

commemoration

ki ne n

2. きねん（祈念）

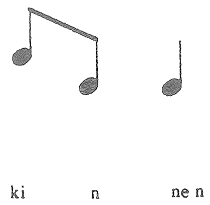
a prayer



ki ne n

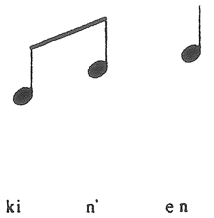
3.　きんねん（近年）

recent years



4.　きんえん（禁煙）

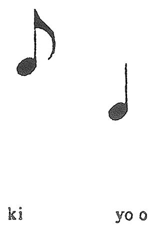
Non Smoking



発音練習（拗音）

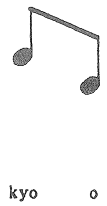
[y]1.　きよう（器用）

skillfulness



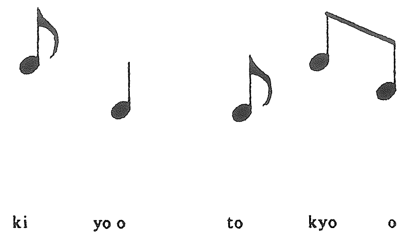
2.　きょう（今日）

today



3.　きよう（器用）と

きょう（今日）

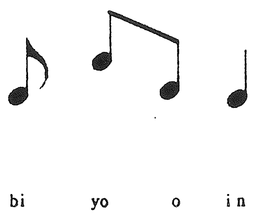


4.　きよう（器用）に作(つく)りました。

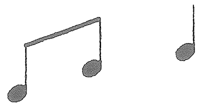
5.　きょう（今日）作(つく)りました。

[y]1.　びよういん（美容院）

beauty parlor



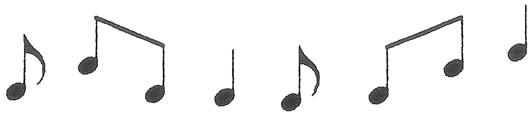
2. びょういん（病院）



hospital

byo o in

3. びよういん(美容院)と



びょういん(病院)

bi yo o in to byo o in

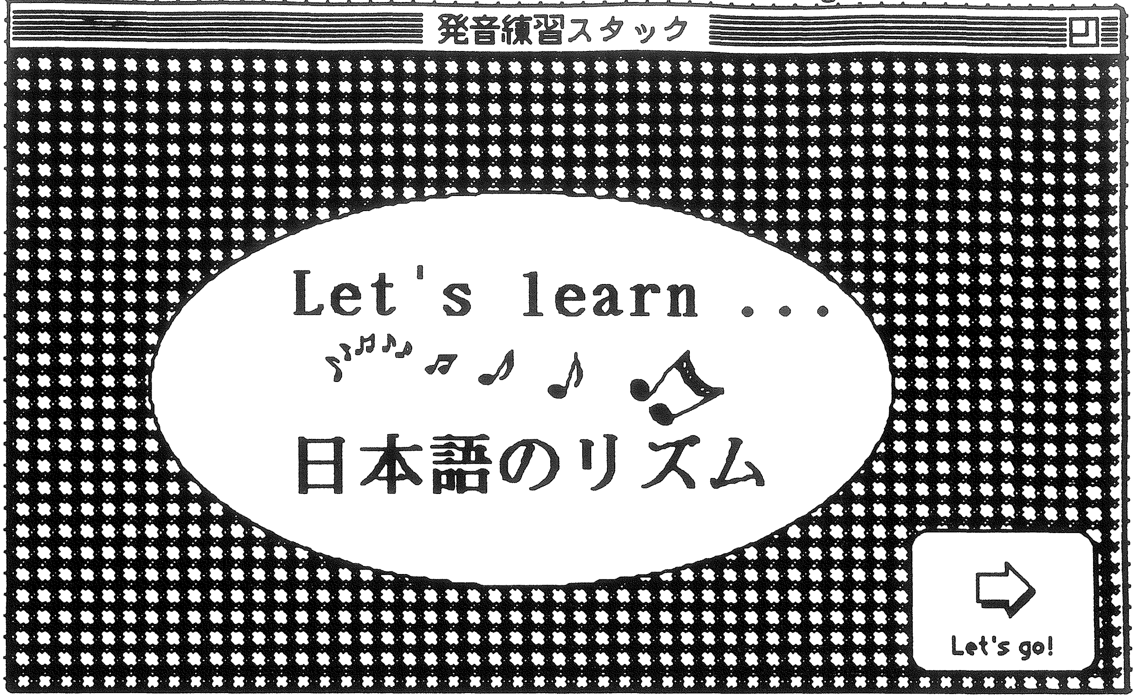
4. びよういん（美容院）まで行(い)きます。

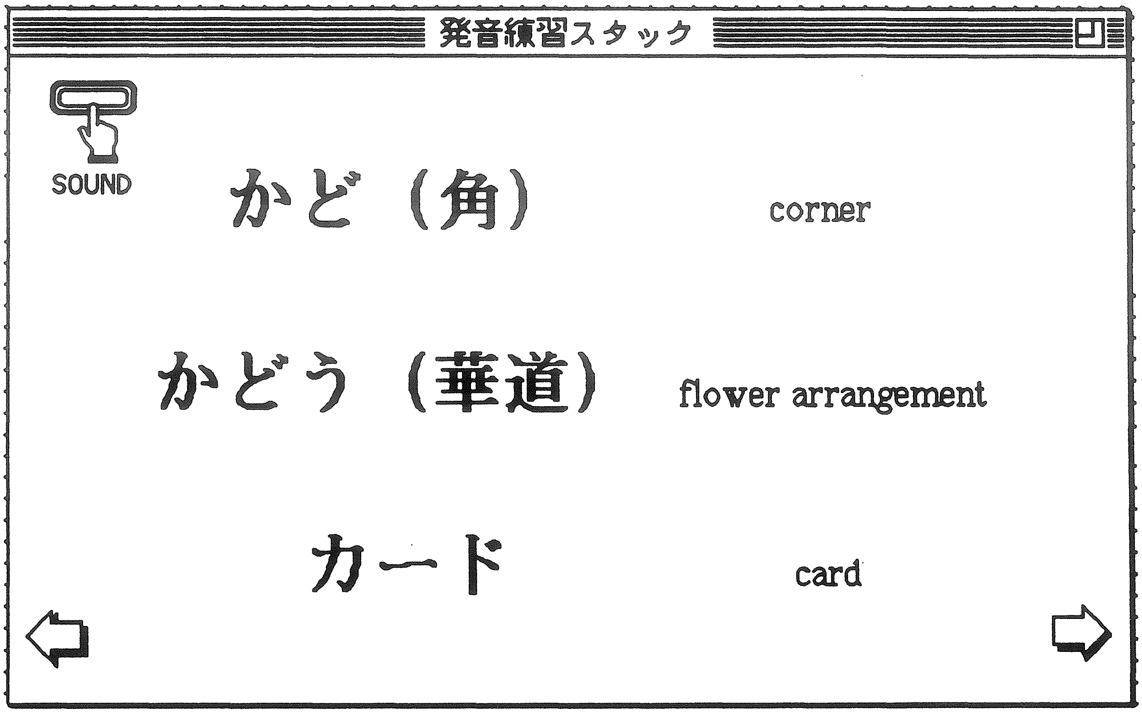
5. びょういん（病院）まで行(い)きます。

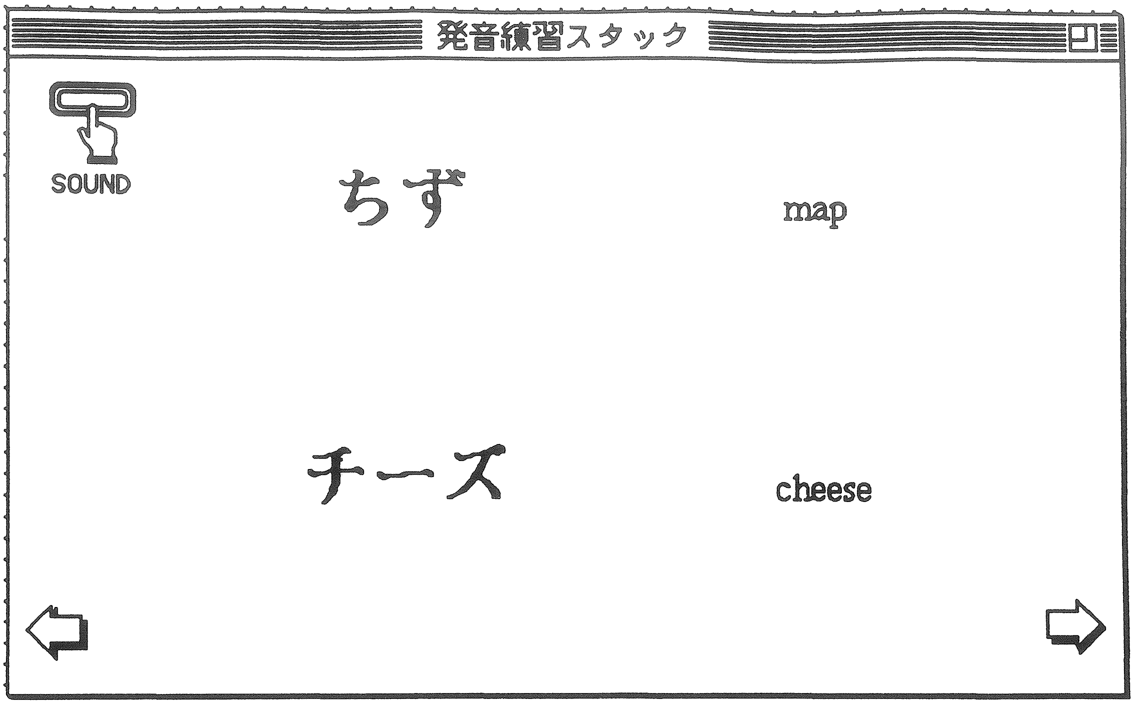
文章課題

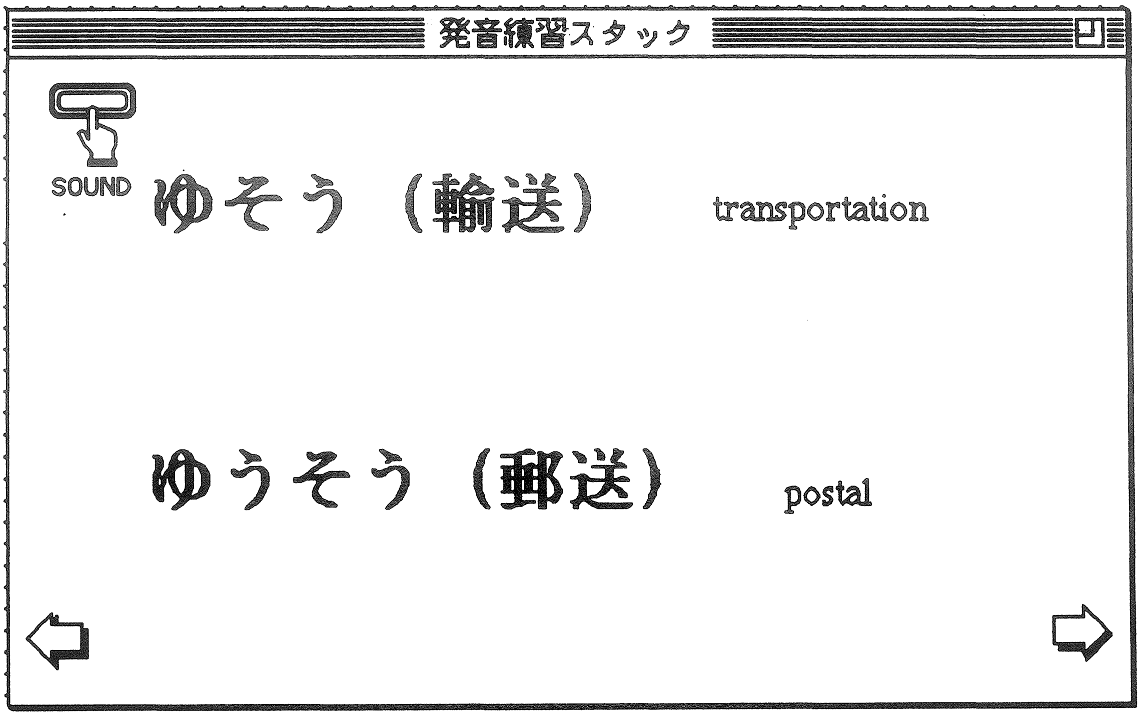
1. おばさんとおじさんはおばあさんとおじいさんの家(うち)へ行(い)きました。
2. 近年禁煙(きんねんきんえん)ブームです。
3. 今日子(きょうこ)さんは器用(きよう)な人(ひと)です。
4. 病院(びょういん)の後(あと)で美容院(びよういん)へ行(い)きます。

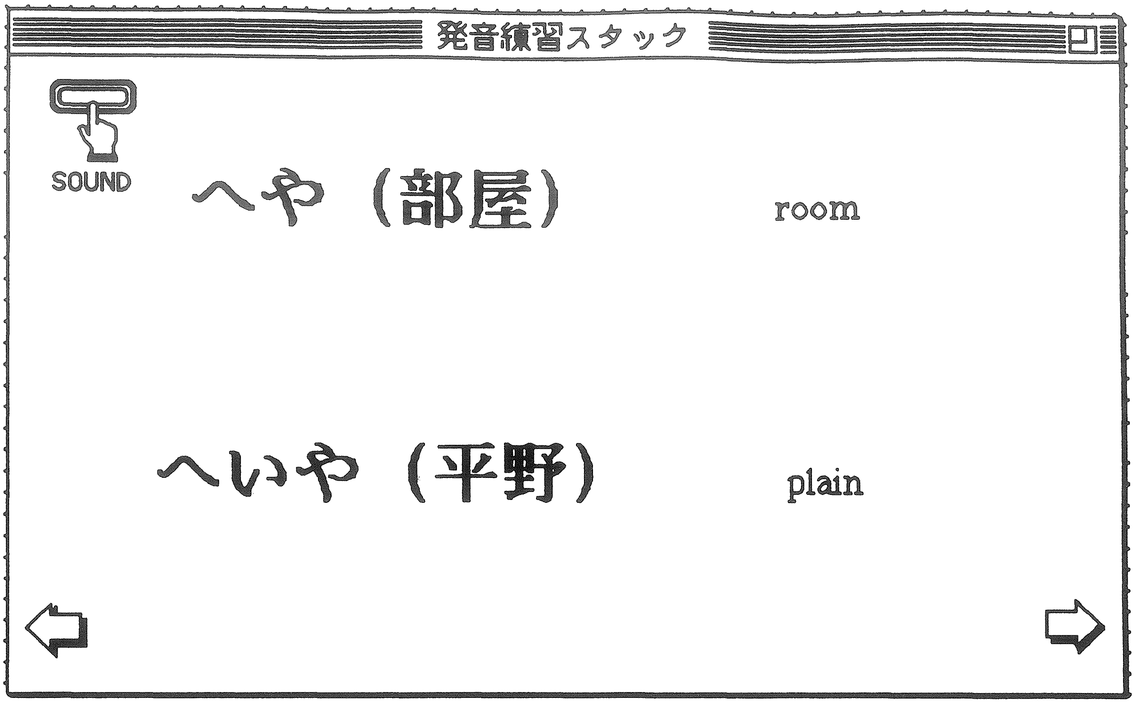
Appendix 8.2 Computer Associated Language Learning (CALL)



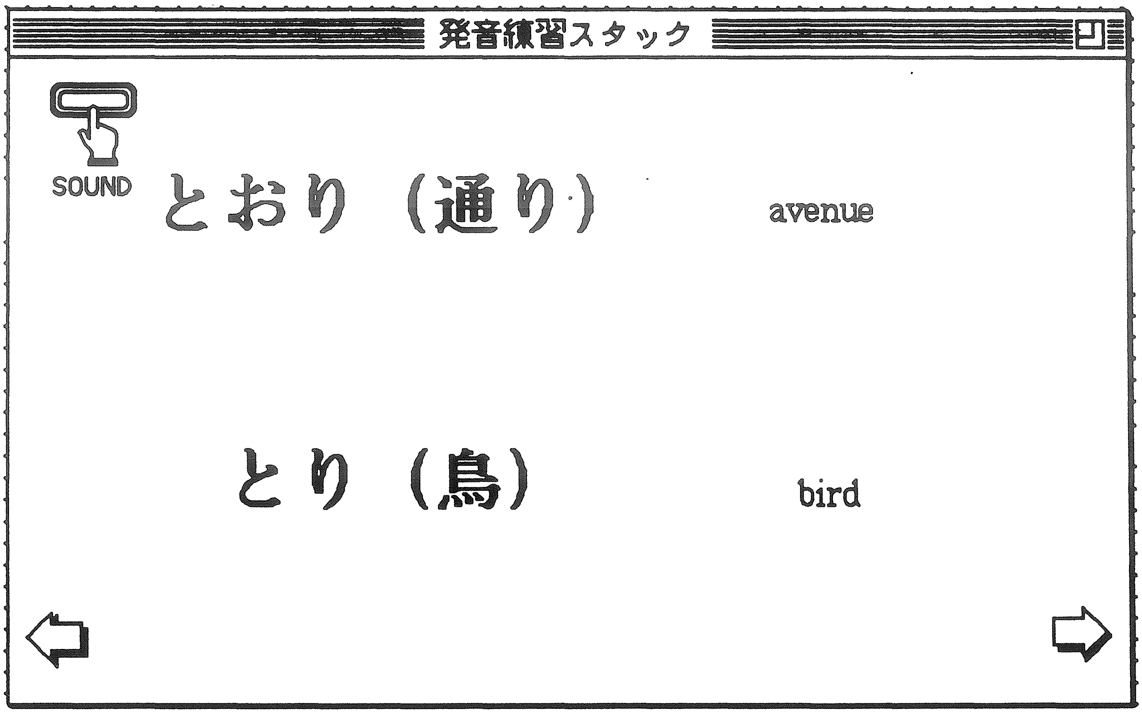



発音練習スタック
SOUND
ちず
map
チーズ
cheese


発音練習スタック
SOUND
ゆそう（輸送）
transportation
ゆうそう（郵送）
postal

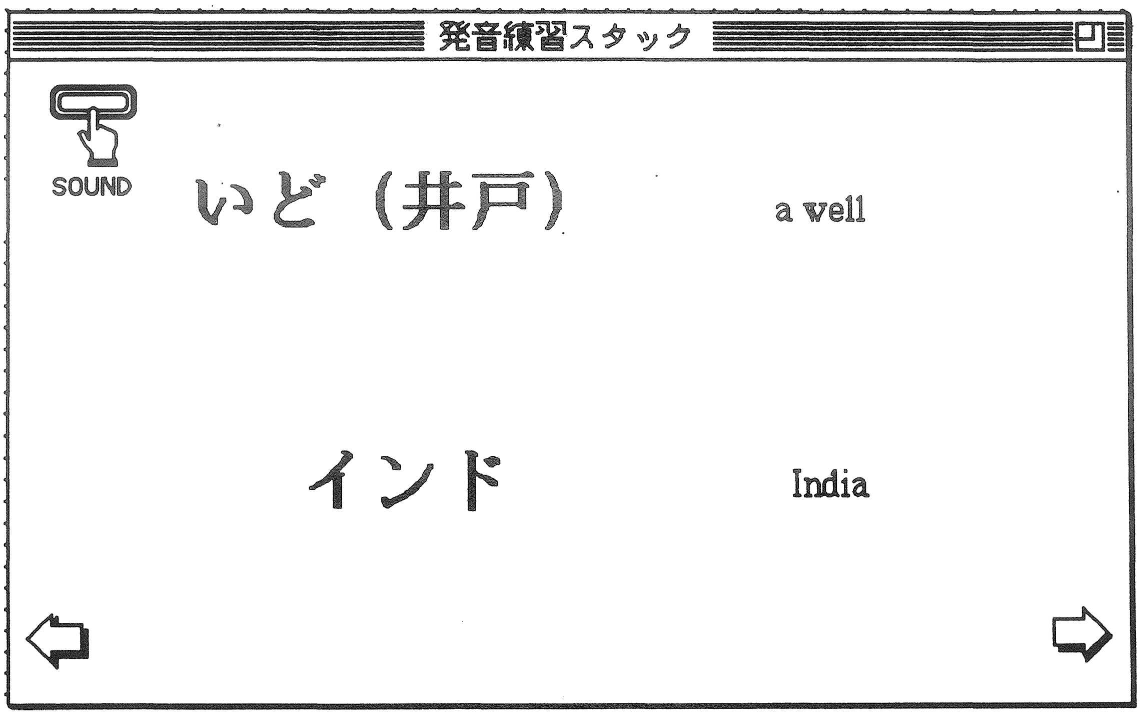


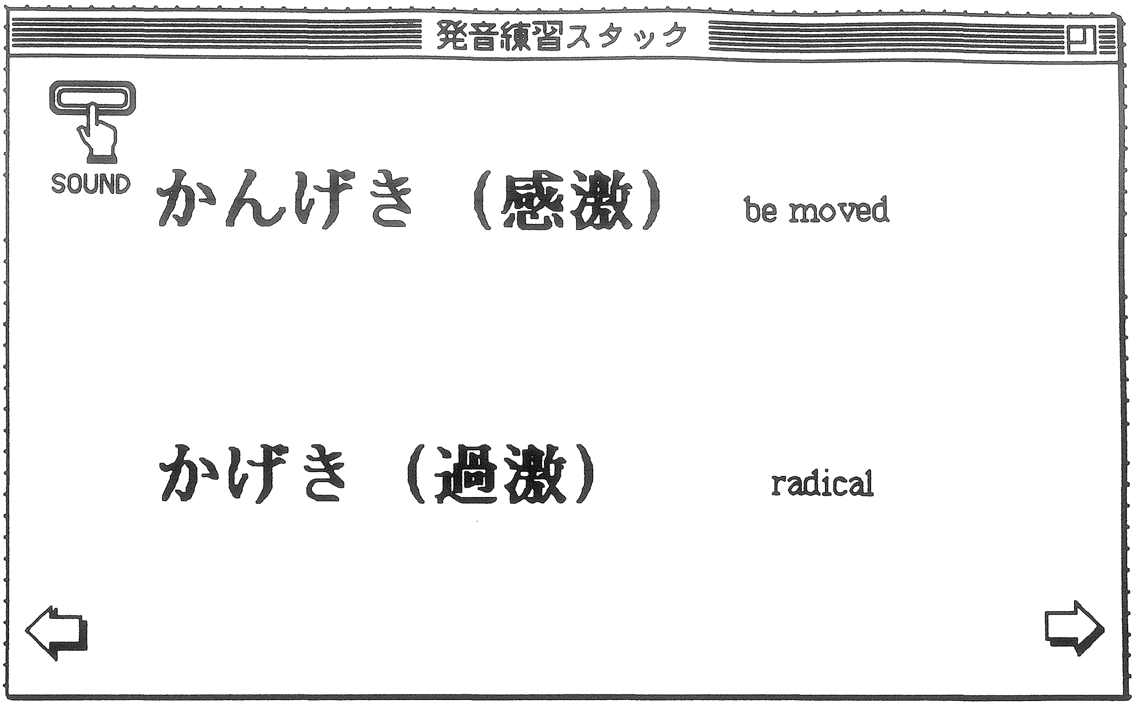
発音練習スタック
SOUND
へや（部屋）
room
へいや（平野）
plain

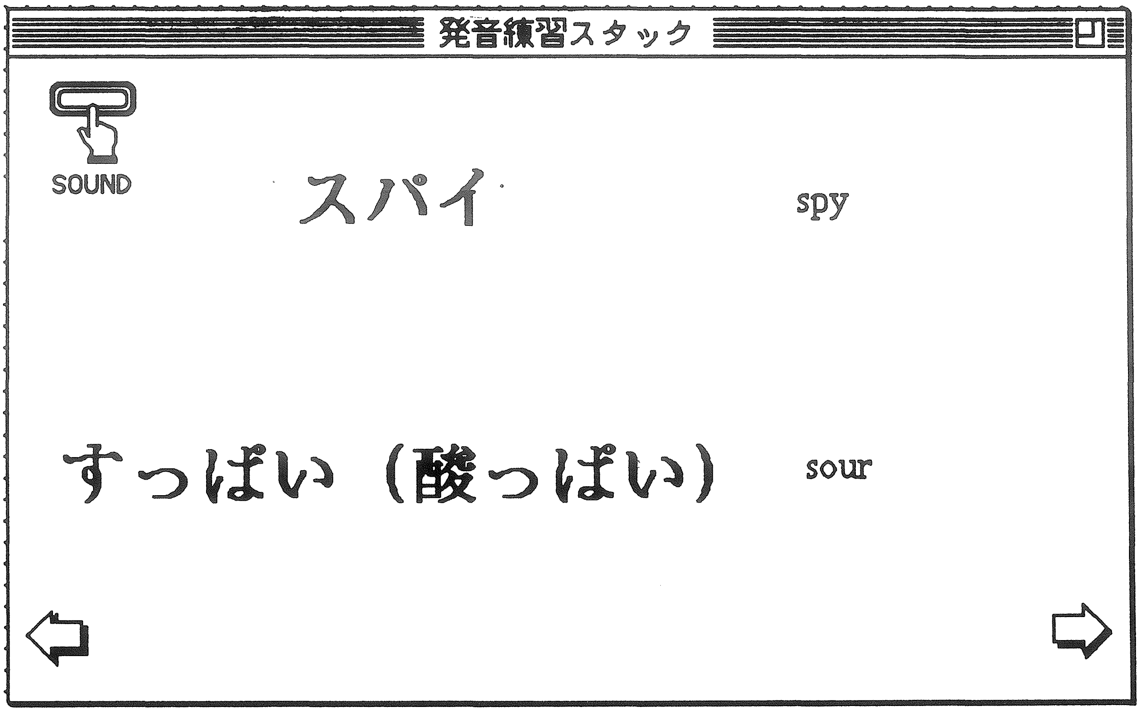


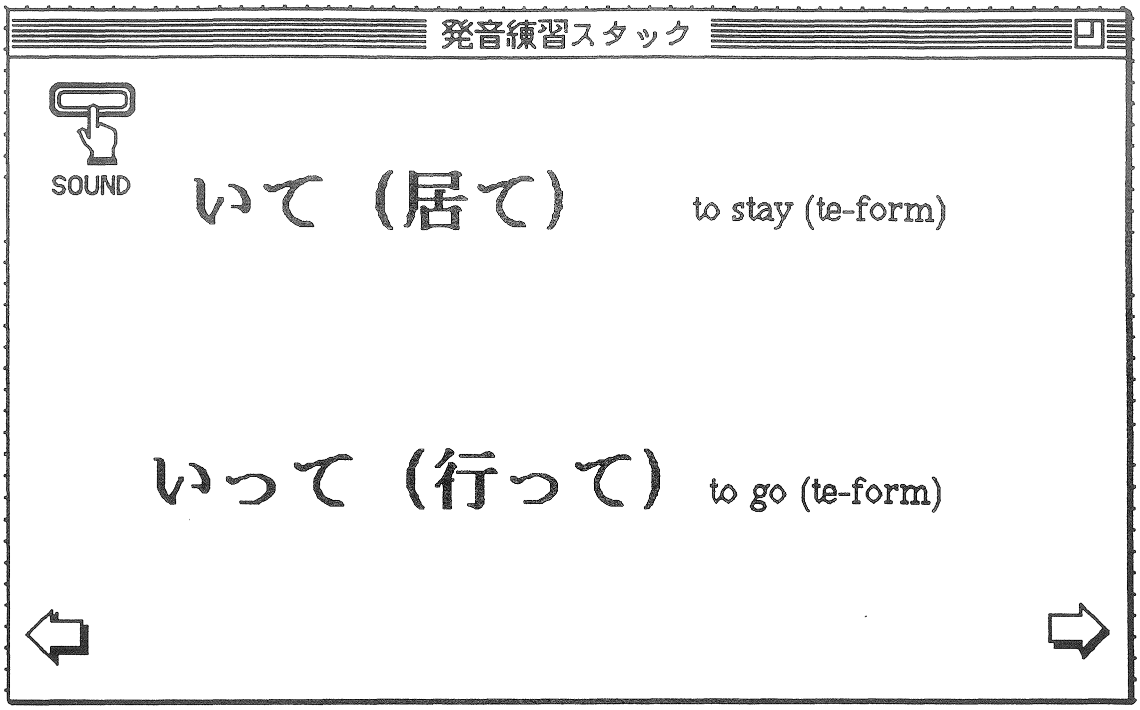
発音練習スタック
SOUND
とおり（通り）
avenue
とり（鳥）
bird

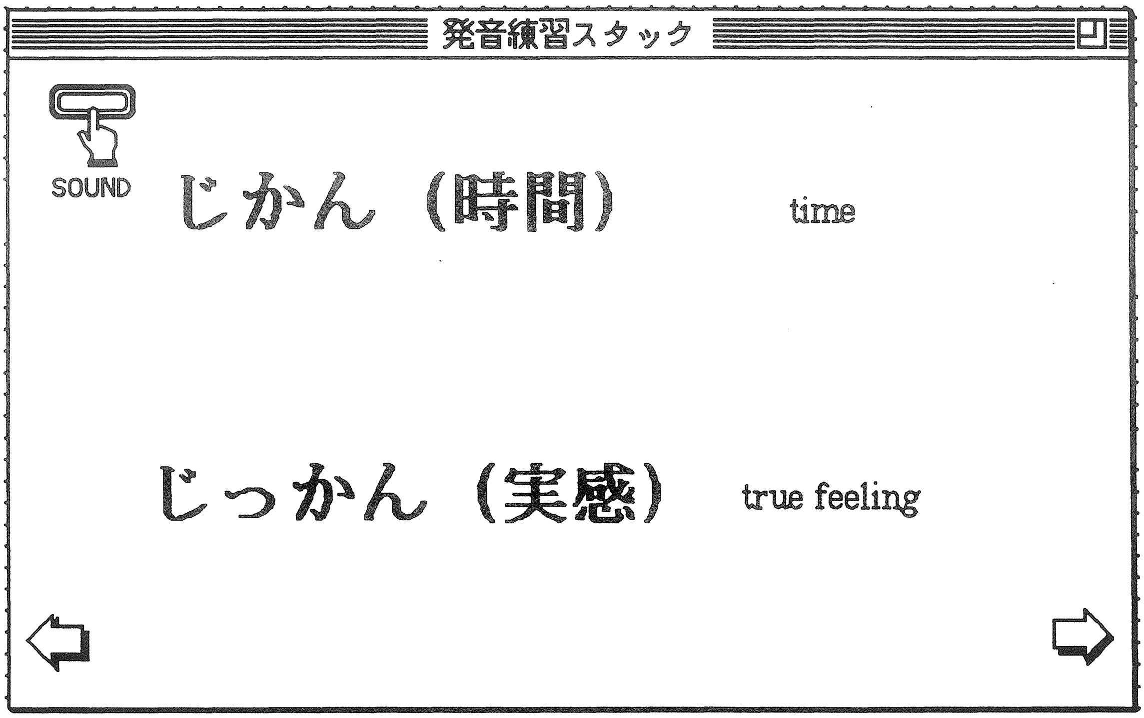

発音練習スタック
SOUND
さま（様）
Mr. Mrs. Ms.
さんま（秋刀魚）
mackerel

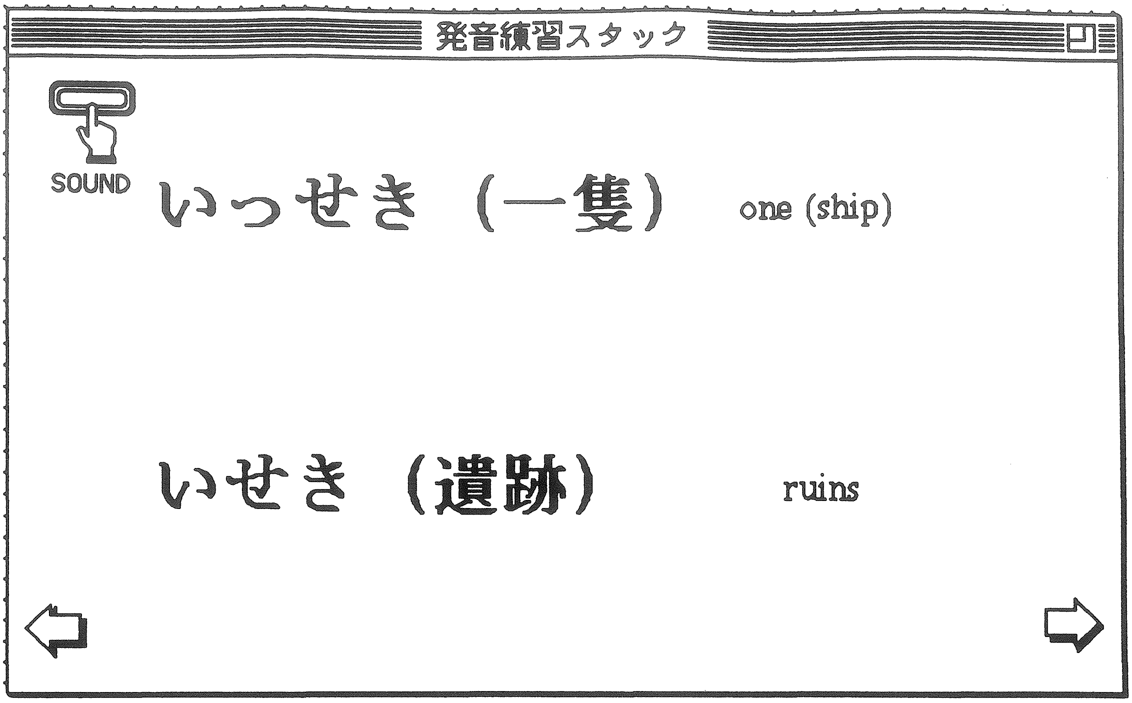

発音練習スタック
SOUND
いど（井戸）
a well
インド
India

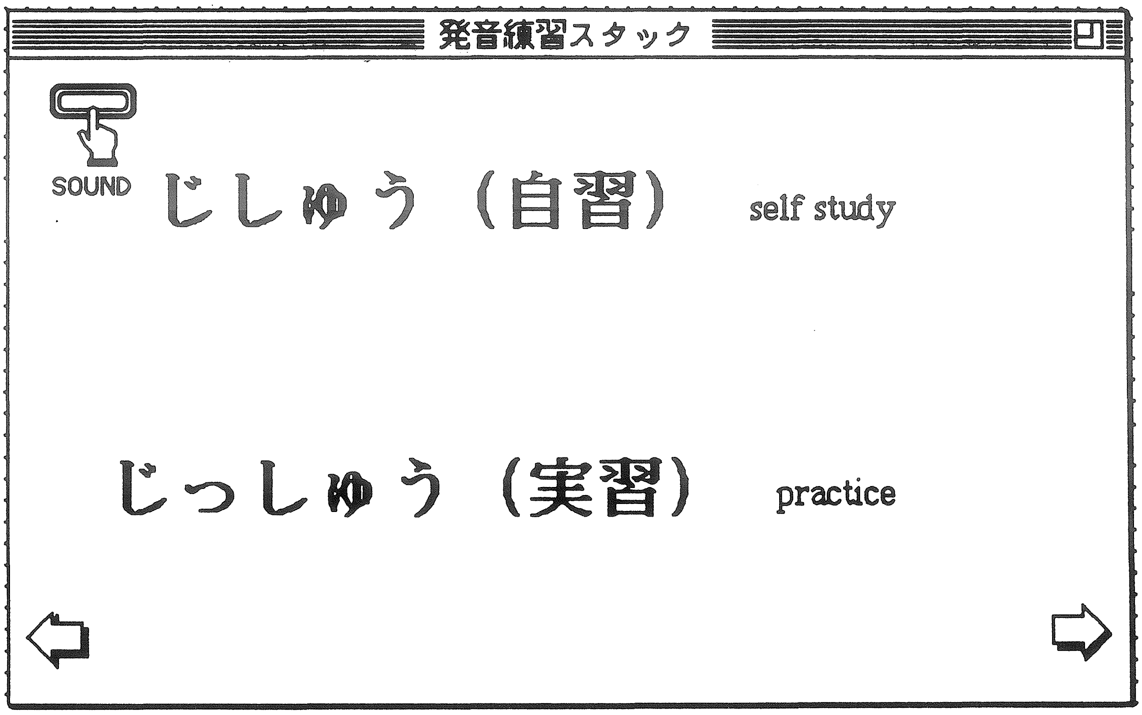

発音練習スタック
SOUND
かんげき（感激）
be moved
かげき（過激）
radical


発音練習スタック
SOUND
スパイ
spy
すっぱい（酸っぱい）
sour

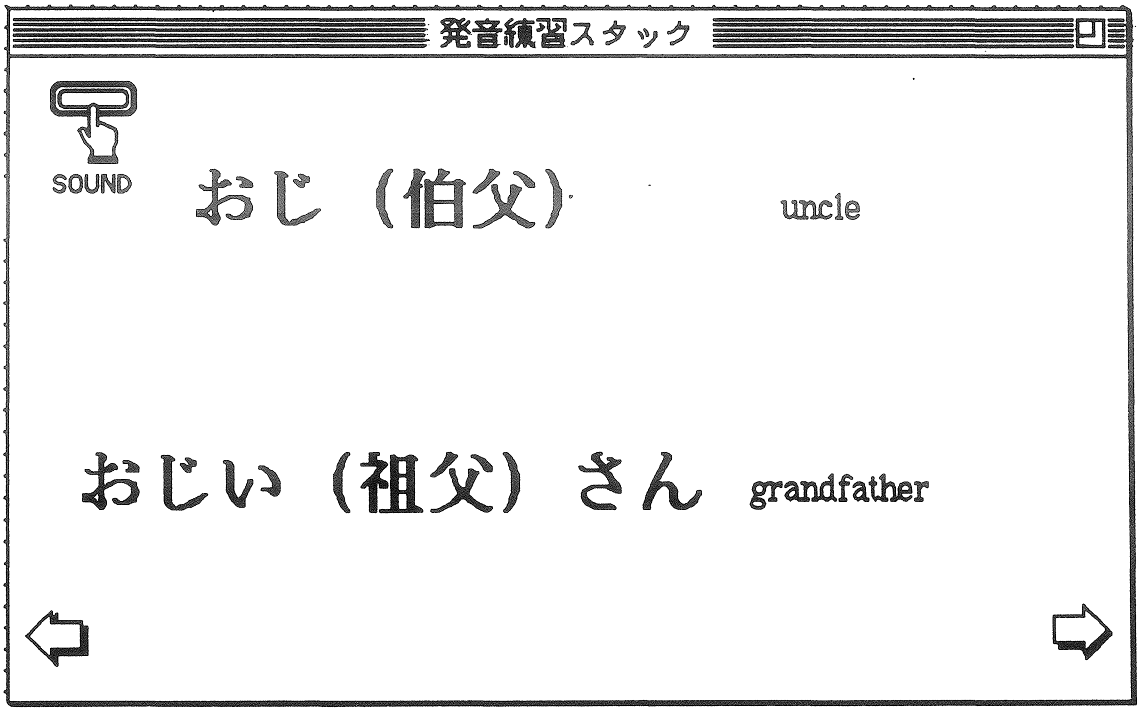

発音練習スタック
SOUND
いて（居て）
to stay (te-form)
いって（行って）
to go (te-form)

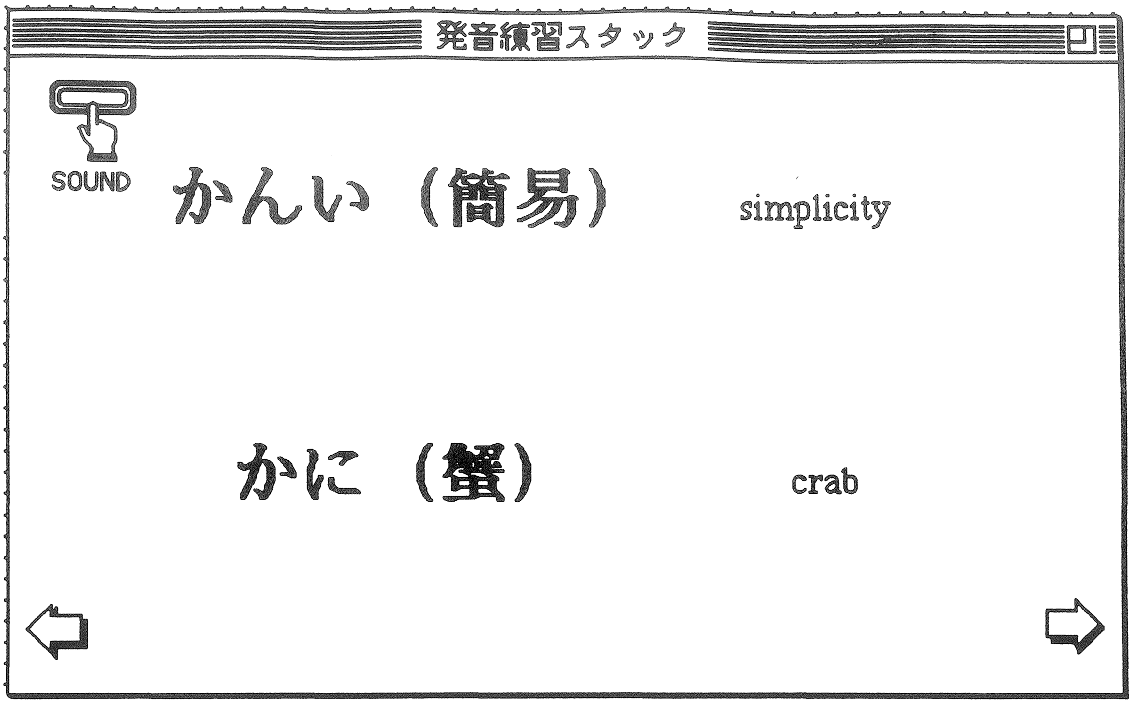

発音練習スタック
SOUND
じかん（時間）
time
じっかん（実感）
true feeling

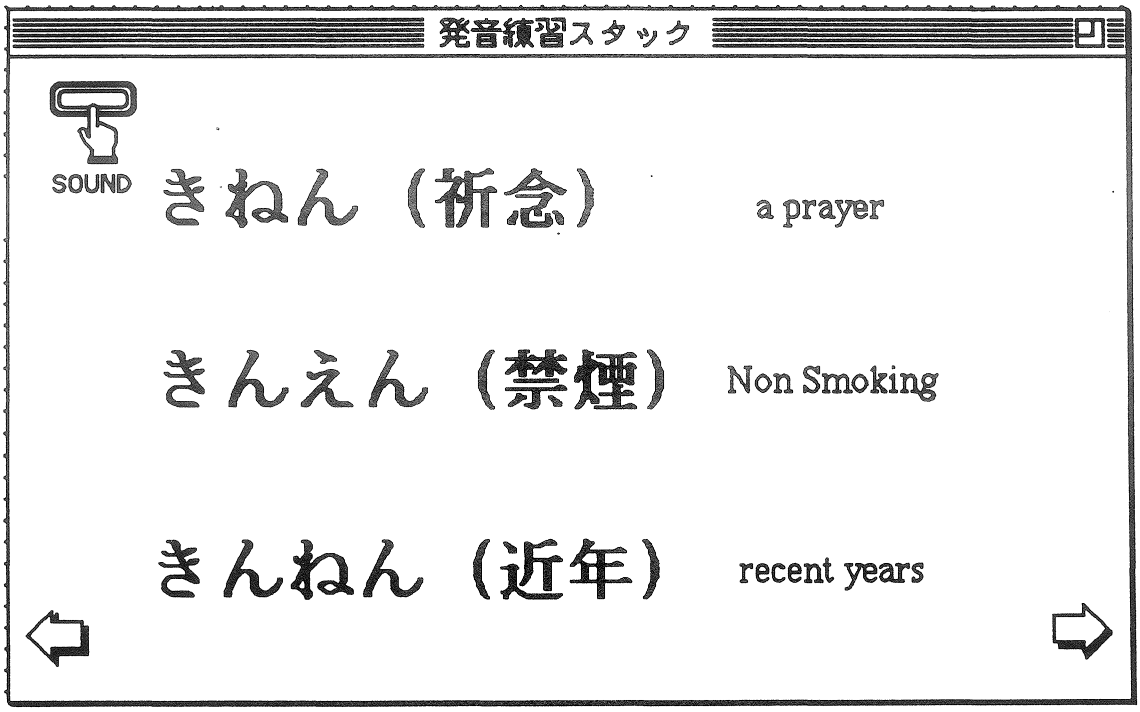

発音練習スタック
SOUND
いっせき（一隻）
one (ship)
いせき（遺跡）
ruins


発音練習スタック
SOUND
じしゅう（自習）
self study
じっしゅう（実習）
practice

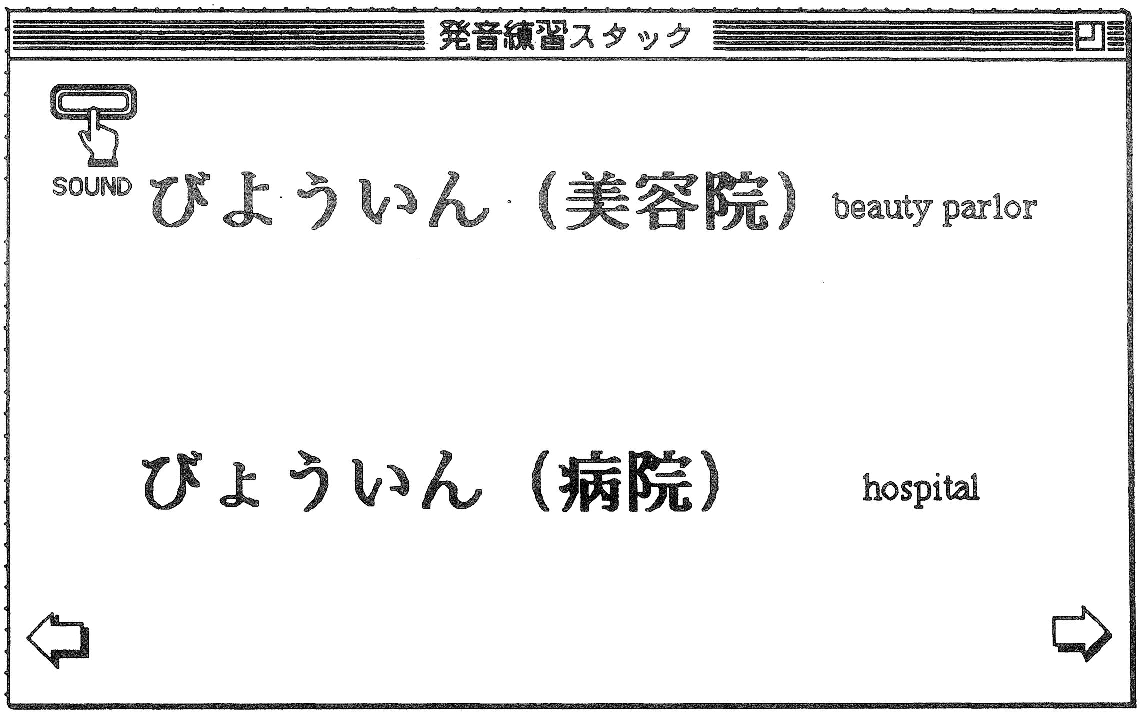

発音練習スタック
SOUND
おば（叔母）
aunt
おばあ（祖母）さん
grandmother


発音練習スタック
SOUND
おじ（伯父）
uncle
おじい（祖父）さん
grandfather


発音練習スタック
SOUND
かんい（簡易）
simplicity
かに（蟹）
crab


発音練習スタック
SOUND
きねん（祈念）
a prayer
きんえん（禁煙）
Non Smoking
きんねん（近年）
recent years


発音練習スタック
SOUND
きよう（器用）
skillfulness
きょう（今日）
today


発音練習スタック
SOUND
びようい ん（美容院）
beauty parlor
びょういん（病院）
hospital

Bibliography

Abe, Y. (1986) "Metrical structure and compounds in Japanese" In T. Imai & M. Saito eds., 5-52

Abercrombie, D. (1965) "Syllable quantity and enclitics in English" In *Studies in Phonetics and Linguistics*, 26-34, Oxford, Oxford University Press

Abercrombie, D. (1967) *Elements of General Phonetics*, Edinburgh, Edinburgh University Press

Abramson, A. & L. Lisker (1973) "Voice timing perception in Spanish word-initial stops" *Journal of Phonetics* 1, 1-8

Akinaga, K. (1968) "Iwayuru tokushu onsetsu (tokushu haku) ni tsuite" *Kooza Nihongo Kyooiku* 4, 36-51, Waseda University [in Japanese]

Al-Ani, S. (1970) *Arabic Phonology: An Acoustical and Physiological Investigation*, The Hague, Mouton

Arisaka, H. (1957) *Kokugo On'inshi no Kenkyuu*, Tokyo, Sanseido [in Japanese]

Asher, J. (1977) *Learning Another Language Through Actions: The Complete Teacher's Guidebook,* Los Gatos, California, Sky Oaks Productions

Assman, P.F., T.M. Nearey & J.T. Hogan (1982) "Voiced identification: orthographic, perceptual, and acoustic aspects" *Journal of the Acoustical Society of America* 71, 975-989

Ausubel, D.P., J.D. Novak & H. Hanesian (1978) *Educational Psychology: A Cognitive View*, Second Edition, New York, Holt, Rinehart and Wilston

Azuma, J. & Y. Tsukuma (1990) "Prosodic features marking the major syntactic boundary of Japanese" *Proceedings of the 1990 International Conference on Spoken Language Processing*, 453-455

Azuma, J. & Y. Tsukuma (1991) "Role of F0 and pause in disambiguating syntactically ambiguous Japanese sentences" *Proceedings of the Twelfth International Congress of Phonetic Sciences* 3, 274-277, University of Province

Barry, W. (1974) "Language background and the perception of foreign accent" *Journal of Phonetics* 2, 65-89

Bastian, J., P. Eimas & A.M. Liberman (1961) "Identification and discrimination of a phonemic contrast induced by silent interval" *Journal of the Acoustical Society of America* 33, 842 (Abstract)

Beebe, L. (1980) "Sociolinguistic variation and style shifting in second language acquisition" *Language Learning* 30, 433-447

Beebe, L. ed. (1988) *Issues in Second Language Acquisition: Multiple Perspectives*, Newbury House, Rowley, Mass.

Beckman, M.E. (1982a) "Effects of accent on vowel duration in Japanese" *Journal of the Acoustical Society of America* 71, S23

Beckman, M.E. (1982b) "Segment duration and the 'mora' in Japanese" *Phonetica* 39, 113-135

Beckman, M.E. (1984) "Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese" *Phonetica* 41, 61-71

Beckman, M.E. (1992) "Evidence for speech rhythms across languages" In Y. Tohkura et al. eds., 457-463

Beckman, M.E. (1994) "Foot and phrase in poetry" *Paper presented at the Twelfth National Conference of the English Linguistic Society of Japan*, 13, 91-94, November, University of Tokyo

Beckman, M.E. & J. Edwards (1989) "Lengthening and shortening and the nature of prosodic constituency" In J. Kingston & M.E. Beckman eds., 152-178

Bekku, S. (1973) "Nihongo Yonbyooshiron", *Asahi Shimbun* (Newspaper), Tokyo [in Japanese]

Bekku, S. (1977) *Nihongo no Rizumu*, Tokyo, Koodansha [in Japanese]

Bell-Berti, F. & K.S. Harris (1981) "A temporal model of speech production" *Phonetica* 38, 9-20

Best, C.T. (1992) "The emergence of language-specific phonemic influences in infant speech perception" In H.C. Nusbaum & J. Goodman eds., *The Transition from Speech Sounds to Spoken Words: The Development of Speech Perception*, MIT Press

Best, C.T., G.W. McRoberts & N.M. Sithole (1988) "Examination of perceptual reorganisation for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants" *Journal of Experimental Psychology: Human Perception and Performance* 14, 345-360

Best, C.T. & W. Strange (1992) "Effects of phonological and phonetic factors on cross-language perception of approximants" *Journal of Phonetics* 20, 305-330

Bloch, B. (1942) "Studies in colloquial Japanese IV: Phonemics" *Language* 26, 86-125

Bloomfield, L. (1933) *Language*, New York, Holt

Bornstein M. H. (1987) "Perceptual categories in vision and audition" In S. Harnad ed., *Categorical Perception: The Groundwork of Cognition*, 287-300, Cambridge, Cambridge University Press

Bradlow, A.R., R.F. Port & K. Tajima (1995) "The combined effects of prosodic variation on Japanese mora timing" *Proceedings of the Thirteenth International Congress of Phonetic Sciences*, Stockholm, Sweden

Browman, C. & L. Goldstein (1985) "Dynamic modelling of phonetic structure" In V. A. Fromkin ed., *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, 35-53, Orlando, Academic Press

Browman, C. & L. Goldstein (1990) "Representation and reality: Physical systems and phonological structure" *Journal of Phonetics* 18, 411-424

Campbell, W.N. (1992) "Speech timing in English and Japanese" *Proceedings of the International Symposium on Japanese Prosody*, 207-216, Nara, Japan

Campbell, W.N. & S.D. Isard (1991) "Segment durations in a syllable frame" *Journal of Phonetics* 19/1, 37-47

Campbell, W.N. & Y. Sagisaka (1991) "Mora-level effects in Japanese speech timing" *Paper presented at the Kinki Onsei Gengo Kenkyuukai*, April 1991

Celce-Murcia, M. (1977) "Phonological factors in vocabulary acquisition: A case study of a two-year-old English-French bilingual" *Working Papers in Bilingualism* 13, 27-95

Chen, M. (1970) "Vowel length variation as a function of the voicing of the consonantal environment" *Phonetica* 22, 129-159

Chomsky, N. & M. Halle (1968) *Sound Pattern of English*, New York, Harper & Row

Clark, M.M. (1986) "Japanese as a tone language" In T. Imai & M.Saito eds., 53-105

Cook, V. (1986) *Experimental Approaches to Second Language Learning*, Pergamon Press, Oxford

Corder, S.P. (1971) "Idiosyncratic and error analysis" *International Review of Applied Linguistics* 9, 147-160

Corder, S.P. (1978) "Language distance and the magnitude of the language learning task" *Studies in Second Language Acquisition* 2, 27-36

Crompton, A. (1981) "Syllable and segments in speech production" *Linguistics* 19, 663-716

Cutler, A., J. Mehler, D. Norris & J. Segui (1983) "A language-specific comprehension strategy" *Nature* 304, 159-160

Cutler, A., J. Mehler, D. Norris & J. Segui (1986) "The syllable's differing role in the segmentation of French and English" *Journal of Memory and Language* 25, 385-400

Cutler, A., J. Mehler, D. Norris & J. Segui (1992) "The monolingual nature of speech segmentation by bilinguals" *Cognitive Psychology* 24, 381-410

Cutler, A. & D. Young (1994) "Rhythmic structure of word blends in English" *Proceedings of the International Conference on Spoken Language Processing* 3, 1407-1410, Yokohama

Davis, S. & W. V. Summers (1989) "Vowel length and closure duration in word medial VC sequences" *Journal of Phonetics* 17, 339-353

Denes, P. (1955) "Effect of duration on the perception of voicing" *Journal of the Acoustical Society of America* 27, 761-764

Dickerson, L.J. (1974) *Internal and External Patterning of Phonological Variability in the Speech of Japanese Learners of English : Toward a Theory of Second-Language Acquisition* Unpublished Ph.D dissertation, University of Illinois

Dulay, H.C. & M.K. Burt (1973) "Should we teach children syntax?" *Language Learning* 23, 245-258

Dulay, H.C. & M.K. Burt (1974a) "Errors and strategies in child second language acquisition" *TESOL Quarterly* 8, 129-136

Dulay, H.C. & M.K. Burt (1974b) "A new perspective on the creative construction process in child second language acquisition" *Language Learning* 24, 253-278

Dulay, H., M. Burt & S. Krashen (1982) *Language Two*, New York, Harper

Eckman, F. (1977) "Markedness and the contrastive analysis hypothesis" *Language Learning* 27, 315-330

Eckman, F. (1981) "On the naturalness of interlanguage phonological rules" *Language Learning* 31, 195-216

Eckman, F., E. Moravcsik & J. Wirth eds. (1986) *Markedness*, New York, Plenum Press

Edwards, J. & M.E. Beckman (1988) "Articulatory timing and the prosodic interpretation of syllable duration" *Phonetica* 45, 2-4, 156-174

Eimas, P.D. (1975) "Auditory and phonetic coding of the cues for speech: Discrimination of the (r-l) distinction by young infants" *Perception and Psychophysics* 18, 341-347

Ellis, R. (1990) *Instructed Second Language Acquisition- Learning in the Classroom*, Basil Blackwall, Oxford

Ellis, S.M. & G. Weismer (1978) "More on temporal compensation" *Journal of the Acoustical Society of America* 63, S54

Enomoto, K. (1992) "Interlanguage phonology: the perceptual development of durational contrasts by English-speaking learners of Japanese" *Edinburgh Working Papers in Applied Linguistics* 3, 25-36

Felix, S. ed. (1980) *Second Language Development: Trends and Issues*, Gunter Narr Verlag

Felix, S. (1980) "Interference, interlanguage, and related issues" In S. Felix ed., 93-107

Fellbaum, M. (1986) "Markedness and allophonic rules" In F. Eckman et al. eds., 291-308

Flege, J.E. (1979a) "Temporal correlates of [voice] in Arabic-accented English" In Wolf & Klatt eds., *Speech Communication Papers*, 97th meeting of the Acoustical Society of America, 171-174

Flege, J.E. (1979b) *Phonetic Interference in Second Language Acquisition*, Unpublished Ph.D dissertation, Indiana University, Bloomington

Flege, J.E. (1980) "Phonetic approximation in second language acquisition" *Language Learning* 30, 117-134

Flege, J.E. (1981) "The phonological basis of foreign accent" *TESOL Quarterly* 15, 443-455

Flege, J.E. (1984) "The detection of French accent by American listeners" *Journal of the Acoustical Society of America* 76, 692-707

Flege, J.E. (1987a) "The instrumental study of L2 speech production: Some methodological considerations" *Language Learning* 37, 285-296

Flege, J.E. (1987b) "The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification" *Journal of Phonetics* 15, 47-65

Flege, J.E. (1987c) "The production and perception of speech sounds in a foreign language" In H. Winits ed., *Human Communication and Its Disorders* 3, Norwood, New Jersey, Ablex

Flege, J.E. (1987d) "Effects of equivalence classification on the production of foreign language speech sounds" In A. James & J. Leather eds, 9-39

Flege, J.E. (1990) "Perception and production: the relevance of phonetic input to L2 phonological learning" In C. Ferguson & T. Huebner eds, *Crosscurrents in Second Language Acquisition and Linguistic Theories*, Philadelphia, John Benjamins

Flege, J.E. (1992) "The intelligibility of English vowels spoken by British and Dutch talkers" In R. Kent ed., 565-604

Flege, J.E. & R. Davidian (1984) "Transfer and developmental processes in adult foreign language production" *Journal of Applied Psycholinguistic Research* 5, 323-347

Flege, J.E. & W. Eefting (1986) "Linguistic and developmental effects on the production and perception of stop consonants" *Phonetica* 43, 155-171

Flege, J.E. & W. Eefting (1987a) "The production and perception of English stops by native Spanish speakers" *Journal of Phonetics* 15, 67-83

Flege, J.E. & W. Eefting (1987b) "Cross-language switching in stop consonant production and perception by Dutch speakers of English" *Speech Communication* 6, 185-202

Flege, J.E. & W. Eefting (1988) "Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation" *Journal of the Acoustical Society of America* 83, 729-740

Flege, J.E. & R. Hammond (1982) "Mimicry of non-distinctive phonetic differences between language varieties" *Studies in Second Language Acquisition* 5, 1-17

Flege, J.E. & J. Hillenbrand (1984) "Limits on pronunciation accuracy in adult foreign language speech production" *Journal of the Acoustical Society of America* 76, 708-721

Flege, J.E. & J. Hillenbrand (1985) "Differential use of temporal cues to the [s-z] contrast by native and non-native speakers of English" *Journal of the Acoustical Society of America* 79, 508-517

Flege, J.E. & J. Hillenbrand (1987) "A differential effect of release bursts on the stop voicing judgments of native French and English listeners" *Journal of Phonetics* 15, 203-208

Flege, J.E., M. McCutcheon & S.C. Smith (1987) "The development of skill in producing word-final English stops" *Journal of the Acoustical Society of America* 82, 433-447

Flege, J.E. & R.F. Port (1981) "Cross-language phonetic interference: Arabic to English" *Language and Speech* 24, 125-146

Flege, J.E. & A.M. Schmidt (1995) "Native speakers of Spanish show rate-dependent processing of English stop consonants" *Phonetica* 52, 90-111

Flege, J.E. & C. Wang (1989) "Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t/ - /d/ contrast" *Journal of Phonetics* 17, 299-315

Fourakis, M. & G.K. Iverson (1985) "On the acquisition of second-language timing patterns" *Language Learning* 35, 431-442

Fowler, C.A. (1977) *Timing Control in Speech Production* Bloomington, Indiana University Linguistics Club

Fowler, C.A. (1980) "Coarticulation and theories of extrinsic timing" *Journal of Phonetics* 8, 113-133

Fowler, C.A. (1983) "Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in sequences of monosyllabic stressed feet" *Journal of Experimental Psychology: Human Perception and Performance* 112, 386-412

Fowler, C.A. (1990) "Calling a mirage a mirage: direct perception of speech produced without a tongue" *Journal of Phonetics* 18, 529-541

Fowler, C.A., P. Rubin, R.E. Remez & M.T. Turvey (1980) "Implications for speech production of a general theory of action" In Butterworth ed., *Language Production*, 373-420, New York, Academic Press

Fries, C. (1945) *Teaching and Learning English as a Foreign Language*, Ann Arbor, MI, The University of Michigan Press

Fry, D.B. (1955) "Duration and intensity as physical correlates of linguistic stress" *Journal of the Acoustical Society of America* 27, 765-768

Fry, D.B. (1958) "Experiments in the perception of stress" *Language and Speech* 1, 126-152

Fry, D.B., A.S. Abramson, P.D. Eimas & A.M. Liberman (1962) "The identification and discrimination of synthetic vowels" *Language and Speech* 5, 171-189

Fujisaki, H., K. Nakamura & T. Imoto (1973) "Auditory perception of duration of speech and non- speech stimuli" *Annual Bulletin of Research Institute of Logopaedics and Phoniatrics* 7, 45-64, University of Tokyo

Fujisaki, H. & N. Higuchi (1979) "Temporal organisation of segmental features in Japanese disyllables" *Proceedings of the Ninth International Congress of Phonetic Sciences* 2, 275-282, University of Copenhagen

Fukui, N. (1986) "Leftward spreading: Compensatory lengthening and Gemination in Japanese" *Linguistic Inquiry* 17, 356-363

Fukui, S. (1978) "Perception for the Japanese stop consonants with reduced extended durations" *The Bulletin of the Phonetic Society of Japan* 159, 9-12

Gaitenby, J. (1965) "The elastic word" Haskins Lab. Status Rep. Speech Res., SR-2, 3.1-3.12, Haskins Laboratories, New Heaven

Garnes, S. (1973) "Phonetic evidence supporting a phonological analysis" *Journal of Phonetics* 1, 273-284

Gass, F. (1984) "Development of speech perception and speech production in adult second language learners" *Applied Psycholinguistics* 5, 51-74

Goldsmith, J. (1976) "An overview of autosegmental phonology" *Linguistic Analysis* 2, 23-68

Goto, H. (1971) "Auditory perception by normal Japanese adults of the sounds 'L' and 'R'" *Neuropsychologia* 9, 381-410

Grieser, D. & P.K. Kuhl (1989) "Categorization of speech by infants: Support for speech-sound prototypes" *Developmental Psychology* 25, 577-588

Guilford, J.P. (1959[1954]) *Psychometric Methods*, Trans. Y. Akie, Tokyo, Baifuukan [In Japanese]

Habein, Y.S. (1984) *The History of the Japanese Written Language*, University of Tokyo Press

Hakuta, K. (1976) "A case study of a Japanese child learning English as a second language" *Language Learning* 26/2, 321-351

Han, M.S. (1962a) "The feature of duration in Japanese" *Study Sounds* 10, 65-80

Han, M.S. (1962b) *Japanese Phonology: An Analysis Based upon Sound Spectrograms*, Tokyo, Kenkyuusha

Han, M.S. (1992) "The timing control of geminate and single stop consonants in Japanese: a challenge for nonnative speakers" *Phonetica* 49, 102-127

Han, M.S. (1994) "Acoustic manifestation of mora timing in Japanese" *Journal of the Acoustical Society of America* 96/1, 73-82

Haraguchi, S. (1977) *The Tone Patterns of Japanese: An Autosegmental Theory of Tonology*, Tokyo, Kaitakusha

Harnad, S. ed. (1987) *CategoricalPerception*, Cambridge, Cambridge University Press

Harris, K.S., B. Tuller & J.A.S. Kelso (1986) "Temporal invariance in the production of speech" In J. Perkell & D. Klatt eds., 243-252

Hashimoto, M. (1977) "On'in no taikei to koozoo" *Nichi-ei Hikaku Kooza 1 - Onsei to Keitai*, Tokyo, Taishuukan [in Japanese]

Hashimoto, M. (1980) "On'in taikei no hikaku" *Iwanami Kooza Nihongo 5 - On'in*, Tokyo, Iwanami Shoten [in Japanese]

Hattori, S. (1960) *Gengogaku no hoohoo;* Tokyo, Iwanami Shoten [in Japanese]

Hattori, S. (1984) *Onseigaku*, Tokyo, Iwanami Shoten [in Japanese]

Hayes, B. (1989) "Compensatory lengthening in moraic phonology" *Linguistic Inquiry* 20, 2, 253-306

Healy, A.F. & Repp, B.H. (1982) "Context independence and phonetic mediation in categorical perception" *Journal of Experimental Psychology: Human Perception and Performance* 8, 68-80

Hirata, Y. (1988) *Nihonjin to Eigowasha ni yoru Sokuon no Kikitori to Seisei-Tango Reberu, Bun Reberu deno Kikitori Jikken o Chuushin ni*, Unpublished MA thesis, Tsukuba University, Japan [in Japanese]

Hirata, Y. (1990a) "Perception of geminated stops in Japanese word and sentence levels" *The Bulletin of the Phonetic Society of Japan* 194, 23-28 [in Japanese]

Hirata, Y. (1990b) "Perception of geminated stops in Japanese word and sentence levels by English-speaking learners of Japanese language" *The Bulletin of the Phonetic Society of Japan* 195, 4-10 [in Japanese]

Hirayama, T. (1979) "Zen nihon no hatsuon to akusento [Pronunciation and accent in Japan]" In *Nihongo Hatsuon Akusento Jiten* [A Dictionary of Japanese Pronunciation and Accent], 103-138, Tokyo, Japan Broadcasting Cooperation [in Japanese]

Hirose, H. (1971) "The activity of the adductor laryngeal muscles in respect to vowel devoicing in Japanese" *Phonetica* 23, 156-170

Hirozane, Y. (1992) "Perception by Japanese Speakers of some English sounds as the Japanese choked sound /Q/" *The Bulletin of the Phonetic Society of Japan* 201, 15-19

Hoequist, C. (1983a) "Durational correlates of linguistic rhythm categories" *Phonetica* 40, 19-31

Hoequist, C. (1983b) "Syllable duration in stress, syllable, and mora-timed languages" *Phonetica* 40, 203-237

Hockett, C.F. (1955) *A Manual of Phonology*, Indiana University Publications in Anthropology and Linguistics, Memoir 11, Bloomington

Homma, Y. (1981) "Durational relationship between Japanese stops and vowels" *Journal of Phonetics* 9, 273-281

Homma, Y. (1982) "The rhythm of Tanka, short Japanese poems" *Proceedings of the Thirteenth International Congress of Linguistics*, 618-624

Homma, Y. (1985) *Acoustic Phonetics in English and Japanese*, Kyoto, Yamaguchi

House, A.S. & Fairbanks, G. (1953) "The influence of consonant environment upon the secondary acoustical characteristics of vowels" *Journal of the Acoustical Society of America* 25, 105-113

Huggins, A.W.F. (1971) "Just noticeable differences for segment duration in natural speech" *Journal of the Acoustical Society of America* 51, 1270-1278

Hyltenstam, K. & M. Pienemann eds. (1985) *Modelling and Assessing in Second Language Acquisition*, Clevedon, Multilingual Matters

Hyman, L. (1985) *A Theory of Phonological Weight*, Dordrecht, Foris Publications

Imai, T. & M.Saito eds. (1986) *Issues in Japanese Linguistics*, Dordrecht, Foris Publications

Ingram, D. (1974) "Fronting in child phonology" *Journal of Child Language* 1, 233-241

Ingram, D. (1979) "Phonological patterns in the speech of young children" In P. Fletcher & M. Garman eds., *Language Acquisition: Studies in First Language Development*, 133-148, Cambridge, Cambridge University Press

International Phonetic Association (1984[1949]) *The Principles of the International Phonetic Association*, IPA, London

Ioup, G. (1984) "Is there a structural foreign accent? A comparison of syntactic and phonological errors in second language acquisition" *Language Learning* 34, 2, 1-17

Ioup, G. & S.H. Weinberger eds. (1987) *Interlanguage Phonology. The Acquisition of Second Language Sound Systems*, Newbury House, Rowley, Mass.

Jakobson, R. (1968[1941]) *Child Language, Aphasia and Phonological Universals*, Trans.A.R. Keiler, The Hague, Mouton

James, A. & J. Leather eds. (1987) *Sound Patterns in Second Language Acquisition*, Dordrecht, Foris Publications

Janson, T. (1979) "Vowel duration, vowel quality, and perceptual compensation" *Journal of Phonetics* 7, 93-103

Johansson, F.A. (1973) *Immigrant Swedish Phonology: A Study in Multiple Contact Analysis*, Lund, Sweden, CWK Gleerup

Jooo, H. (1988) "Kotoba no rizumu" *Gekkan Gengo* 17/3, 24-31 [in Japanese]

Kahn, D. (1975) *Syllable-based Generalizations in English Phonology*, Indiana University Linguistics Club, Bloomington

Kaiki, N., K. Takeda & Y. Sagisaka (1990) "Statistical analysis for segmental duration rules in Japanese speech synthesis" *Proceedings of the International Conference on Spoken Language Processing* 1, 17-20, Tokyo

Kaiki, N. & Y. Sagisaka (1992) "The control of segmental duration in speech synthesis using statistical methods" In Y. Tohkura et al. eds., 391-402

Kamei, T. (1956) "Onsei no gainen wa Nihongo ni yuuyoo nari ya" *Kokubungakukoo* 15 [in Japanese]

Kashima, T. (1987) *Acoustic and Auditory Study of the Pronunciation of Japanese by Australian Students*, Unpublished Ph.D dissertation, Monash University

Kashima, T. (1992) "Nihongo no rizumu tan'i to sono kata ni tsuite-Nihongo kyooiku e no ooyoo o mezashite" In H.C.Quackenbush et al. eds., *Nihongo Kenkyuu to Nihongo Kyooiku*, 305-319, Nagoya University [in Japanese]

Katada, F. (1990) "On the representation of moras: Evidence from a language game" *Linguistic Inquiry* 21, 641-646

Kawagoe, I. (1995) "Shakuyoogo ni miru sokuonka to rizumu shoototsu (Consonant gemination in loanword phonology and rhythm clash" *Gengo Kenkyuu* 108, 46-73 [in Japanese]

Kawakami, S. (1982) "Nihongo no rizumu no genri" *Kokugakuin Zasshi*, 48-55 [in Japanese]

Kawakami, S. (1984) "Rizumu-kata no ninshiki to nihongo kyooiku (The Metrical Distinction of Japanese Words- an Aid of Teaching Japanese)" *The Bulletin of the Phonetic Society of Japan* 175, 7-9 [in Japanese]

Kawakami, S. (1987) "Facts about the so-called Japanese nasal syllabic consonants" *The Bulletin of the Phonetic Society of Japan* 185, 18-20 [in Japanese]

Kawashima, T. (1983) "A comparative analysis of British & American vowels spoken by Japanese university students" *The Bulletin of the Phonetic Society of Japan* 172, 12-14 [in Japanese]

Keller, E. (1992) *Signalyze: Signal Analysis for Speech and Sound,* Manual, Seattle, Info Signal Inc.

Kelso, J.A.S., B. Tuller & K.S. Harris (1983) "A 'dynamic pattern' perspective on the control and coordination of movement" In P. MacNeilarge ed., 137-173

Kent, R. ed (1992) *Intelligibility in speech disorders: Theory, measurement and management* Amsterdam, John Benjamins

Kindaichi, H. (1942) "Gagyoo bionron" (Reprinted in Kindaichi 1967) [in Japanese]

Kindaichi, H. (1967) *Nihongo On'in no Kenkyuu*, Tokyo, Tokyoodoo Shuppan [in Japanese]

Kingston, J. & M.E. Beckman eds. (1989) *Papers in Laboratory Phonology* 1: *Between the Grammar and the Physics of Speech*, Cambridge, Cambridge University Press

Klatt, D.H. (1975) "Voice onset time, frication and aspiration in word-initial clusters" *Journal of Speech and Hearing Research* 18, 686-705

Klatt, D.H. (1976) "The linguistic use of segmental duration in English: Acoustic and perceptual evidence" *Journal of the Acoustical Society of America* 59, 1208-1221

Kohler, K. (1979a) "Parameters in the production and the perception of plosives in German and French" In Barry & Kohler *'Time' in the Production and Perception of Speech*, Abeitsberichte 12, 261-280, Institute of Phonetics, Kiel University

Kohler, K. (1979b) "Dimensions in the perception of fortis and lenis plosives" *Phonetica* 36, 332-343

Krashen, S.D. (1981) *Second Language Acquisition and Second language Learning,* Oxford, Pergamon Press

Krashen, S.D. (1982) *Principles and Practice in Second Language Acquisition,* New York, Pergamon Press

Krashen, S.D., M. Long & R. Scarcella (1979) "Age, rate, and eventual attainment in second language acquisition" *TESOL Quarterly* 13, 125-129

Kubozono, H. (1985) "Speech errors and syllable structure" *Linguistics and Philology* 6, 220-243

Kubozono, H. (1989) "The mora and syllable structure in Japanese: Evidence from speech errors" *Language and Speech* 32/3, 249-278

Kubozono, H. (1990) "Phonological constraints on blending in English as a case for the phonology-morphology interface" *Yearbook of Morphology* 3, 1-20

Kuhl, P.K. (1981) "Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech sound categories" *Journal of the Acoustical Society of America* 70, 340-349

Kuhl, P.K. (1982) "Speech perception: An overview of current issues" In N.J. Lass et al. eds. 286-322

Kuhl, P.K. (1986) "Reflections on infants' perception and representation of speech" In J. Perkell & D. Klatt eds., 19-30

Kuhl, P.K. (1990) "Towards a new theory of the development of speech perception" *Proceedings of the International Conference on Spoken Language Processing* 2, 745-748, Tokyo

Kuhl, P.K. (1991) "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not" *Perception and Psychophysics* 50, 93-107

Kuhl, P.K. (1993) "Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception" *Journal of Phonetics* 21, 125-139

Kuhl, P.K. (1994) "Forming the brain's perceptual maps: Effects of language experience on speech perception" *Lecture given at Yakult International Symposium, Brain and Communication*, 26 May 1994, Tokyo

Kurisu, K. (1994) "Further evidence for bi-moraic foot in Japanese" *Proceedings of the International Conference on Spoken Language Processing* 1, 367-370, Yokohama

Ladd, D.R. (1989) "Metrical representation of pitch register" In J. Kingston & M.E. Beckman eds., 35-57

Ladefoged, P. (1967) *Three Areas of Experimental Phonetics*, Oxford, Oxford University Press

Ladefoged, P. (1982[1975]) *A Course in Phonetics*, Second Edition, New York, Harcourt Brace Jovanovich

Lado, R. (1957) *Linguistics across Cultures*, Ann Arbor, MI, The University of Michigan Press

Lane, H. (1965) "The motor theory of speech perception: A critical review" *Psychological Review* 7, 275-309

Lange, R.A. (1973) *The Phonology of Eighth-century Japanese: A Reconstruction Based on Written Records*, Tokyo, Sophia University

Larsen-Freeman, D. & M.H. Long (1991) *An Introduction to Second Language Acquisition Research*, Longman, UK

Lass, N.J., L.V. McReynolds, J.L. Northern & D.E. Yoder (1982) *Speech, Language, and Hearing* 1, W.B. Saunders Company

Lehiste, I. (1970) *Suprasegmentals*, MIT Press

Lehiste, I. (1972) "The timing of utterances and linguistic boundaries" *Journal of the Acoustical Society of America* 51, 2018-2024

Lehiste, I. (1976) "Influence of fundamental frequency pattern on the perception of duration" *Journal of Phonetics* 4, 113-117

Lehiste, I. (1977) "Isochrony reconsidered" *Journal of Phonetics* 5, 253-264

Lehiste, I. (1989) "Experimental studies of poetic rhythm" *Hungarian Papers in Phonetics* 21, 365-368, Linguistics Institute of the Hungarian Academy of Sciences, Budapest

Lehiste, I. (1990) "Phonetic investigation of metrical structure in orally produced poetry" *Journal of Phonetics* 18, 123-133

Lehiste, I. (1994) "Poetic metre, prominence, and the perception of prosody: A case of intersection of art and science of spoken language" *Paper presented at the plenary lectures, International Conference on Spoken Language Processing*, 20, 2237-2243, Yokohama

Lehiste, I., K. Morton & M.A.A. Tatham (1973) "An instrumental study of consonant gemination" *Journal of Phonetics* 1, 131-148

Liberman, A.M. (1979) *The Intonational System of English*, Indiana University Linguistics Club, Bloomington

Liberman A.M., K.S. Harris, H.S. Hoffman & B.C. Griffith (1957) "The discrimination of speech sounds within and across phoneme boundaries" *Journal of Experimental Psychology: Human Perception and Performance* 54, 358-368

Liberman, A.M., K.S. Harris, P.D. Eimas, L. Lisker & J. Bastian (1961a) "An effect of learning on speech perception: The discrimination of durations of silence with and without significance" *Language and Speech* 54, 175-195

Liberman, A.M., K.S. Harris, J.A. Kinney & H. Lane (1961b) "The discrimination of relative onset time of the components of certain speech and nonspeech patterns" *Journal of Experimental Psychology: Human Perception and Performance* 61, 379-388

Liberman, A.M., F.S. Cooper, D.P. Shankweiler & M. Studdert-Kennedy (1967) "Perception of the speech code" *Psychological Review* 74, 431-461

Liberman, A.M. & A. Prince (1977) "On stress and linguistic rhythm" *Linguistic Inquiry* 8-2, 249-336

Lieberman, P. (1977) *Speech Physiology and Acoustic Phonetics*, New York, Macmillan

Lieberman, P. & S.E. Blumstein (1988) *Speech Physiology, Speech Perception, and Acoustic Phonetics*, Cambridge, Cambridge University Press

Lindblom, B. & K. Rapp (1973) "Some temporal regularities of spoken Swedish" *Papers from the Institute of Linguistics* 21, University of Stockholm

Lindblom, B., B. Lyberg & K. Holmgren (1981) *Durational Patterns of Swedish Phonology: Do They Reflect Short-term Memory Processes?* Indiana University Linguistics Club, Bloomington

Lisker, L. (1957) "Closure duration and the intervocalic voiced-voiceless distinction in English" *Language* 33, 42-49

Lisker, L. (1974) "On time and timing in speech" In Sebeok & Abramson *Current Trends in Linguistics* 12, The Hague, Mouton

Lisker, L. & A. Abramson (1964) "A cross-language study of voicing in initial stops: acoustical measurements" *Word* 20, 384-422

Lisker, L. & A. Abramson (1967) "Some effects of context on voice onset time in English stops" *Language and Speech* 10, 1-29

Lisker, L. & A. Abramson (1971) "Distinctive features and laryngeal control" *Language* 47, 767-785

Long, M. H. (1983) "Does second language instruction make a difference? A review of research" *TESOL Quarterly* 17/3, 359-382

Long, M. H. (1985) "A role for instruction in second language acquisition: Task-based language teaching" In K. Hyltenstam & M. Pienemann eds., 77-99

Long, M. H. (1988) "Instructed interlanguage development" In L. Beebe ed., 115-141

Long, M. H. (1993) "Sensitive periods in second language acquisition" *Paper presented at the Eighteenth Congress of Applied Linguistics Association of Australia*, University of Adelaide, September 1993, South Australia

Mabuchi, K. (1973) "Manyooshuu no on'in" *Hisamatsu Manyooshuu Kooza* 3 161-186, Tokyo, Yuuseidoo [in Japanese]

MacKain, K., C. Best & W. Strange (1981) "Categorical perception of English /r/ and /l/ by Japanese bilinguals" *Applied Psycholinguistics* 2, 369-390

Macken, M. & D. Barton (1977) "A longitudinal study of the acquisition of the voicing contrast in American English word initial stops, as measured by VOT" *Papers and Reports on Child Language Development* 14, 74-121

MacNeilarge, P. ed. (1983) *The Production of Speech*, New York, Springer

Maeda, S. (1979) "Timing control in Japanese speech production" *Tenri Daigaku Gakuhoo* 121, Tenri University Report, September, 1-21 [in Japanese]

Maekawa, K. (1989) "Boin no musei-ka" In M. Sugito ed., *Kooza Nihongo to Nihongo Kyooiku* 2, Tokyo, Meiji Shoin, 135-153 [in Japanese]

Major, R. (1987) "A model for interlanguage phonology" In G. Ioup & S.H. Weinberger eds., 101-124

Mehler, J., J.Y. Dommergues, U. Frauenfelder & J. Segui (1981) "The syllable's role in speech segmentation" *Journal of Verbal Learning and Verbal Behaviour* 20, 298-305

McCawley, J.D. (1968) *The Phonological Component of a Grammar of Japanese*, The Hague, Mouton

McCawley, J.D. (1978) What is a tone language? In V.A. Fromkin ed., *Tone*, New York, Academic Press, 113-131

Miller, J.L., C.M. Connine, T.M. Schermer & K.R. Kluender (1983) "A possible auditory basis for internal structure of phonetic categories" *Journal of the Acoustical Society of America* 73, 2124-2133

Miller, J.L. & P.D. Eimas (1977) "Studies on the perception of place and manner of articulation: A comparison of the labial-alveolar and nasal-stop distinctions" *Journal of the Acoustical Society of America* 61, 835-845

Milon, J. (1974) "The development of negation in English by a second language learner" *TESOL Quarterly* 8, 137-143

Min, K. (1987) "Kankoku-jin no nihongo no sokuon no chikaku ni tsuite" *Nihongo Kyooiku* 62, 179-193 [in Japanese]

Minagawa, Y. (1994) "Bogokanshoo sareta heisaon no musei jikan, VOT ni tsuite- nanakakokugo kaku bogo washa no hatsuwa shiryoo yori" *Paper presented at the Autumn Conference of the Society for Teaching Japanese as a Foreign Language,* 9 October, Toohoku University [in Japanese]

Minagawa, Y. (1995) "Kankokujin gakushuusha ni okeru akusento, tokushuhaku chooshu nooryoku to gakushuunensuu ni yoru sono henka" *Paper presented at the Second Language Acquisition Research Forum*, International Christian University, September 1995, Tokyo [in Japanese]

Miura, I. (1982) "Syllabic duration in relation to speaking rate" *The Bulletin of the Phonetic Society of Japan* 171, 12-14 [in Japanese]

Miyawaki, K., W. Strange, R. Verbrugge, A. Liberman, J. Jenkins & O. Fujimura (1975) "An effect of linguistic experience: the discrimination of /r/ and /l/ by native speakers of Japanese and English" *Perception and Psychophysics* 18, 331-340

Mizutani, O. & N. Mizutani (1987) *An Introduction to Modern Japanese*, Tokyo, Japan Times

Mochizuki, M. (1981) "The identification of /r/ and /l/ in natural and synthesised speech" *Journal of Phonetics* 9, 283-303

Mochizuki-Sudo, M. & S. Kiritani (1991) "Production and perception of stress-related durational patterns in Japanese learners of English" *Journal of Phonetics* 19, 231-248

Muraki, M. & N. Nakaoka (1990) "Hatsuon to sokuon - Eigo, Chuugokugo washa no hatsuon" In M. Sugito ed., *Kooza Nihongo to Nihongo Kyooiku* 3, Tokyo, Meiji Shoin [in Japanese]

Nagano-Madsen, Y. (1992) *Mora and Prosodic Coordination: A phonetic study of Japanese, Eskimo and Yoruba*, Sweden, Lund University Press

Nakajoo, O. (1992) *Nihongo no Onsei*, Tokyo, Keisoo Shoboo [in Japanese]

Nakamichi, M. (1980) "Nihongo no rizumu no tan'i ni tsuite (A study in Rhythmic unit of Japanese)" *Onsei Gengo no Kenkyuu* 1, Tokyo University of Foreign Languages, 15-22 [in Japanese]

Nakaoka, N. (1990) "Nihongo no fuhensei to kobetsusei ni kansuru rironteki oyobi jisshooteki kenkyuu" *Kenkyuu Seika Hookokusho* (6A) Juutenryooiki Kenkyuu (Grant-in-Aid for Scientific Research on Priority Areas), Tokyo, 179-182 [in Japanese]

Nearey, T.M. (1990) "The segment as a unit of speech perception" *Journal of Phonetics* 18, 347-373

Nemser, W. (1971a) "Approximative systems of foreign language learners" *International Review of Applied Linguistics* 9, 115-123

Nemser, W. (1971b) *An Experimental Study of Phonological Interference in the English of Hungarians*, Indiana University, Bloomington

Nolan, F. (1990) "Who do phoneticians represent?" *Journal of Phonetics* 18, 453-464

Oasa, H. (1980) *A Quantitative Study of Regional Variations in Australian English*, Unpublished MA thesis, Australian National University

O'Dell, M. & R.F. Port (1983) "Discrimination of word-final voicing in German" *Journal of the Acoustical Society of America* 73 suppl.1, S31

Odlin, T. (1989) *Language Transfer: Cross-linguistic Influence in Language Learning*, Cambridge, Cambridge University Press

Ohala, J. (1974) "Phonetic explanation in phonology" In A. Bruck et al. eds. *Papers from the Parasession on Natural Phonology*, 251-274, Chicago, Chicago Linguistic Society

Ohala, J. (1975) "The temporal regulation of speech" In Fant & Tatham, *Auditory Analysis and Perception of Speech*, London, Academic Press

Ohala, J. (1980) "The application of phonological universals in speech pathology" In N.J. Lass ed., *Speech and Language: Advances in Basic Research and Practice* 3, 75-97, New York, Academic Press

Ohala, J. (1981) "Speech timing as a tool in phonology" *Phonetica* 38, 204-212

Ohala, J. & S. Lyberg (1976) "Comments on 'Temporal interactions within a phrase and sentence'" *Journal of the Acoustical Society of America* 59, 990-992

Oller, O.K. (1973) "The effect of position in utterance on speech segment duration in English" *Journal of the Acoustical Society of America* 54, 1235-1247

Omae, N. (1983) "Duration of English vowel and its reduced phase" *The Bulletin of the Phonetic Society of Japan* 172, 7-8 [in Japanese]

Otaka, H. (1987) "Nihongo no onsetsu koozoo to rizumu" *Gekkan Gengo* 16/6, 82-92, Tokyo, Taishuukan [in Japanese]

Otaka, H. (1988) "Onsetsu to wa nani ka, mora to wa nani ka" *Gekkan Gengo* 17/3, 60-68, Tokyo, Taishuukan [in Japanese]

Otake, T. (1989a) "Nihongo gakushuusha no Nihongo onsei no rizumu" *Denshi Jyoohoo Tsuushin Gakkai-shi* SP88-156, 15-20 [in Japanese]

Otake, T. (1989b) "A temporal compensation effect in Arabic and Japanese" *The Bulletin of Phonetic Society of Japan* 189, 19-24

Otake, T. (1989c) "Eigo o bokokugo to suru Nihongo gakushuusha no onsei no jikan-teki tokusei" *Bulletin of Kokusai Budoo Daigaku* 4 [in Japanese]

Otake, T. (1989d) "A cross linguistic contrast in the temporal compensation effect" *The Bulletin of Phonetic Society of Japan* 191, 14-19

Otake, T. (1991) "Nihongo onsei no segmentation unit" *Denshi Jyoohoo Tsuushin Gakkai-shi* SP90-108, 41-46 [in Japanese]

Otake, T. (1992) "Nihongo onsei no chikakujoo no tan'i: Onsetsu to moora" *Nihongo no Moora to Onsetsu ni kansuru Soogooteki Kenkyuu* 1, 38-47 [in Japanese]

Otake, T. (1993) "Nihongo onsei no chikaku no kokoo: onsetsu to moora no kyoozon no kanoosei" *Nihongo no Moora to Onsetsu ni kansuru Soogooteki Kenkyuu* 2, 51-71 [in Japanese]

Otake, T., G. Hatano, A. Cutler & J. Mehler (1993) "Mora or syllable? Speech segmentation in Japanese" *Journal of Memory and Language* 32, 258-278

Ootsubo, K. (1980) "Nihonjin no chooboin, tanboin no handan nooryoku ni tsuite" *Gengo Bunka Ronshuu* 2/1, Nagoya University [in Japanese]

Ootsubo, K. (1981) "Nihonjin no sokuon no umu no handan nooryoku ni tsuite" *Gengo Bunka Ronshuu* 3/1, Nagoya University [in Japanese]

Ootsubo, K. (1990) "Nihongo onsei no kyooikuhoo no kaihatsu" *Nihongogaku* 8, 72-79 [in Japanese]

Pastore, R.E., W.A. Ahroon, K.J. Baffuto, C. Friedman, J.S. Puleo & E.A. Fink (1977) "Common factor model of categorical perception" *Journal of Experimental Psychology: Human Perception and Performance* 3, 686-696

Perkell, J. & D. Klatt eds. (1986) *Invariance and Variety in Speech Processes*, Hillsdale, NJ, Erlbaum

Peterson, G. & I. Lehiste (1960) "Duration of syllabic nuclei in English" *Journal of the Acoustical Society of America* 32, 693-703

Pienemann, M. (1984) "Psychological constraints on the teachability of languages" *Studies in Second Language Acquisition* 6 (2), 186-214

Pienemann, M. (1985) "Learnability and syllabus construction" In K. Hyltenstam & M. Pienemann eds., 23-75

Pienemann, M. & M. Johnston (1986) "An acquisition-based procedure for second language assessment" *Australian Review of Applied Linguistics* 9, 92-122

Pienemann, M. (1987) "Determining the influence of instruction of L2 speech processing" *Australian Review of Applied Linguistics* 10/2, 83-113

Pienemann, M. (1989) "Is language teachable?" *Applied Linguistics* 10/1, 52-79

Pienemann, M. & M. Johnston (1987) "Factors influencing the development of language proficiency" In D. Nunan ed., *Applying Second Language Acquisition Research*, 45-141, National Curriculum Resource Centre, Adult Migrant Education Program, Adelaide

Pienemann, M., M. Johnston & G. Brindley (1988) "Constructing an acquisition-based procedure for second language assessment" *Studies in Second Language Acquisition* 10/2, 217-243

Pierrehumbert, J.B. (1990) "Phonological and phonetic representation" *Journal of Phonetics* 18, 375-394

Pierrehumbert, J.B. & M.E. Beckman (1988) *Japanese Tone Structure*, Linguistic Inquiry Monograph, Cambridge

Pike, K.L. (1945) *The Intonation of American English*, Ann Arbor, University of Michigan Press

Pike, K.L. (1947) *Phonemics*, Ann Arbor, University of Michigan Press

Pike, K. (1979) "Universals and phonetic hierarchy" *Proceedings of Ninth International Congress of Phonetic Science* 2, 48-52, Institute of Phonetics, University of Copenhagen, Copenhagen

Port, R.F. (1977) *The Influence of Speaking Tempo on the Duration of Stressed Vowel and Medial Stop in English Trochee Words*, Indiana University Linguistics Club, Bloomington

Port, R.F. (1979a) "The influence of tempo on stop closure duration as a cue for voicing and place" *Journal of Phonetics* 7, 45-56

Port, R.F. (1979b) "Combination of timing factors in speech production" In Wolf & Klatt eds., *Speech Communication Papers*, 97th meeting of the Acoustical Society of America, 193-196

Port, R.F. (1981) "Linguistic timing factors in combination" *Journal of the Acoustical Society of America* 69, 262-274

Port, R.F., S. Al-Ani & S. Maeda (1980) "Temporal compensation and universal phonetics" *Phonetica* 37, 235-252

Port, R.F., J. Dalby & M. O'Dell (1987) "Evidence for mora timing in Japanese" *Journal of the Acoustical Society of America* 81/5, 1574-1585

Port, R.F. & R. Rotunno (1979) "Relation between voice-onset time and vowel duration" *Journal of the Acoustical Society of America* 66, 654-662

Poser, W.J. (1990) "Evidence for foot structure in Japanese" *Language* 66, 78-105

Prince, A.S. (1980) "A metrical theory for Estonian quantity" *Linguistic Inquiry* 11, 511-562

Quackenbush, H.C., A. Ozaki, M. Fujiwara, T. Kashima & Y. Momiyama (1992) *Nihongo Kenkyuu to Nihongo Kyooiku*, Nagoya University [in Japanese]

Raphael, L. (1972) "Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English" *Journal of the Acoustical Society of America* 51, 1296-1303

Renard, R. (1975) *Introduction to the Verbo-tonal Method of Phonetic Correction*, Paris, Didier

Reichardt, C. & T. Cook (1979) *Beyond Qualitative and Quantitative Methods in Education Research*, Sage Publications, Beverly Hills

Repp, B.H. (1984) "Categorical perception: Issues, methods, findings" In N.J. Lass ed., *Speech and Language: Advances in Basic Research and Practice* 10, 243-335, New York, Academic Press

Rischel, J. (1990) "What is phonetic representation?" *Journal of Phonetics* 18, 395-410

Roberge, C. & M. Kimura (1990) *Nihongo no Hatsuon Shidoo: VThoo no Riron to Jissai*, Tokyo, Bonjinsha

Rosch, E (1977) "Human categorisation" In N. Warren ed., *Studies in Cross-Cultural Psychology*, 1-49, London, Academic Press

Roug, L., I. Landberg & L.J. Lundberg (1989) "Phonetic development in early infancy- a study of Swedish during the first 18 months of life" *Journal of Child Language* 16, 19-40

Sagisaka, Y. & Y. Tohkura (1984) "Phoneme duration control for speech synthesis by rule [in Japanese]" *Journal of Electronic Information Communication Engineering* 67-A, 629-636

Samuel, A.G. (1982) "Phonetic prototypes" *Perception and Psychophysics* 31, 307-314

Sawashima, M., H. Hirose, K. Honda & M.Sugito (1980) "Relative timing of articulatory and phonatory controls in Japanese word accent: a study on nonsense two-mora words" *Annual Bulletin of Research Institute of Logopaedics and Phoniatrics* 14, 139-147, University of Tokyo

Saita, I. (1987) "Gaikokujin nihongo gakushuusha no tame no onsei gakushuu shien shisutemu no kenkyuu" *Kenkyuu Seika Hookokusho - Sookatsuhan, Juutenryooiki Kenkyuu* (Grant-in-Aid for Scientific Research on Priority Areas), 97-100 [in Japanese]

Saito, S. (1986) "Hanashi kotoba ni okeru ra-gyoo on oyobi na-gyoo on no mora-ka" *Nihongo Kyooiku* 60, 205-219 [in Japanese]

Sato, C. (1984) "Phonological processes in second language acquisition: another look at interlanguage syllable structure" *Language Learning* 34, 43-57

Sato, H. (1993) "Gairaigo o zairyoo to shita akusento no kentoo- rizumu ni chakumoku shite-" *Gaikokujin o Taishoo to Suru Nihongokyooiku ni okeru Onseekyooiku no Hoosaku ni kansuru Kenkyuu*, Tokyo, 5-18 [in Japanese]

Sato, H. (1994) "Bimoraic rhythm and accentuation in Japanese" *Paper presented at the Twelfth National Conference of the English Linguistic Society of Japan*, 85-90, 13 November, University of Tokyo

Sato, Y. (1990) "A spectrographic analysis of the duration of the mora nasal in Japanese" *The Bulletin of the Phonetic Society of Japan* 193, 12-17

Sato, Y. (1993) "The durations of syllable-final nasals and the mora hypothesis in Japanese" *Phonetica* 50, 44-67

Sawashima, M. (1971) "Devoicing of vowels" *Annual Bulletin of Research Institute of Logopaedics and Phoniatrics* 5, 7-13, University of Tokyo

Schachter, J. (1974) "An error in error analysis" *Language Learning* 24, 205-214

Schachter, J. & M. Celce-Murcia (1977) "Some reservations concerning error analysis" *TESOL Quarterly* 11, 441-451

Schmidt, A.M. & J.E. Flege (1995) "Effects of speaking rate changes on native and nonnative speech production" *Phonetica* 52, 41-54

Schmidt, R. (1977) "Sociolinguistic variation and language transfer in phonology" *Working Papers in Bilingualism* 12, 79-95

Schults, R.W. (1960) "Problem solving behaviour and transfer" *Harvard Educational Review* 30, 61-77

Scovel, T.S. (1969) "Foreign accents, language acquisition and cerebral dominance" *Language Learning* 19, 245-254

Scovel, T.S. (1976) "Contrastive analysis is alive and well living in phonology" *Paper presented at the Tenth Annual National TESOL Convention*, New York

Seliger, H.W. & E. Shohamy (1989) *Second Language Research Methods*, Oxford, Oxford University Press

Selinker, L. (1972) "Interlanguage" *International Review of Applied Linguistics* 10, 209-231 (Reprinted in J.C. Richards ed., *Error Analysis: Prospectives on Second Language Acquisition*, London, Longman, 31-54)

Sheldon, A. & S. Strange (1982) "The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception" *Applied Psycholinguistics* 3, 243-261

Shibata, T. (1958) "Conditions controlling standardization" *Nihon no Hoogen* (The Dialects of Japan), Tokyo, Iwanami Shoten

Shibata, T. (1962) "On'in" In Kokugogaku Gakkai ed., *Hoogengaku Gaisetsu*, 137-61 [in Japanese]

Shibatani, M (1990) *Languages of Japan*, Cambridge, Cambridge University Press

Smith, B.L. (1978) "Temporal aspects of English speech production: A developmental perspective" *Journal of Phonetics* 6, 37-67

Smith, C.L. (1995) "Prosodic patterns in the coordination of vowel and consonant gestures" In B. Connell & A. Arvaniti eds., *Phonology and Phonetic Evidence: Papers in Laboratory Phonology* IV, 205-222, Cambridge, Cambridge University Press

Soli, S.D. (1981) "Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation" *Journal of the Acoustical Society of America* 70, 976-984

Stampe, D. (1969) "The acquisition of phonetic representation" *Papers from the Fifth Regional Meeting*, Chicago Linguistic Society

Stevens, K.N. & A. House (1972) "Speech perception" In Tobias ed., *Foundations of Modern Auditory Theory* 2, 1-62, New York, Academic Press

Stevick, E.W. (1982) *Teaching and Learning Languages*, Cambridge, Cambridge University Press

Strange, W. (1992) "Learning non-native phoneme contrasts: Interactions among subject, stimulus, and task variables" In Y. Tohkura et al. eds., 197-219

Strange, W. & S. Dittman (1984) "Effects of discrimination training on the perception of /r-l/ by Japanese adults" *Perception and Psychophysics* 36, 131-145

Strange, W. & J. Jenkins (1978) "The role of linguistic experience in the perception of speech" In R. Walk & H. Pick, Jr. eds., *Perception and Experience*, 125-126, New York, Plenum

Studdert-Kennedy, M. (1976) "Speech perception" In N.J. Lass ed., *Contemporary Issues in Experimental Phonetics*, 234-295, New York, Academic Press

Sugito, M. (1980) "Akusento, intoneeshon no hikaku" *Nichiei Hikaku Kooza 1 - Onsei to Keitai*, Tokyo, Taishuukan, 107-183 [in Japanese]

Sugito, M. (1982) *Nihongo Akusento no Kenkyuu*, Tokyo, Sanseido [in Japanese]

Sugito, M. (1984) "Nihongo no onsetsu (haku) niwa dooiu Nihongo rashisa ga aru ka" *Kokubungaku Kaishaku to Kyoozai no Kenkyuu*, 29/6, 32-39 [in Japanese]

Sugito, M. (1989) "Onsetsu ka, haku ka" *Kooza Nihongo to Nihongo Kyooiku* 2, Tokyo, Meiji Shoin, 154-177 [in Japanese]

Sugito, M. (1990) "Nihongo to Eigo no akusento to intoneeshon" *Kooza Nihongo to Nihongo Kyooiku* 3, Tokyo, Meiji Shoin, 349-378 [in Japanese]

Sugito, M. (1991) "Nihongo onsei no tokushitsu" *Nihongogaku* 10, 40-50 [in Japanese]

Tabata, T. (1986) "Nihongo no onsetsu koozoo to moora" *Ooyoo Gengogaku Kooza 2 - Gaikokugo to Nihongo*, Tokyo, Meiji Shoin [in Japanese]

Takahashi, M. (1986) "Perception of Japanese flaps by American listeners" *The Bulletin of the Phonetic Society of Japan* 183, 18-22 [in Japanese]

Takeda, K., Y. Sagisaka & H. Kuwabara (1989) "On sentence-level factors governing segmental duration in Japanese" *Journal of the Acoustical Society of America* 86, 2081-2087

Tanaka, Y. (1977) *Shinrigakuteki Sokuteihoo*, Tokyo, Tokyo University Press [in Japanese]

Tarone, E. (1972) "A suggested unit for interlingual identification in pronunciation" *TESOL Quarterly* 6/4, 325-331

Tarone, E. (1978) "The phonology of interlanguage" In J. Richards ed., *Understanding Second and Foreign Language Learning*, 15-33, Newbury House, Rowley, Mass.

Tarone, E. (1980) "Some influences on the syllable structure of interlanguage phonology" *International Review of Applied Linguistics* 18, 139-152

Teranishi, R. (1980) "Two-mora-cluster as a rhythm unit in spoken Japanese sentence or verse" *Journal of the Acoustical Society of America* 67/1, 40

Terrell, T. (1977) "A natural approach to second language acquisition and learning" *Modern Language Journal* 61, 325-337

Toda, T. (1991) "Application of sound digitizer in teaching the timing of geminate consonants in Japanese" *Proceedings of the Seventh Biennial Conference of Japanese Studies Association of Australia*, 123-128, Canberra

Toda, T. (1993a) "Speech production by Australian DFAT learners of Japanese: Durational contrast in vowels, obstruents and nasal consonants" *Paper presented at the Eighth Biennial Conference of Japanese Studies Association of Australia*, University of Newcastle, July 1993, Newcastle

Toda, T. (1993b) "Acquisition of timing control in Japanese" *Paper presented at the Eighteenth Congress of Applied Linguistics Association of Australia*, University of Adelaide, September 1993, South Australia

Toda, T. (1994) "Interlanguage phonology: Acquisition of timing control in Japanese", *Australian Review of Applied Linguistics* 17/2, 51-76, Applied Linguistics Association of Australia

Toda, T. (1995a) "Aspects of English-Japanese interlanguage phonology: Development of inaccurate perceptual categorization", *Paper presented at the Ninth Biennial Conference of Japanese Studies Association of Australia*, University of Queensland, July 1995, Brisbane

Toda, T. (1995b) "Eigo-washa ni yoru tokushuhaku no shuutoku no genkai ni tsuite (Limitations in the acquisition of the perception of Japanese mora by English-speaking learners)" *Paper presented at the Second Language Acquisition Research Forum*, International Christian University, September 1995, Tokyo [in Japanese]

Toda, T. & H.C. Quackenbush (1992) "English loan words in Japanese-Acquisition of phonological rules by Australian students" *Paper presented at the Inaugural Pacific Second Language Research Forum*, Sydney University, July 1992, Sydney

Tohkura, Y., E. Vatikiotis-Bateson & Y. Sagisaka eds. (1992) *Speech Perception, Production and Linguistic Structure*, Tokyo, Ohmsha

Trehub, S.E. (1976) "The discrimination of foreign speech contrasts by infants and adults" *Child Development* 47, 466-472

Trubetzkoy (1969 [1939]) *Principles of Phonology*, Trans. C. Baltaxe, Berkeley, University of California Press

Uchida, T. (1993a) "Characteristics of auditory cognition of long vowels and double consonants for Chinese students in learning Japanese language" *Japanese Journal of Educational Psychology* 41, 414-423 [in Japanese]

Uchida, T. (1993b) "Onsei o atsukau tame no paasonaru compyuutaa o mochiita shinrigaku jikken soochi" *Bulletin of the School of Education* 40, 227-237 [in Japanese]

Vance, T.J. (1987) *An Introduction to Japanese Phonology*, Albany, State University of New York Press

Vatikiotis-Bateson, E. & S. Kelso (1993) "Rhythm type and articulatory dynamics in English, French and Japanese" *Journal of Phonetics* 21, 231-265

Wardhaugh, R. (1970) "The contrastive analysis hypothesis" *TESOL Quarterly* 4, 123-130

Watanabe, S. & N. Hiratoo (1985) "Nionsetsu-go ni okeru musei haretsuon to sokuon no handan kyookai to senkoo boin no nagasa no kankei" *Onsei Gengo 1*, Kinki Onsei Gengo Kenkyuukai, 1-8 [in Japanese]

Weinreich, U. (1953) *Language in Contact*, Linguistic Circle of New York

Weismer, G. (1979) "Sensitivity of voice-onset-time (VOT) measures to certain segmental features in speech production" *Journal of Phonetics* 7, 197-204

Werker, J.F., J. Gilbert, I. Humphry & R.C. Tees (1981) "Developmental aspects of cross-language speech perception" *Child Development* 52, 349-355

Werker, J.F. & R.C. Tees (1984) "Phonemic and phonetic factors in adult cross-language speech perception" *Journal of the Acoustical Society of America* 75, 1866-1878

Werker, J.F. & L. Polka (1993) "Developmental changes in speech perception: New challenges and new directions" *Journal of Phonetics* 21, 83-101

Werker, J.F. & J.S. Logan (1985) "Cross-language evidence for three factors in speech perception" *Perception and Psychophysics* 37, 35-44

Williams, P. (1990) "Evidence for foot structure in Japanese" *Language* 66-1, 78-105

Williams, P.D. & D.J. Sharf (1979) "Effect of adaptation on the perception and production of vowel duration preceding stop consonants" *Journal of Phonetics* 7, 81-92

Wilson, M. (1987) "Brain mechanisms in categorical perception" in S. Harnad ed., *Categorical Perception: The Groundwork of Cognition*, Cambridge, Cambridge University Press

Wode, H. (1976) "Developmental sequences in naturalistic second language acquisition" *Working Papers in Bilingualism* 11,.1-31

Wode, H. (1977) "The L2 acquisition of /r/" *Phonetica* 34, 200-217

Wode, H. (1981) "Language-acquisitional universals: A unified view of language acquisition" In H. Winitz ed., *Native Language and Foreign Language Acquisition*, Annals of the New York Academy of Sciences 379, 218-234

Wode, H. (1990) "Continuity in the development of language acquisitional abilities" In H. Burmeister & P.L. Rounds eds., *Variability in Second Language Acquisition: Proceedings of the Tenth Meeting of the Second Language Research Forum* 1, 85-116, University of Oregon

Wode, H. (1991) *Speech Perception: A Developmental Perspective*, Mimeo, Kiel University

Wode, H. (1992) "Categorical perception and segmental coding in the ontogeny of sound systems: A universal approach" In C.A. Ferguson, L. Menn & C. Stoel-Gammon eds., *Phonological Development: Models, Research, Implications*, 605-631, Timonium, MD: York Press

Wode, H. (1993) "The development of phonological abilities" In K. Hyltenstam & A. Viberg eds., *Progression and regression in Language: Sociocultural, Neuropsychological and Linguistic Perspectives,* 415-438, Cambridge, Cambridge University Press

Wode, H. (1994) "Speech perception and the learnability of languages" *International Journal of Applied Linguistics* 4/2, 143-168

Wolf, C.G. (1978) "Voicing cues in English final stops" *Journal of Phonetics* 6, 299-309

Yatabe, S. (1987) "English syllabic consonants and Japanese nasal syllabic consonants" *The Bulletin of the Phonetic Society of Japan* 185, 14-16 [in Japanese]

Yeni-Komshian, G.H., A. Caramazza & M. Preston (1977) "A study of voicing in Lebanese Arabic" *Journal of Phonetics* 5, 35-49

Yeni-Komshian, G.H., J. Kavanagh & C.A. Ferguson eds. (1980) *Child Phonology* 2 *Perception*, New York, Academic Press

Yoshiba, H. (1983) *Moraic Phonology: Toward the Establishment of a New Phonological Unit*, Ph.D dissertation, University of Washington

Yoshida, S. (1990) "A government-based analysis of the 'mora' in Japanese" *Phonology* 7, 331-351

Yoshida, Y., N. Kuratani & S. Okunishi (1976) *Japanese for Beginners*, Tokyo, Gakken

Yoshimitsu, K. (1980) "Gaikokujin no Nihongo no jittai (4)- Gaikokujin gakushuusha no akusento" *Nihongo Kyooiku* 45, 63-74 [in Japanese]