OUKA
The University of Osaka
Institutional Knowledge Archive

Title	Markedness of Prepausal Vowels in Japanese Rhythmic Structure
Author(s)	森, 庸子
Citation	大阪大学, 2003, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/51189
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

博士論文

題目　Markedness of Prepausal Vowels in Japanese Rhythmic Structure

提出年月　2002年12月

言語社会研究科　　言語社会専攻

氏名　森　庸子

Acknowledgements

I would like to express my deep gratitude to Shiro Kôri, my supervisor, for his valuable suggestions and comments, and supervisions, on this dissertation. I am also grateful to Isao Ueda and Tadaharu Tanomura, my associate supervisors, for their valuable comments.

From designing the phonetic experiments to checking my English, I owe a great deal to many people for their suggestions, help, kindness, and encouragements, that have led me to accomplish this study. I would like to thank them all for their time and/or labor that they have kindly spent for me at various stages of this investigation (in alphabetical order): Nick Campbell, Gerry Docherty, Vincent J. van. Heuven, Tony Higgins, Kazue Hiraoka, Mitsuyo Iwasa, Eiji Kaneda, Shigeru Kiritani, Morio Kohno, Haruo Kubozono, Kenji Matsui, Yukihiko Nakai, Akio Nasu, Yoshinori Sagisaka, Keiichi Tajima.

要旨

日本語リズム構造におけるポーズ前母音の有標性

森　庸子

本論文は、モーラリズム言語とされてきた日本語が、モーラリズムとは対照的なもう一つの時間構造を持つことを論究するものである。それは、モーラリズムの制約外にあり、大幅な長さの変化が許される、ポーズ前母音である。本研究では、モーラリズムの制約により日本語では一般に観察されないとされてきたアクセントの持続時間への影響を検討することによって、ポーズ前母音の日本語リズム構造における有標性を解明しようとする。

第１章では、日本語の音声研究の流れと課題を紹介し、本研究の目的を述べる。従来から、モーラリズムの日本語では、各モーラ（CV・V 音節および特殊拍）内または隣接モーラ間で音素の長さを調節し、各モーラの長さをほぼ一定に保とうとする傾向が顕著に観察されてきた。また、このモーラリズムの制約から、日本語のアクセントは音素長に影響を及ぼさないとされてきた。従って、日本語の音声研究においては、音声の高さの変化（アクセントやイントネーション）に重点が置かれ、各音素の持続時間は、モーラリズムによりほぼ規則的に実現されると考えられてきた。従って、アクセントと長さの相関は一般に否定され、その関係を詳細に検討した研究はなかった。

これに対して、本研究では、延べ 13 人の東京方言話者を対象とした音声実験を８回繰り返し、話者の特別な発話意図を持たない語・句・平叙文の発話を音響分析することにより、アクセントと長さとの関係を詳細に検討する。そして、それらの実験結果から、アクセントによるピッチ下降の有無により、ポーズ前母音長に一貫して大きな有意差が観察されることを明らかにする。さらにその結果に基づき、ポーズ前という位置の日本語リズム構造における有標性とその意義を論究することを目的とする。

第２章では、日本語の時間構造に影響を及ぼす要因について検討する。実験１より、発話の長さが日本語の音素の長さに大きな影響を与えることが確認され、５モーラまでは、発話長が長くなるほど、構成要素の音素長が短くなることが観

察された。また、「マ」に置換した発話を用いた実験２により、アクセントのアクセント核のあるモーラの持続時間への影響は否定された。しかし、位置の影響は観察され、第１子音が発話内の子音中、最も短いのに対して、発話末母音は、アクセントによるピッチ下降のない場合の方が、アクセントによるピッチ下降のある場合より、有意に伸長されることが明らかになった。

第３章では、実際にある音素構造が同じでアクセント型の違う語および句の単独発話の実験により、第２章で観察されたアクセントと発話末母音長の相関を検討する。実験３、４A、４B より、発話末にアクセントによるピッチ下降のない語がある場合は、アクセントによるピッチ下降のある語がある場合に比べて、発話末母音長が有意に約 40ms、平均 1.40 倍から 1.43 倍長いことが明らかになった。また、発話末の母音の長さとピッチには相関があり($p<0.0001$)、発話末でピッチが下がると、母音長が短くなる傾向が示唆された。

第４章では、短い平叙文(実験５)と、ポーズを含む長い平叙文(実験６) におけるアクセントと長さの相関を検討した。まず、NP+VP の構造を持つ短文の発話においては、VP におけるアクセントによるピッチ下降の有無により、文末母音長に約 40ms の有意差が生じる。このアクセントによる影響は文末に限定された現象であり、NP におけるアクセントによるピッチ下降の有無は、句末母音長にも文末母音長にも影響しない。

ポーズを含む長文の発話では、ポーズ前母音長がアクセントと位置により有意に異なることがわかった。アクセントの影響によるポーズ前母音長の差は、文末では、語・句・短文の場合とほぼ同じだが(1.42:1)、非文末のポーズ前では、アクセントの影響も若干観察されるものの、アクセントによるピッチ下降に関らず、母音長が大きく伸長される傾向が観察された。

また、位置の影響によりポーズ前母音長は、非文末では文末に比べて、アクセントによるピッチ下降のない場合は 1.24 倍に、アクセントによるピッチ下降のある場合は、1.54 倍に伸長されていた。アクセント・位置の２つの影響による差は、各セットで最大 66%から 93%に及び、最長の母音は最短の母音の平均約 1.74 倍となっている。

さらに、ポーズ前母音における長さとピッチの相関は、文末では顕著だが、非文末では観察されない。従って、非文末では、ピッチに関らず意図的に母音を伸

長しようとする傾向が示唆され、文末と非文末では、異なる韻律制御機構が機能していると考えられる。

第５章では、第４章の実験結果から、平叙文における文末・非文末のポーズ前母音の長さと高さを予測する韻律モデルを提案する。

文末では、文末のフレーズにアクセントによるピッチ下降のない場合には母音長が「長」となり、アクセントによるピッチ下降のある場合には「短」となる。一方、非文末のポーズ前では、アクセントによるピッチ下降のない場合には「最長」、ある場合には「長」となった。この韻律モデルは、相対的持続時間を予想するモデルだが、さまざまな音素構造からなる文に共通して適用できるモデルであり、「最長」「長」「短」の各レベルの設定は、本研究で得られたデータにおける高い統計的有意差に基づいている（$p<0.0001$）。

さらに、平叙文の発話におけるポーズ前母音のピッチは、ポーズ前および先行フレーズにおけるアクセントによるピッチ下降の有無と、文末か非文末かにより予測することができる。すなわち、非文末では、先行およびポーズ前のフレーズにアクセントによるピッチ下降のない場合は、「最高」であり、先行フレーズのみにアクセントによるピッチ下降のある場合は「高」、ポーズ前フレーズにアクセントによるピッチ下降のある場合は、先行フレーズにおけるアクセントによるピッチ下降の有無に関らず、「中」となる。一方、文末では、先行フレーズにおけるアクセントによるピッチ下降の有無に関らず、ポーズ前フレーズにアクセントによるピッチ下降のない場合は「高」、ある場合は「低」に２分できる。持続時間の場合と同様に、各レベルの間には、高い統計的有意差（$p<0.0001$）が検出され、これらのレベルは、文の音素構造に関らず適用できる。

この韻律モデルは、アクセントによるピッチ下降の有無と文末・非文末の区別という２つの条件の組み合わせだけで、平叙文の発話におけるポーズ前母音の長さと高さの両方を予測するものである。

第６章では、前章までで明らかになったアクセントによるポーズ前母音長の変化を、日本語における他の伸縮現象と比較することによって、ポーズ前の母音は、モーラリズムの制約および音韻的長さの制約を受けていない特殊な存在であることを論じる。

実験７により、日本語の１モーラ語は助詞を伴わずに、フレーズ末で孤立する

と、「韻律語は最小でも２モーラ必要」という最小語条件を満たすため、40%から50%伸長されることが明らかになった。しかし、この場合にも、50%を超える伸長は観察されないことから、日本語では母音の長短が弁別的機能を果たすという音韻的制約が強く働いていることが示唆される。さらに、実験７および先行研究より、ポーズを伴わないフレーズ末でもわずかな伸長は観察されるが、その程度は、約 10%以内である。また、アクセントによる持続時間の差異はほとんど全く観察されないが、例外的に観察される場合も、その程度は約 10ms 以内である。

これに対して、本研究で明らかになったポーズ前の母音長の差異は、50%を超える例も少なからず観察され、その他の時間長の伸縮現象に比べて、その伸縮の程度がはるかに大きい。音素構造が同一であるにも関らず、ポーズ前の母音長は、アクセントにより一貫して有意に大きく異なり、しかも、先行子音の長さを補償しモーラ長を一定に保とうとするモーラリズムの補償効果も、ポーズ前という位置では観察されない。すなわち、これらのことから、ポーズ前母音は、日本語のモーラリズムの制約外にあると考えられ、同時にポーズ前という位置では、母音長の長短の区別もあいまいになる傾向にあることが示唆される。

第７章では本研究で明らかになったポーズ前母音の大幅な伸縮性とその有標性が、日本語および日本語のリズム構造に対して、どのような意味を持つかを論究する。

日本語では、文末またはポーズ前の韻律が大きな役割を果たすとされてきた。例えば、疑問文では顕著な上昇調を実現するため母音を伸長するように、顕著なピッチ変化には、音素（特に母音）の伸長が伴う傾向にある。従って、ポーズ前の母音では大幅な伸縮が許されるため、長さとピッチにより様々な韻律を実現し、各韻律に特有な機能を果たしていることが示唆される。

また、日本語では従来から、「そうでしょ」「ただいまー」などの発話における長母音の短縮、短母音の伸長現象が観察されてきた。モーラリズムとされる日本語において、なぜポーズ前にのみ、短呼、長呼現象が許されるか、という本質的問題は、本研究で明らかになったポーズ前母音の有標性で説明すべき問題である。すなわち、ポーズ前母音では、モーラリズムの制約が解除され、母音の伸縮が自在になり、母音の長短の音声的対立があいまいになるため、このような短呼、長呼現象が生起すると考えられる。

さらに、終助詞およびポーズ前に生起する間投助詞も、母音長の大幅な変化が許されるというポーズ前の特殊性により、多様な韻律を実現し、各終助詞および間投助詞特有の様々な機能を果たすことができると考えられる。

また、これらの終助詞・間投助詞や、前述のポーズ前母音の韻律が表すものには、共通点があるように思われる。それらは、主として、聞き手、もしくは話し手自身の発話に対する、話し手の態度、意図、感情、発話スタイル、コメントなどを表すものである。

これまで、日本語の時間構造については、モーラリズムの主張を支持する報告が繰り返されてきた。2モーラフットなどの韻律単位が提唱されることもあるが、これらの韻律単位もすべて、モーラリズムの制約下にあるモーラにより構成されていた。

しかし、本研究で明らかになったポーズ前の母音の大幅な伸縮性は、ポーズ前がモーラリズムの制約外であることを示すものであると考える。すなわち、日本語が、モーラの等時性を保とうとするモーラリズム以外に、モーラリズムの範囲外にあって自在な伸縮性を持つポーズ前という時間構造を持つことを示すものと考える。この時間構造は、ポーズ前に限定されるものであるが、その多様な韻律により、日本語の話しことばにおいて音声的に卓立し、日本語によるコミュニケーションにおいて、大きな役割を果たしている可能性が示唆される。

Contents

Acknowledgements ii

要旨 iii

1 Introduction

1.1 Roles of pitch and duration in Japanese prosody 1

1.1.1 Japanese as a pitch-accent language 1

1.1.2 Japanese as a mora-timed language 4

1.2 Intonation in Japanese prosody 5

1.2.1 Pitch contour and utterances 5

1.2.2 Sentence-final pitch contour 7

1.2.3 Fujisaki's Model 8

1.2.4 The Beckman and Pierrehumbert Model 9

1.3 Final lengthening in Japanese 11

1.4 Goals of this study 12

2. Temporal structure of Japanese 15

2.1 Introduction 15

2.2 Experiment 1 17

2.2.1 Subjects 18

2.2.2 Materials 18

2.2.3 Procedure 19

2.3 Results & discussion 21

2.3.1 Mora timing in a frame sentence 21

2.3.2 Effect of utterance length on mora duration 22

2.3.3 Effect of utterance length on segmental duration 23

2.3.4 Mora timing in longer utterances 27

2.4 Experiment 2 28
2.4.1 Subjects 29
2.4.2 Materials 29
2.4.3 Procedure 30
2.5 Results & discussion 31
2.5.1 Effect of position on mora duration 31
2.5.2 Effect of accent on accented mora 36
2.5.3 Effect of accent on final moras 37
2.5.4 Isochrony in moras 43
2.5.5 Bimoraic unit 44
2.6 Summary 47

3 Effect of accent on duration in isolated words and phrases 50
3.1 Introduction 50
3.2 Experiment 3 51
3.2.1 Subjects 51
3.2.2 Materials 51
3.2.3 Procedure 52
3.3 Results & discussion 54
3.3.1 Effect of accent on segmental duration 54
3.3.2 Effect of accent on word final vowels 60
3.3.3 Effect of accent on final lengthening in isolated words 63
3.3.4 Interspeaker variation 64
3.3.5 Correlation between F0 and duration in isolated words 66
3.4 Experiment 4A 67
3.4.1 Subjects 68
3.4.2 Materials 68
3.4.3 Procedure 68
3.5 Results & discussion 69
3.5.1 Effect of N2's accent on utterance-final vowels 69

3.5.2 Effect of N1's accent on utterance-final vowels 72
3.5.3 Utterance-final lengthening vs. word-final lengthening 75
3.6 Experiment 4B 76
3.6.1 Subjects 77
3.6.2 Materials 77
3.6.3 Procedure 77
3.7 Results & discussion 78
3.7.1 Effect of segmental variation on utterance-final vowels 78
3.7.2 Phrase-final lengthening vs. utterance-final lengthening 79
3.7.3 Correlation between F0 and duration in isolated phrases 83
3.8 Summary 84

4 Effect of accent on duration in declarative sentences 87
4.1 Introduction 87
4.2 Experiment 5 88
4.2.1 Subjects 88
4.2.2 Materials 88
4.2.3 Procedure 89
4.3 Results & discussion 90
4.3.1 Effect of accent on the duration of sentence-final vowels 90
4.3.2 Effect of accent on phrase-final vowels 94
4.3.3 Correlation between F0 and duration in sentence-final vowels 96
4.4 Experiment 6 98
4.4.1 Subjects 98
4.4.2 Materials 98
4.4.3 Procedure 99
4.5 Results & discussion 101
4.5.1 Effect of accent and position on the duration of prepausal vowels 101

4.5.2 Effect of preceding consonants on prepausal vowels 105
4.5.3 Effect of position on the duration of penultimate vowels 107
4.5.4 Effect of accent and position on the pitch of prepausal vowels 112
4.5.5 Correlation between F0 and duration in prepausal vowels 116
4.6 Summary 119

5. Prosodic control in declarative sentences 122
5.1 Introduction 122
5.2 Prosodic control of sentence-final vowels in short declarative sentences 122
5.3 Prosodic control of prepausal vowels in declarative sentences with pauses 128
5.3.1 Effect of test words in long declarative sentences 128
5.3.2 Duration control of prepausal vowels in long sentences 129
5.3.3 Pitch control of prepausal vowels in long sentences 131
5.3.4 Prosodic control of prepausal vowels in declarative sentences with intervening pauses 133
5.4 Summary 135

6. Markedness of prepausal vowels in Japanese rhythmic structure 136
6.1 Introduction 136
6.2 Lengthening of monomoraic nouns 137
6.2.1 Introduction 137
6.2.2 Experiment 7 139
6.2.3 Results and discussion 142
6.2.4 General discussion 149
6.2.5 Summary 152
6.3 Lengthening of moras under accent or focus 153

6.4 Phrase-final lengthening with no following pauses 154
6.5 Lengthening of penultimate vowels in non-utterance-final position 156
6.6 Durational variability in prepausal vowels 156
6.7 Markedness of prepausal vowels in Japanese rhythmic structure 159
6.7.1 Prepausal position outside of mora-timing constraint 159
6.7.2 Phonological constraint on vowel length in prepausal position 160
6.7.3 Markedness of prepausal vowels in Japanese rhythmic structure 162

7. Theoretical implications and concluding remarks 164
7.1 Introduction 164
7.2 Final shortening in Japanese vs. final lengthening in English 165
7.3 Durational variability and intonation in prepausal position 167
7.4 Phonetic contrast between short and long vowels in prepausal position 169
7.5 Prosody of sentence-final and interjectional particles 171
7.6 Functions of prosody in prepausal vowels 172
7.7 Prepausal vowels in Japanese rhythmic structure 173
7.8 Concluding remarks 175

References 177

Appendix 184

CHAPTER 1

Introduction

1.1 Roles of pitch and duration in Japanese prosody

1.1.1 *Japanese as a pitch-accent language*

Japanese has long been known as a pitch-accent language in which pitch[1] is what matters in determining word's accent (cf. Kindaichi, 1967; McCawley, 1968). Possible accent patterns of three-mora words followed by the particle /ga/ in Tokyo Japanese, for instance, are as follows: ("H" denotes high pitch, while "L" denotes low pitch)

Table 1-1. Accent pattern of three-mora words (Tokyo Japanese)

word +/ga/	accent pattern	position of accent
sakana ga	unaccented	0
L H H H		
kabuto ga	initially-accented	1
H L L L		
kokoro ga	medially-accented	2
L H L L		
kotoba ga	finally-accented	3
L H H L		

When a mora is "accented" or an "accent (nucleus)" is located on a mora, the sudden pitch fall is assumed to take place at or around the boundary between the accented mora (or the mora with the accent nucleus) and the following mora. In this thesis, this sharp pitch fall caused by accent is called an "accentual (F0) fall" (Kubozono, 1993), and the signal ' is used to mark the boundary between an accented mora and the next one.

A word without an accentual fall is called an "unaccented" word, while a word with an accentual fall is referred to as "initially-accented" or "medially-accented" word, depending on the position of the accented mora within a word. Accordingly, the term "finally-accented" refers to those words that have accentual falls between the final mora and the following particle, if any. Thus, if there is no following particle, a finally-accented word is assumed to have no accentual fall. Also, the position in a word of an accented mora occasionally represents accent patterns (Table 1-1). If the first mora of a word is accented, for instance, the accent pattern can be represented by "1". If there is no accented mora, "0" is used to indicate it. In this thesis, any accent pattern mentioned follows the above usage of accent and definitions of accent patterns.

As a pitch-accent language, it is generally believed that the primary phonetic correlate of word accent in Japanese is a sudden drop in pitch (Kubozono, 1993). However, the questions arise whether the sharp fall in pitch corresponds with acoustic drop in F0 and whether the other elements of sound such as intensity or duration have anything to do with the manifestation of Japanese accent.

Beckman (1986) conducts acoustic and perceptual experiments on the realization of accents in Japanese and English. She finds that total amplitude and duration are reliable phonetic correlates

and perceptual cues to accent in English, but not in Japanese, pointing out the distinctive differences in the phonetic manifestation of "stress accents" (as found in English) and "non-stress accents" (as found in Japanese).

Hoequist (1983b) reports only a small but consistent increase of 2% in duration of Japanese moras under accent, which shows a highly significant difference from that found in a stress- or syllable-timed language such as English or Spanish. Kaiki, Takeda & Sagisaka (1992) also assert that there is no effect of accent on the duration of Japanese syllables. Sakaguchi & Yoshida (1999) report no correlation observed between accent and duration in their corpus.

Thus, their results support the traditional assumption that Japanese accent is realized solely by pitch or F0 and that neither duration nor intensity plays significant roles in the manifestation of accent. Why is accent in Japanese manifested exclusively by pitch? The most widely accepted answer to this question is that Japanese is a mora-timed language in which moras maintain a roughly constant duration. It is assumed, therefore, that Japanese moras under accent are not allowed to be substantially lengthened due to the constraint of mora timing.

In addition, the length of vowel (long vs. short) is one of the distinctive features in Japanese phonology. The word "*obaasan*" which includes a long vowel /a:/, for instance, means an old woman, whereas the word "*obasan*" whose second vowel is a short vowel /a/ refers to an aunt or a middle-aged woman. If an accented mora is uttered twice as long, therefore, the word may be perceived as a different word.

Moreover, previous studies (Beckman, 1982; Hoequist, 1983a) have reported that the ratio of a heavy (two-mora) and light (one-mora) Japanese syllables ranges from 1.66:1 to 1.80:1. It is

suggested, therefore, that under this phonological constraint accented moras are not allowed to exceed more than about one and a half times the non-lengthened forms.

1.1.2 *Japanese as a mora-timed language*

The mora functions as an abstract unit of duration in Japanese phonology (Bloch, 1950) by which the length of words, phrases, or utterances is measured. A light syllable constitutes one mora composed of an optional consonant and a following short vowel (V or CV). Although they do not constitute a syllable on their own, some other elements also serve as a mora. They are a moraic nasal, a moraic obstruent (or the first half of a geminate consonant), and the second half of a long vowel or diphthong. Thus, the word *sanpo*^{2} "a walk", for instance, consists of two syllables but of three moras, while *gakkoo* "school" has two syllables but is four moras in length.

Japanese has long been described as a mora-timed language, in contrast with stress-timed or syllable-timed languages. By this, it is meant that isochrony in Japanese is manifested at the mora level. The strong version of the mora-timing hypothesis claims the equal phonetic duration of moras (Han, 1962). This assertion has been refuted by acoustic investigations showing that the actual duration of moras is subject to considerable variation (Beckman, 1982; Hoequist, 1983a).

Instead, some researchers observe that Japanese speakers tend to compress or stretch segments in an attempt to keep each mora roughly equal in duration (Sagisaka & Tohkura, 1984; Campbell, 1999). In other words, they claim temporal compensation within the unit of a mora in Japanese. Some have further observed temporal compensation across several moras or in larger units of production such as words (Port, Al-Ani & Maeda, 1980; Port, Dalby & O'Dell,

1987; Han, 1994). They report that Japanese words with the same number of moras are almost equal in duration, and that word duration increases in proportion to the number of moras contained, suggesting acoustic realization of mora timing at the word level.

Other researchers report no large increase in duration of Japanese moras under accent or focus, compared with that found in so-called stress-timed or syllable-timed languages, apparently due to the constraint of mora timing (cf. 1.1.1). Hoequist (1983a, b) has carried out cross-linguistic studies on the tendency to maintain a constant duration of the proposed rhythmic unit by measuring the durations of words set in a frame sentence. He reports significantly less lengthening of Japanese moras under accent than that found in English or Spanish. Kaiki et al. (1992) have found no effect of accent on the duration of vowels in their corpus, suggesting a mora-timing constraint and a control mechanism of duration independent from that of F0. Kôri (1989) reports that the role of duration in the realization of focus in Japanese is inconsistent, and thus secondary.

As is clear from the above, researchers have agreed that there is no correlation between accent and duration in Japanese partly due to the manifestation of accent by pitch and partly due to the constraint of mora timing. Thus, Japanese prosody has been mainly explored or discussed by means of its accent and/or pitch contour, assuming that duration plays no significant role in the prosody.

1.2 Intonation in Japanese prosody

1.2.1 *Pitch contour and utterances*

What the term "intonation" refers to differs among researchers. In the present study, "intonation" is defined as the contour of fundamental frequency of an utterance.

Before reviewing previous studies on Japanese intonation, it is important to discuss the various units of production observed in spoken Japanese and the definitions of terms that refer to them. An "utterance" is a term widely used to refer to a stretch of speech in linguistic discussion. Bloomfield (1926) introduced the word by defining it as "An act of speech is an utterance" (p. 154). Bloch (1949) pointed out that "(this definition) fails to tell us just how much of speech an utterance is supposed to include"(p.7), while Hattori (1968) has presented a more concrete definition of the domain of "an utterance". According to Hattori (1968), "an utterance" is considered to end when a speaker finishes talking and keeps silent, or starts doing something else (p.3). Fries (1949) has incorporated "a shift of speaker" into his definition of an utterance, and called all those stretches of speech that were bounded by a change of speaker an "utterance unit".

In accordance with the above definitions, the domain of an "utterance" is defined in this thesis as any stretch of connected speech produced by a single speaker. The term "pause-bounded unit (PBU)" is, on the other hand, used here to refer to a unit bounded by pauses.

In all the experiments carried out in the present investigation, each subject was presented a card, and asked to read the word/phrase/sentence on it one at a time. Tokens from these experiments were produced in succession by each speaker, but the production was interrupted by a presentation of a new card every time the speaker finished reading one card. Thus, each production was isolated and had no connection with another. Therefore, it is assumed that each token obtained in this way constitutes an independent utterance.

Every token from Experiments 1 to 5 and Experiment 7 is an

independent utterance and can also be called a pause-bounded unit (PBU) since it is bounded by pauses. In Experiment 6, however, sentential utterances with intervening pauses after main syntactic boundaries were obtained. In this experiment, each production of a long test sentence is considered to constitute a single utterance made up of two PBUs.

1.2.2 *Sentence-final pitch contour*

In previous studies on Japanese intonation, the pitch contour in utterance- or sentence-final position has been widely discussed, since this final pitch pattern is assumed to play a key role in the intonation of a sentential utterance in Japanese. It is said, for instance, that in Tokyo Japanese a declarative sentence has a falling intonation towards the end of the sentence, whereas an interrogative sentence has a sentence-final rising intonation (Sugitô, 1990; Maekawa, 1999). However, even in the intonational pattern of a declarative sentence asserted by researchers, there seems to be considerable inconsistency.

Yoshizawa (1960) observes two distinctive F0 contours in the final vowel of a declarative sentence in Tokyo Japanese: one with a remarkable F0 fall within the final syllable; the other without. He considers the latter a basic intonational pattern of a typical declarative sentence and calls it a *"heityoo"* (level tone). He distinguishes the former pattern as a *"kootyoo"* (fall tone) from the latter and assumes that the *kootyoo* is used to display the speaker's dissatisfaction or contempt. Uemura (1989) claims that in a declarative sentence, pitch falls gradually towards the end of the sentence due to declination and that no additional pitch fall takes place at the end of the sentence.

By contrast, some researchers (Sugitô, 1983,1990; Fujisaki,

1992; Maekawa, 1999) have pointed out a tendency of a sharp pitch fall towards the end of a declarative sentence. Sugitô (1983) states that "in the final position of a Japanese sentence, pitch falls abruptly, which in most cases exhibits the end of sentence" (pp. 64), although she does not provide enough phonetic data to prove her assumption.

As is clear from the above, researchers' explanations of the final intonational patterns of a declarative sentence are not only inconsistent but also contradictory to one another. In addition, there has been no in-depth research to explore the relationship between the accent pattern of a sentence-final word and the pitch contour of the sentence, although both accent and intonation are closely related to each other.

1.2.3 *Fujisaki's Model*

Fujisaki's phonetic intonation model[3] (Fujisaki & Nakashima, 1969; Fujisaki, 1992; 2002) is the only quantitative model that attempts to describe F0 contours of Japanese words or sentences. By analyzing sentential utterances with a declarative intonation and a number of other F0 contours of sentences, Fujisaki (2002) states that "an F0 contour of a sentence can be considered to consist of two kinds of elements" (p.2): a phrase component and an accent component. The phrase component is "a slowly varying component which may or may not show a slight initial rise and then gradually decay towards an asymptotic baseline[4]" (p.2). This component is a voicing component which is responsible for the natural declination of the F0 contour, and can be resumed or reinforced at certain syntactic boundaries such as major phrasal boundaries between NPs and VPs. The accent component is, by contrast, a local hump (peak or plateau) "closely corresponding to the accent patterns of words constituting the sentence" (p.2).

Fujisaki has shown in many papers that the model can fit real F0 contours very accurately. However, it is often the case that an F0 contour falls to the minimal possible F0 value before the glottal oscillation stops after the last nuclear accent in an utterance or a sentence. Fujisaki (personal communication) considers it to be not an intentional lowering of F0, but a process leading to the total cessation of glottal oscillation, and approximates it by postulating a negative phrase component that allegedly causes the final F0 drop. However, Liberman & Pierrehumbert (1984) have pointed out that his model cannot correctly predict the utterance (or sentence) final lowering effect triggered by the nuclear accent of the utterance-final word even by using this negative phrase component.

The question arises whether the negative "phrase" component he proposes really is a "phrase" component. It seems that the abrupt F0 drop towards the end of a declarative sentence tightly combines with an accentual fall induced by a nuclear accent of the final word. According to his model, what is responsible for accent-related variation in pitch is the "accent" component, not the "phrase" component. This problem will be discussed in depth later in Chapters 4 and 5, by showing what a crucial role the last word's accent plays in the final F0 drop in a declarative sentence.

1.2.4 *The Beckman and Pierrehumbert Model*

In her doctoral dissertation, Pierrehumbert (1980) has proposed describing an intonation contour of English as a series of high and low tones, reducing the traditional four level descriptions to only two tones, which she designates as H (high tone) and L (low tone). Further distinction is made between tones associated with a pitch accent and those related to a boundary of a postlexical domain. The former is represented as *H or L*, the latter as L% (boundary L) or

H% (boundary H). In her theory, F0 contours are constructed by using each tone as a target and a set of rules to interpolate between the targets. In addition, both H and L are of an abstract height and cannot be mapped to a single F0 value. Furthermore, Pierrehumbert argues that there is no one-to-one mapping between the target level and the phonological level in her system.

Beckman & Pierrehumbert (1986; 1988) have compared the F0 contour of Japanese with that of English and attempted to apply that phonological description system to Japanese pitch contour. They also posit L% in Japanese unaccented phrases as well as in accented ones. They (1986) observe that, in a sequence of unaccented words, "F0 merely glides smoothly downward in a simple phonetic interpolation between the initial phrasal H and the following boundary L" (p.262). Thus, they posit "boundary L" instead of H-tone "spreading"[5].

It is evident that a large variation is observed in phonetic realization of the fundamental frequency of L% which is assumed to represent a low tone that allegedly appears at the edge of a domain in contrast with a preceding and/or following high tone. Even phonologically, however, there is doubt that Japanese perceive a phrase boundary tone as "low" in contrast with a phrasal high tone when the phrase consists of unaccented words only, that is, when it has no accentual fall.

Their model has been widely accepted, presumably because their system of describing an F0 contour can successfully describe any F0 contour of spoken Japanese or English. If an F0 contour is defined as an F0 change as a function of time, it should consist of a rise and/or a fall of F0 or its repetitions. Then, it is no wonder that the peak and valley of an F0 contour can be described as H and L, respectively, in Japanese as well as in English. The problem is that

their description cannot predict any pitch contour or a target H or L tone not only phonetically but also phonologically. Beckman & Pierrehumbert (1986; 1988) present neither typology nor classification of target H and/or L tones that play crucial roles in their phonological description of pitch contours.

1.3 Final lengthening in Japanese

Sagisaka et al. (1984) found word-final lengthening in prepausal position based on the analysis of 5,240 isolated word utterances produced by a male professional narrator. Takeda, Sagisaka, & Kuwabara (1989) analyzed continuous speech composed of 282 sentences uttered by the same speaker. In the recording two kinds of test sentences were employed: one consisted of sentences extracted from junior high school textbooks. The other came from experimental telephone conversations. In reading conversational materials, the speaker was allegedly instructed to "utter these samples naturally in a conversational style as far as possible" (p.2083).

They then applied the word-level duration model formulated on the large-scale database of 5,240 isolated words to the durations of the sentential utterances. Through error analysis, Takeda et al. have found the following two effects in the sentential utterances (p. 2086): (1) "at prepausal position, vowel durations are greatly underestimated" and (2) "in sentence-final position, vowel durations are overestimated in read declarative sentences; however, no overestimation is found in conversational sentences". These results are contradictory since all the sentence-final vowels are in prepausal position. Takeda et al. conclude that "there are different duration control mechanisms in sentence-final and prepausal position" (p.

2085) and in the sentence-final vowels of conversational sentences and those read.

Several questions arise from their report. What factor makes it possible for a prepausal vowel to lengthen or shorten in the sentential utterances? Does the conversational or reading speaking style really play a decisive role in the phenomenon? The effect of different speaking styles is doubtful since the "declarative" and "conversational" sentences they analyzed were all the readings of texts by a professional narrator. Their report does not show clearly to what extent his speaking styles differed. In addition, there is a possibility that the differences in the text materials, not in the speaking style, may result in the observed durational variation in prepausal position. In short, the factors contributing to the prepausal lengthening and sentence-final shortening in the sentential utterances have not been clarified.

1.4 Goals of this study

The foregoing sections have reviewed the past studies on Japanese intonation and final lengthening. Although some of the studies provide quantitative models to generate an F0 contour or segmental durations of an utterance, the duration and pitch of a sentence-final vowel and the factors that contribute to these two variables have not been clarified. Moreover, as has been pointed out, intonation and temporal structure have been explored separately, assuming no correlation between them. In other words, no detailed research has been performed on the relationship between pitch (or accent) and duration in Japanese prosody.

This study will explore in detail the relation between accent and duration, and clarify a consistent effect of accent on the duration of

prepausal vowels in Japanese. The series of phonetic experiments will make it clear that a prepausal vowel is lengthened when a prepausal (content) word has no accentual fall, whereas it is not the case when the prepausal word has an accentual fall. In addition, it will be clearly shown that no accent-related durational difference of a similar amount is observed in any other segment that consists of utterances.

These experimental results will lead to provide a prosodic control model that predicts the relative pitch and duration of prepausal vowels in sentence-final and non-sentence-final positions differentiated by the number and position of accentual falls in declarative sentences. Although this prosodic model does not predict absolute values of F0 and duration, it does describe a relative level of F0 and duration of prepausal vowels based on the significant differences observed in the present experimental data.

Moreover, the other shortening and lengthening phenomena observed in Japanese, such as phrase-final lengthening or lengthening of monomoraic nouns without particles will be examined one by one, showing that substantial shortening and lengthening are specific to prepausal position. From all these findings, it will be claimed that the prepausal position is outside of the mora-timing constraint in Japanese prosody, which accounts for a large durational variability restrictively observed in prepausal vowels. Finally, it will be suggested that Japanese has two distinct temporal structures: an invariable structure governed by the mora-timing hypothesis, and a variable and elastic structure free from any mora timing constraint.

The following chapter (Chapter 2) presents an overview of the temporal structure of Japanese. Chapter 3 highlights the strong effect of the absence or presence of an accentual fall on the duration of utterance-final vowels in isolated words and phrases. In Chapter

4, the same effect of accent will be examined in sentential utterances, with or without intervening pauses. These findings will support the proposal for a prosodic control model of sentential utterances in Chapter 5. Chapter 6 will claim that a wide durational variability in prepausal vowels indicates the markedness of the prepausal position in Japanese rhythmic structure. Finally, Chapter 7 addresses the theoretical implications of the elasticity of prepausal vowels in Japanese prosody.

---

[1] "F0" is used in the current study for the abbreviation of fundamental frequency that is acoustically measurable. The term "pitch" is used to more generally refer to the height of sound including its phonological and perceptual aspects. Thus, the phonological representation of word tone will be referred to as "low pitch" or "high pitch", while the acoustical control of voice height will be called "F0 control".

[2] The transcription of Japanese words used in this thesis mostly follows the orthography recognized by International Standard Organization (ISO), except that the long vowel is denoted by two successive vowels to make the representation of accentuation on each mora available (e.g., *oo, ee*). However, the names of authors, publishers, and places are mostly represented in the Hepburn system, or by their established styles, if any.

[3] Fujisaki has made not a few modifications in his quantitative model for generating F0 contours since his first proposal of the model in 1969, including the alternation of the definitions of the parameters. Thus, the description of his model in this thesis is mostly from his latest paper in 2002.

[4] According to his theory, the 'baseline' refers to the asymptotic F0 value to which the phrase and accent components approach.

[5] Autosegmental accounts have associated the H tone with every tone-bearing unit that succeeds an initial H tone in a phrase until the L of an accent occurs, and have called it "spreading" of H tone.

CHAPTER 2

Temporal Structure of Japanese

2.1 Introduction

This chapter provides an overview of the temporal structure of Japanese. More specifically, it explores the weaker version of mora timing hypotheses of Japanese: mora timing at word level and little or no effect of accent and position on mora duration. As has been discussed in Chapter 1, previous studies have shown many phonetic features of Japanese temporal structure, some through laboratory experiments, some through corpus analyses. However, there exists a considerable inconsistency in their results that may partly result from differences in methodology. The aim of the present experiments is to offer solutions to the problems of these contradictory findings in order to better explain the temporal structure of Japanese.

As has been introduced in 1.1.2, isochrony at word level has been claimed by researchers. Port, et al. (1987) have analyzed the duration of Japanese words embedded in a carrier sentence and assert that "[Japanese] words with an increasing number of moras increase in duration by nearly constant increments" and that "word durations should come in steps that are integral multiples of the

duration of a mora at a particular speaking tempo"(pp.1574-1575). Sugitô (1975) has also measured duration of moras in 549 Japanese words uttered in isolation by one speaker of the Tokyo dialect. She reports that word length increases almost linearly with an increasing number of moras in the word. In other words, these phoneticians assert that the mean duration of moras that constitute words is almost constant in Japanese.

Satô (1977a,b) has analyzed the duration of moras uttered in isolation by a pair of male and female subjects. His data indicate, however, that the mean duration of moras decreases as the number of moras in a word increases, until a word size reaches six moras. Oyakawa (1971) has analyzed utterances varying in length between the two-mora word *aka* and eleven-mora clause *akasa'dake'datomi'setemo,* produced by one subject. From the results he draws the conclusion that there is a tendency that "in the utterances of up to five syllables segments become shorter in duration as the utterances become longer"(p.31). Consequently, he asserts that "the rate of speech and the length of utterance are two distinctively different factors" (p.33) that contribute to the duration of constituent segments. Campbell (1992a) has observed a similar relation between mean mora length and the size of 'breath groups'. His data, consisting of paragraphs read by a female announcer, demonstrate "the average mora length to be greater in shorter breath groups"(p.214). However, he admits that the number of shorter breath groups is too small for the trend to be statistically significant, since "the majority of mora are in longer breath groups"(*Ibid.*). These findings clearly contradict the uniform reports that indicate invariability in mean mora duration at word level in Japanese.

Thus, there appears to be a considerable discrepancy between

the timing of Japanese words uttered in isolation and those embedded in a carrier sentence, though the former has not yet been investigated thoroughly enough to provide data comparable with those of the latter. Thus, Experiment 1 explores the temporal structure of isolated Japanese words in contrast with that of the same words embedded in a carrier sentence and examines the mora-timing hypothesis at word level.

The other questions addressed herein are to what extent Japanese moras are equal in duration if segmental variation is eliminated and whether or not position and accent have no effect on the temporal structure of Japanese. Hoequist (1983a,b) reports significantly less lengthening under accent in Japanese compared with that found in stress- and syllable-timed languages. Kaiki et al. (1992) also assert that accent does not affect duration of Japanese syllables.

As with final lengthening, Takeda et al. (1989) report consistent lengthening in prepausal position, while pointing out shortening in the sentence-final position, a phenomenon contradictory to the prepausal lengthening. Their report gives rise to the question of what factors contribute to the occurrences and/or lack of the final lengthening in utterances. Thus, what is examined in Experiment 2 is the effect of position and accent on the temporal structure of Japanese and isochrony of moras by employing the reiterant speech that eliminates segmental variation by replacing syllables with 'ma'.

2.2 Experiment 1

This experiment was conducted to examine the durations of constituent moras (CV syllables) and segments in isolated

words/phrases as a function of the word length, in contrast with those embedded in a carrier sentence.

2.2.1 *Subjects*

The subjects were two middle-aged males (M, N), and three middle-aged females (A, B, E), all native speakers of Tokyo (standard) Japanese[1].

2.2.2 *Materials*

Four monomoraic nouns *na, ma, su, si* were selected to form four sets of test words or phrases. The initial consonants /n, m, s/ were all measurable at utterance-initial position.

The four sets of test items were chosen so as to contain a variety of segments. As a rule, each test word/phrase in each set was constructed by adding a new mora, a light syllable (CV), to the end of the previous one. Syllables constituting the test words/phrases are all light syllables composed of one consonant and one short vowel (CV), ranging from one to seven moras in length. Test items of one to four moras are real Japanese words or compound words. Test items composed of five to seven moras contain real phrases, or phrases that native speakers may easily understand. There was a total of 23 test words/phrases as presented in Appendix 2-1.

Two kinds of lists were prepared. In one list, test words/phrases were embedded in a carrier sentence '*Moo itido* ___ *to itte kudasai*' "Please say __ again". The carrier sentence is the same as the one used in Hoequist (1983b). In the other list, only isolated test words/phrases were written. Both kinds of lists were written in Japanese orthography (*kanji and kana*).

2.2.3 *Procedure*

First, the lists of the carrier sentences containing the test items were presented to each subject, who was asked to familiarize him/herself with all the sentences on the lists. Then each list was presented in random order and read seven times at a subjectively established constant rate. Each speaker was instructed not to insert a pause within a sentence. Next, the lists of the isolated test words/phrases were presented to each subject and read seven times at a self-selected constant tempo.

Subjects' readings were recorded in a sound-proof room or a room quiet enough for recording and digitized at 11-kHz sampling rate on a computer using a Windows 98 operating system. Waveforms and wide-band spectrograms were made of every reading with a Kay Multi-Speech Model 3700. As a rule, the second to the seventh productions of each item were measured. Thus, 6 tokens were obtained for every test word/phrase from each subject. The total number of tokens measured was (23 words/phrases $\times$ 6 repetitions $\times$ 2 conditions $\times$ 5 subjects =) 1380.

Measurements were obtained by using manual cursors for durations of (1) each segment measurable of all the test items; (2) the VOTs for voiced and voiceless plosives and for flap /r/ when they were visible on spectrograms; (3) the entire utterance. However, the VOTs for voiced plosives and flap /r/ were not distinctive enough to be measured objectively in some of the tokens from some speakers. Thus, the aspiration (VOT) for voiced stops and flap /r/ were included in the following vowel when processing the data. Also in the case of the syllable /ya/ in the *su*-set, duration measurement was made of the entire syllable since the segment boundary between /y/ and /a/ could not be objectively determined.

The criteria for determining segment boundaries were dependent upon the particular phonetic features of the segments (cf. Peterson and Lehiste, 1960; Beckman, 1982).

The beginning of nasals /n/ and /m/ was taken at the start of nasal murmur, where the nasal formant at low frequency appears. The boundary between a nasal and the following vowel was defined at the point where the formants of the vowel appear. The start of word-initial fricatives /s/ /ʃ/ was determined to be the starting point of high frequency spectral energy. The period of fricatives were in most cases identical with that of imperiodical waveforms of noise. When the boundary between fricative consonant /s/ or /ʃ/ and the following vowel was not distinct enough to be determined on waveforms, the terminal point of the voiceless fricative was defined to be the point at which the formants of the vowel appear, whether frication has ceased or not. Similarly, VOTs of voiceless stops were measured until the point where the formants of the following vowel appear. When /r/ and /g/ were not produced as a stop, the boundary between them and the following vowels were determined at the point where the second and/or third formants of the vowels appeared.

The end of each utterance was defined as the point at which periodical waveforms of vowels end. Vocal fry register pulses at utterance-final position were included only in cases where they occur within 15 ms from the end of the preceding striation glottal pulse.

All the segmental durations were repeatedly measured to ascertain the accuracy and consistency of the measurements.

2.3 Results & discussion

2.3.1 *Mora timing in a frame sentence*

Figure 2-1 represents mean mora duration averaged across subjects for each test set embedded in a carrier sentence as a function of test word/phrase length in moras. The mean mora duration is obtained by dividing word/phrase duration by the number of moras contained.

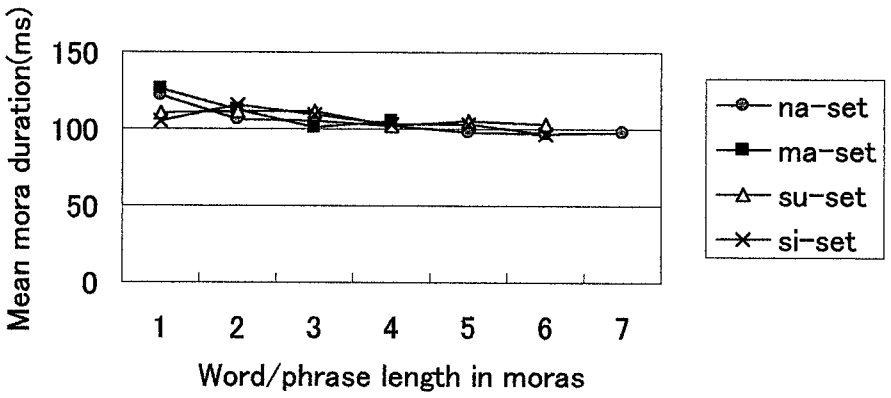


Fig. 2-1. Mean mora duration (ms) of words/phrases set in a frame sentence

As is obvious from the figure, mean mora durations do not differ widely among four sets of items despite variation in their segmental make-ups. Moreover, they do not show great fluctuations as word/phrase length increases in each set. All the mean durations in the four sets vary within a relatively small range from 97.2 to 126.3 ms.. Only noticeable difference in mean duration is observed between the durations of monomoraic words /na/ and /ma/ and those of the others. There is a possibility that the monomoraic words may be lengthened due to the bimoraic minimality constraint formulated on Japanese prosodic words which is discussed in depth in 6.2. This difference seems to be

leveled out when a word contains more than one mora as in the other test items, suggesting the claimed isochrony in moras at word level.

2.3.2 *Effect of utterance length on mora duration*

The mean mora duration in isolated items averaged across speakers is presented in Figure 2-2. The most obvious finding from the figure is that the mean mora duration is much longer in monomoraic words and it gradually decreases as the number of component moras increases up to about four moras. There is a high negative correlation between the mean mora duration and the number of moras contained in items in all the four sets (–0.80 to –0.94, $p<0.0001$).

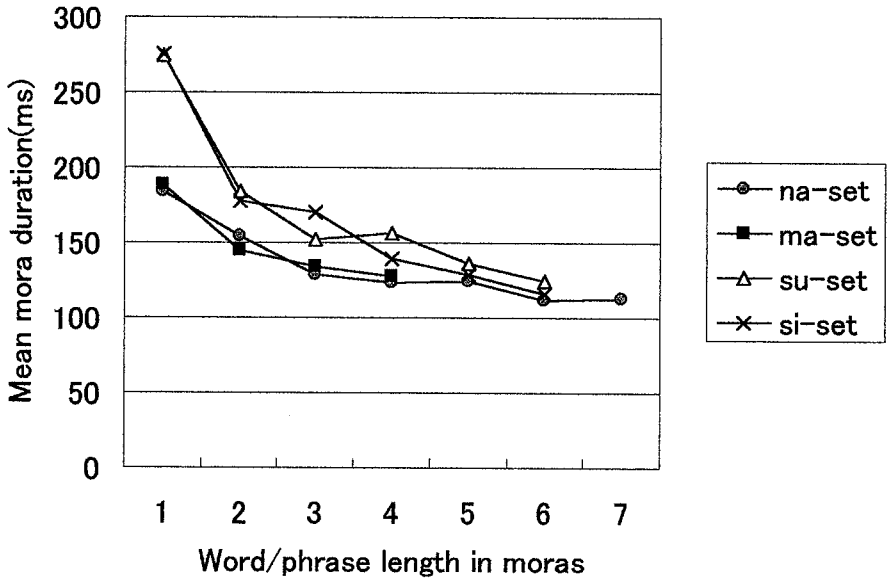


Fig. 2-2. Mean mora duration (ms) of words/phrases uttered in isolation

The difference in mean duration between one- and three-mora words is about 54 to 55ms for *na*- and *ma*-sets. The mean durations in *su*- and *si*-sets exhibit a sharper decrease of 106.2 to

122.8ms between monomoraic words and trimoraic ones. The discrepancy in the degree of shortening among the sets seems to be a reflection of the difference in compressibility or 'elasticity'[2] of the segments constituting the test words. Specifically, the more elastic segment /s/ is assumed to be lengthened and shortened to a greater degree than the less elastic segments /m,n/.

In addition, the four sets show a large variation in mean mora duration even with the same number of moras. The mean mora duration of the isolated bimoraic words, for instance, ranges from 144.4 ms for *ma'do* to 183.8 ms for *sumi'*. The discrepancy again appears to be brought about by differences in segmental content of the words: different inherent durations and/or elasticity or compressibility of the constituent segments.

This great variation in mean mora duration makes it clear that isolated words with the same number of moras vary considerably in duration and that they do not increase in duration by constant increments with the number of moras involved. Thus, the mora-timing model that claims isochrony in moras at word level has turned out to be applicable only to words embedded in a carrier sentence.

2.3.3 *Effect of utterance length on segmental duration*

The next question to be answered may be how the durations of the component segments change as the word gains an increasing number of moras. Figures 2-3A to 2-3H represent the duration of each constituent segment in each isolated test word/phrase averaged across speakers as a function of word/phrase length in moras. The plots in Figs. 2-3A to 2-3H are, however, restricted to the first-to-eighth segments in each set so as to show their duration change with several additional moras.

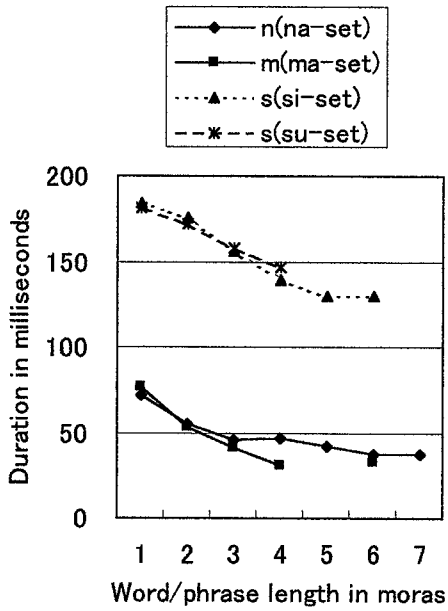


Fig.2-3A Mean duration of 1st segments

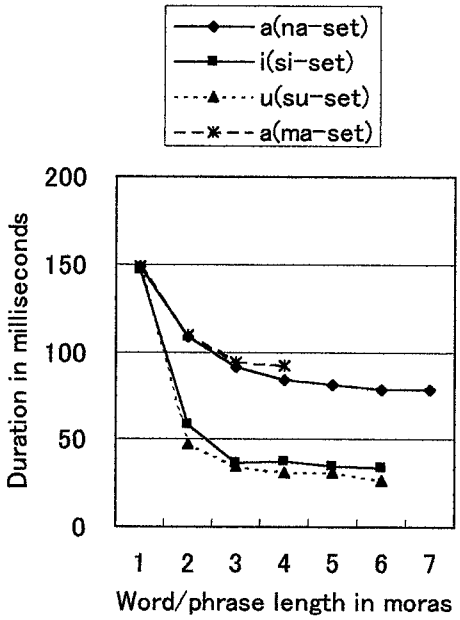


Fig.2-3B Mean duration of 2nd segments

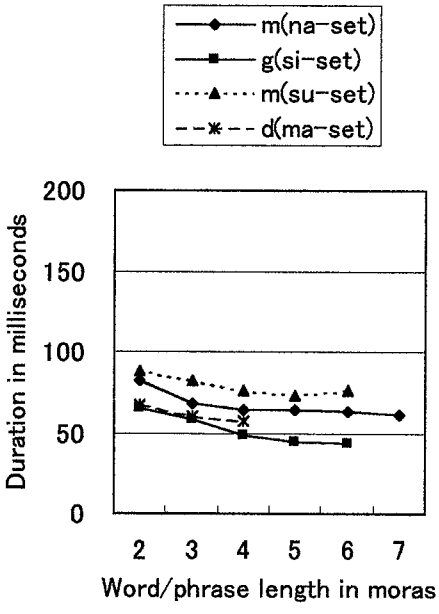


Fig.2-3C. Mean duration of 3rd segment

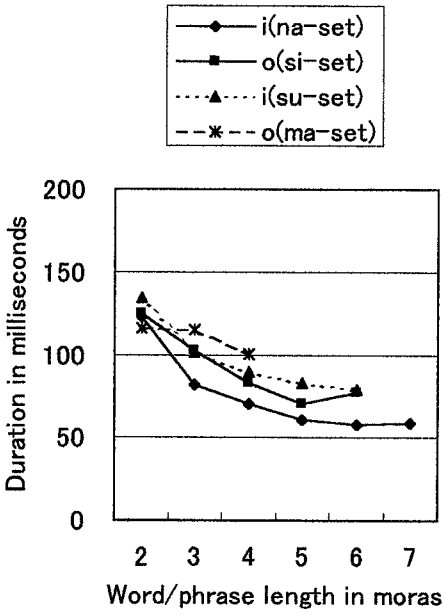


Fig.2-3D. Mean duration of 4th segment

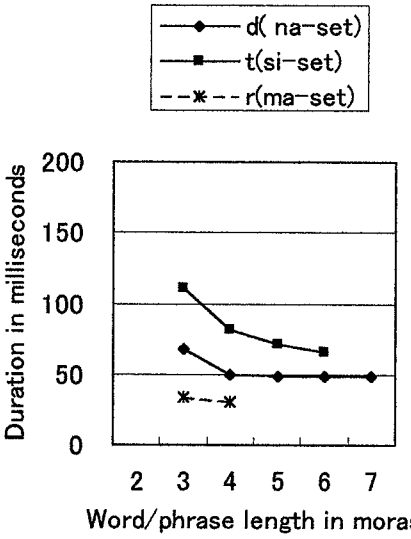


Fig.2-3E. Mean duration of 5th segments

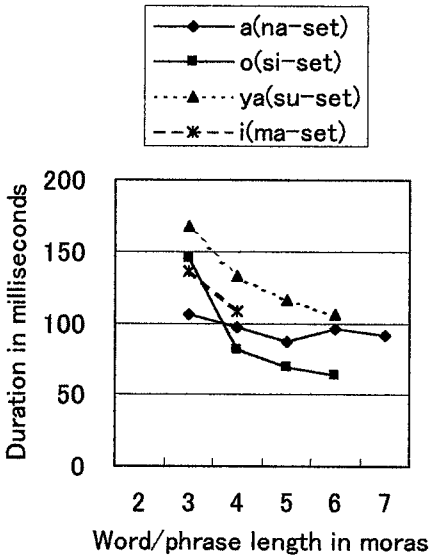


Fig.2-3F. Mean duration of 6th segments

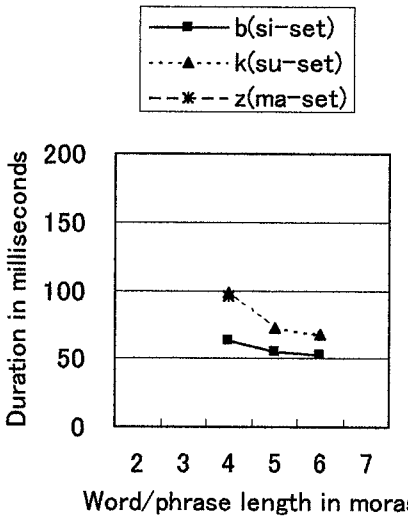


Fig.2-3G. Mean duration of 7th segments

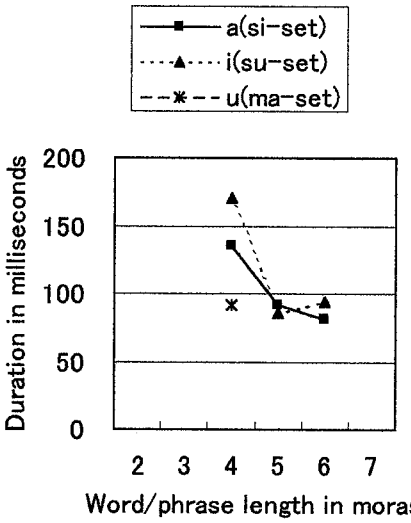


Fig.2-3H. Mean duration of 8th segments

Fig.2-3A exhibits a large difference in mean duration between consonants /s/ and /n, m/, where /s/ is more than 100 ms longer than /m,n/. The discrepancy reflects a difference in inherent duration or elasticity between /s/ and /m,n/ as mentioned above. It can be seen, however, that all the consonants are shortened as the utterance length increases to about four moras. Then, their durations stay almost constant in the longer items.

The vowels /a,i,u/ in the first mora consist of almost the same duration (about 150 ms) in monomoraic words. Then, they are shortened as they are followed by one to three moras, apparently approaching their respective average durations in Japanese reported by Campbell (1992b).

Fig.2-3C indicates that all the consonant duration decreases gradually in utterances of up to four moras. The following vowels in Fig.2-3D, by contrast, show clear shortening in two- to five-mora items. Taken together, Figs. 2-3A to 2-3H suggest that vowels are more stretchable and compressible than consonants.

Figs. 2-3E and 2-3F indicate that the segments in the third mora shrink until the number of moras in an utterance reaches four or five. However, the consonant /r/ of the *ma*-set in Fig. 2-3E shows hardly any shortening. The less elastic segment /r/ is assumed to be lengthened or shortened much less than the more elastic segments /t,d/ (cf. Port et al., 1980; Campbell, 1992a).

The duration of segments in the fourth moras are presented in Figs. 2-3G and 2-3H, in which the vowels largely decrease in duration as the utterance length increases from four to five moras. The 7th and 8th segments are, however, in the word-final position of four-mora test words, and subject to final lengthening (Takeda et al., 1989). Therefore, they are supposed to sharply decrease in duration in the penultimate position of five-mora items.

It can be concluded that component segments in a word produced in isolation tend to be shortened until the word length reaches four or five moras. The only exception to this observation is that the segments of utterance-final syllables tend to gain extra length.

The lengthening observed in all of the segments in utterances of up to four or five moras indicates that the longer mean mora duration in words of one to four moras seen in Fig. 2-2 is not brought about solely by final lengthening taking place in the prepausal syllable. Instead, it can be attributed to the lengthening of all component segments.

These findings have clarified that utterance length is one of the crucial factors that affect the duration of segments or syllables constituting an utterance. The phenomenon is consistent with the effect of mora counts in a breath group on segmental durations that has been pointed out by Sagisaka et al. (1984) and Takeda et al. (1989) based on statistical analyses of a corpus produced by a male professional narrator.

2.3.4 *Mora timing in longer utterances*

It can be seen from Fig. 2-2 that the mean mora duration stops decreasing sharply when the word or phrase size reaches four or five moras in each set. Moreover, items with the same number of moras tend to approach the similar length when the items extends beyond four moras. The mean duration ranges from 111.5 to 124.7 ms in items composed of six or seven moras. Although there still exists a small amount of variation among the four sets, the mean mora duration appears to show much less variability among the sets when word/phrase length exceeds four or five moras.

Campbell (1992a) has reported mean mora duration in much

larger 'breath groups' of up to about 25 moras based on his data from readings of paragraphs. The mean mora duration of his data exhibits variation from 110 to 140 ms in six-mora utterances. Then, it gradually decreases to be about 120 ms on average in 'breath groups' of about fifteen moras. Further in longer utterances the mean mora duration stays almost constant.

All of the above findings indicate that the mean mora duration in utterances of more than approximately four moras varies within a relatively small range from 110 to 140 ms. That is to say, the data suggest that utterances longer than about four moras increase in duration by roughly constant increments for each additional mora. In other words, utterance length increases almost proportionally to the number of component moras. Thus, these results suggest that isochrony in Japanese moras tends to be maintained at the threshold of utterances consisting approximately of more than four moras.

2.4 Experiment 2

The purpose of this test was to examine the effect of accent and position on duration of moras and to clarify the timing of moras in isolated words, phrases, and sentences in Japanese. In order to eliminate the effect of intrinsic and neighboring segmental duration, each subject was instructed to utter reiterant speech that replaced the real items in Experiment 1 with repetitions of the nonsense syllable 'ma'.

The validity of reiterant speech was supported by the view that in normal speech two kinds of factors simultaneously contribute to the temporal structure of speech (cf. Nakatani, O'Connor & Aston, 1981; Hoequist, 1983b). One is a segmental or articulatory

influence. The other is the effect of prosodic features such as accent or rhythm on units larger than segments. Thus, repetitions of the consonant /m/ and vowel /a/ eliminate the influence of inherent segmental duration and adjacent segments on the timing of moras in the test items while their prosodic features are assumed to be preserved.

2.4.1 *Subjects*

The subjects were the same speakers of Tokyo Japanese who had participated in Experiment 1.

2.4.2 *Materials*

The test items were the same as those employed in Experiment 1. In addition, three test sentences were prepared in order to compare the timing pattern of sentences with that of words or phrases (Appendix 2-1). All sentences were composed of five moras with the monomoraic word /na/ or /ma/ sentence-initially. The sentences differ in syntactic/morphological structure or in accent pattern. A syntactic boundary between subject and predicate is located after particle /ga/ in sentences *naga* || *sireru* "The name is known." and *maga* || *no'bita* "A pause is/was prolonged.", whereas a sentence *manobisita* "Too long of a pause is/was inserted." has no corresponding syntactic boundary. Morphologically, all of these sentences can be segmented as three or four morphemes, though the locations of the boundaries are different among them (*na* | *ga* || *sireru, ma* | *ga* || *no'bi* | *ta, ma* | *nobi* | *si* | *ta*). (The vertical lines || and | denote syntactic and/or morphological boundaries, respectively.)

The four sets of test words/phrases and the three sentences were all written on sheets of paper in Japanese orthography (*kanji*

and *kana*).

This experiment was carried out in succession to Experiment 1 with each subject. Each subject was instructed to first read each item as it was, and then to replace each of the moras with the syllable 'ma' just after each normal utterance of a word, phrase, or sentence. Subjects were also told that 'ma' was to be assigned to each mora.

However, some of the reiterant tokens produced by the subjects turned out to be different from the original item in the number of moras or in the location of accent. Whenever such differences were noticed, the subject was asked to do the item again. At least three tokens of reiterant speech for each test item by each subject were obtained.

2.4.3 *Procedure*

Subjects' readings were recorded in a sound-proof room or a room quiet enough for recording. Their readings were digitized at 11-kHz sampling rate on a Windows 98 computer. Waveforms and wide-band spectrograms were made from every reading with a Kay Multi-Speech Model 3700.

One of the benefits in using reiterant speech is that it makes segment measurements easier and more reliable by employing a single set of criteria for segmentation. Measurements were made of the duration of each nasal /m/ and vowel /a/. The beginning of the nasal /m/ was taken at the start of the nasal murmur where the nasal formant at low frequency appears. The boundary between the nasal and the following vowel was defined at the point where the formants of the vowel appear. The end of each utterance was determined by the same criteria as those employed in Experiment 1.

2.5 Results & discussion

2.5.1 *Effect of position on mora duration*

Figures 2-4A to 2-4D represent the effect of position on mora duration averaged across speakers for each set of test items replaced by 'ma'. Although the test words/phrases are presented

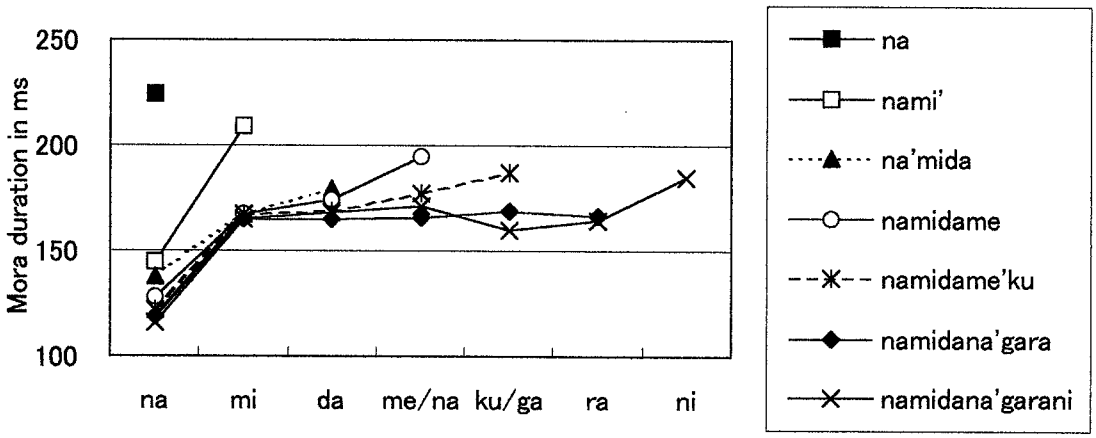


Fig.2-4A. Effect of position on mora duration in isolated reiterant speech for na-set

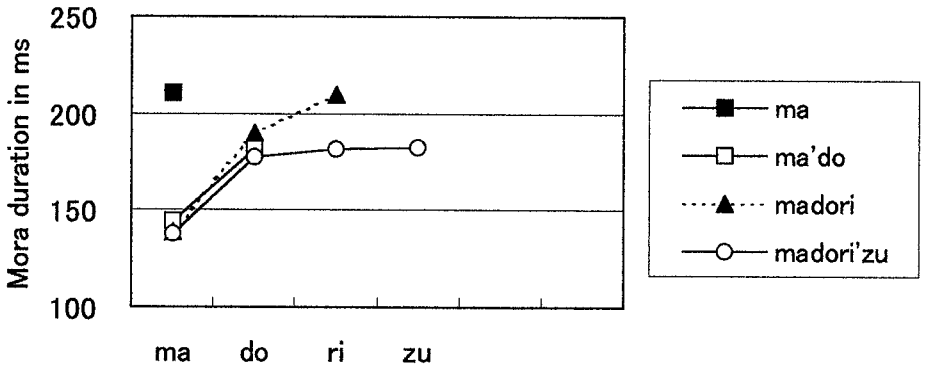


Fig.2-4B. Effect of position on mora duration in isolated reiterant words (ma-set)

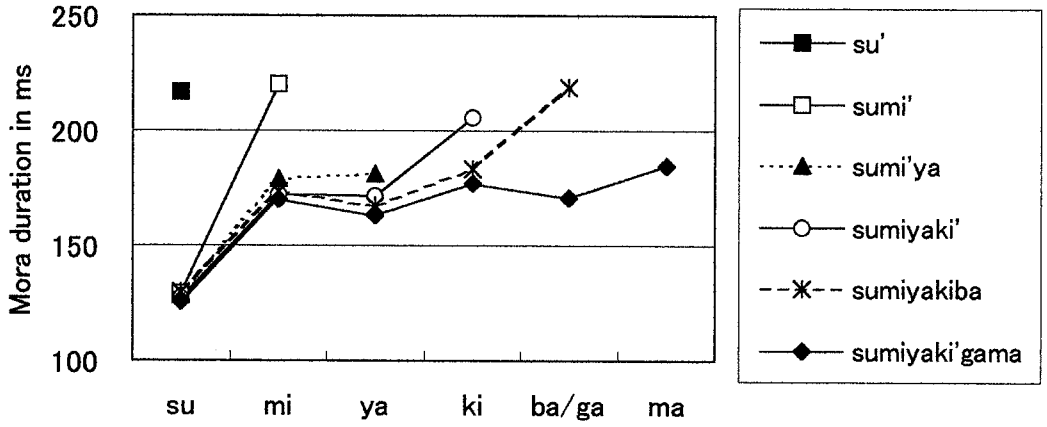


Fig.2-4C. Effect of position on mora duration in isolated reiterant speech for su-set

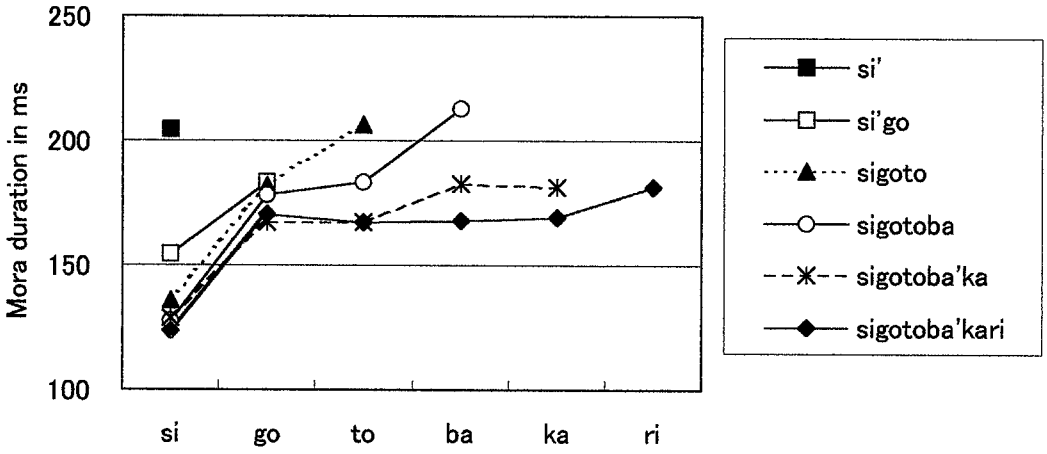


Fig.2-4D. Effect of position on mora duration in isolated reiterant speech for si-set

in the figures using the original real words, the actual tokens were all composed of the syllable /ma/.

Let us first make a comparison of mora lengths in different positions in an utterance. From the figures, it can easily be seen that the duration of a monomoraic word replaced by 'ma' is the

longest mora or as long as the final mora of the other items in each set. When a monomoraic word is followed by one or more moras, however, the initial mora becomes the shortest of all the moras of each test word/phrase. The difference between the first and second moras in each group ranges from 32.7 to 75.9 ms. This discrepancy in duration is consistently observed regardless of accent patterns and utterance length in all four sets of the test items.

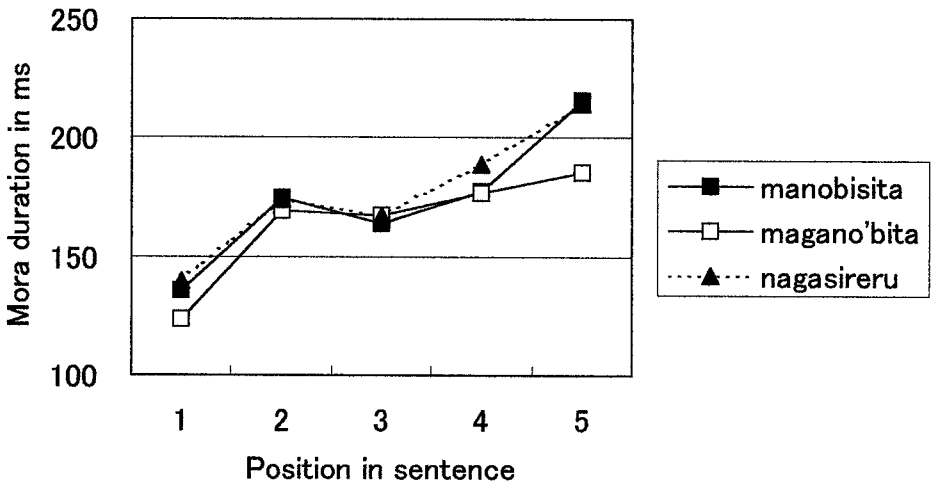


Fig. 2-5. Effect of position on mora duration in isolated reiterant sentences

Fig. 2-5 shows mean mora duration in each position of the three test sentences replaced with 'ma'. Similarly, the figure indicates that the initial moras are the shortest of all the component moras in each sentence.

The shortest duration consistently observed in the initial moras gives support to the validity of the employment of reiterant speech. The original initial moras of the test words contain /si/ and /su/ whose average durations in Japanese are much longer than the initial moras /ma/ and /na/ or most of the other moras constituting the test words/phrases (cf. Campbell, 1992b).

Nevertheless, the tokens that replaced the real words with the syllable /ma/ consistently have much shorter initial moras. That is, the differences in segmental duration of the original test words/phrases do not appear to affect the timing of the reiterant tokens, while the suprasegmental variation in duration depending on position is consistently observed.

Another finding from Figs. 2-4 and 2-5 is that the final moras of reiterant speech without an accentual fall tend to exhibit a large increase in duration from the other component moras. The final moras of reiterant speech with an accentual fall, by contrast, appear to maintain more or less a length similar to their preceding moras. In isolated utterances, unaccented or finally-accented words/phrases are assumed to have no accentual fall within the utterances, while they generally have an accentual fall when they contain initially- or medially-accented words.[3]

Table 2-1 presents mean duration (ms) of moras in each position of words/phrases in each group. As has been clarified in Experiment 1, the mora duration is greatly influenced by the size of an utterance. Accordingly, all of the reiterant tokens composed of one to seven moras are first classified by the number of moras contained. Next, the words/phrases with the same number of moras are further divided into those with an accentual fall and those without.

A one-way analysis of variance of repeated measures[4] (RM-ANOVA) was conducted by a SAS StatView[5] for each group with position as a fixed factor, showing the effect to be highly significant ($p<0.0001$). Post-hoc tests (Fisher's PLSD[6]) indicated that all of the durations of the first moras are significantly shorter than all other mora durations for each group ($p<0.0001$).

Table 2-2 presents mean mora duration for each test sentence

Table 2-1. Mean duration (ms) of moras and SD (bottom line)

No. of moras	Accentual fall	No. of tokens	1st mora	2nd mora	3rd mora	4th mora	5th mora	6th mora	7th mora
1	Without an accentual fall	62	**214.0** 33.3						
2	With an accentual fall	31	**149.5** *** 29.4	**182.2** 23.5					
2	Without an accentual fall	31	**138.3** *** 29.0	**214.3** 20.4					
3	With an accentual fall	34	**134.7** *** 25.5	**172.6** 14.6	**179.3** 18.1				
3	Without an accentual fall	28	**138.3** *** 23.7	**185.6** *** 13.0	**208.2** *** 16.8				
4	With an accentual fall	15	**138.2** *** 18.4	**176.8** 12.2	**181.4** 12.1	**182.4** 23.1			
4	Without an accentual fall	43	**129.1** *** 26.8	**172.5** 15.5	**176.0** 16.6	**203.4** *** 21.5			
5	With an accentual fall	29	**126.0** *** 31.9	**167.2** 17.2	**168.1** 17.2	**180.0** 21.8	**183.9** 17.2		
5	Without an accentual fall	16	**129.9** *** 23.9	**173.8** 13.1	**167.2** 10.9	**184.1** 20.1	**218.1** *** 26.2		
6	With an accentual fall	42	**123.0** *** 20.5	**167.9** 18.2	**164.9** 15.7	**170.5** 15.0	**169.4** 15.6	**178.1** * 22.7	
7	With an accentual fall	14	**116.7** *** 25.7	**165.2** 18.5	**167.7** 12.0	**170.9** 16.8	**159.7** 17.6	**164.6** 14.5	**184.1** * 20.8

*** and * indicate significant differences at the levels of $p<0.0001$, and $p<0.05$, respectively, from all other moras in the same word/phrase.

Table 2-2. Mean mora duration (ms) and SD (bottom line) of reiterant speech for sentences

Sentences replaced by 'ma'	No. of tokens	1st mora	2nd mora	3rd mora	4th mora	5th mora	RM-ANOVA
nagasireru	16	140.3 ***	173.7ns	167.6	188.9 *	213.9 ***	F(4,60)=46.72
SD		29.2	17.7	18.1	21.9	21	p<0.0001
magano'bita	15	123.8 ***	168.9ns	167.4	176.7ns	185.1	F(4,56)=30.21
SD		25.8	14.9	12.2	12.3	19	p<0.0001
manobisita	15	137.2 ***	174.2	163.8	176.6	215.6 ***	F(4,56)=33.61
SD		35.4	12.3	13.5	14.6	19.6	p<0.0001

*** and * indicate significant differences at the levels of p<0.001 and p<0.05. respectively, from all other mora durations in each sentence.

replaced with /ma/. A one-way RM-ANOVA with position as a fixed factor for each sentence showed a highly significant effect of position. Post-hoc tests indicated that all of the initial moras are significantly shorter than the other moras in the same sentence.

This experimental result of the shortest initial mora in an utterance is mostly consistent with the report by Sagisaka et al. (1984) based on a large amount of corpus analyses of varied segments. They report that utterance-initial consonants and vowels are on average shorter by about 20% than the following segments.

2.5.2 *Effect of accent on accented mora*

The results obtained from isolated reiterant speech have revealed that the duration of accented moras is not significantly longer than that of unaccented moras. In Table 2-3, comparisons are made between accented and unaccented moras in the same position of reiterant speech that contain the same number of moras, all

showing no significant difference between them. Moreover, the mean duration of accented vowels is significantly shorter than that of corresponding unaccented vowels, both in the initial mora of three-mora words and in the fourth mora of five-mora words ($p<0.05$).

These results are roughly consistent with earlier observations that the influence of accent on the duration of accented moras in Japanese is secondary or nonsignificant (Mitsuya & Sugitô, 1978; Hoequist, 1983b; Kaiki et al., 1992).

Table 2-3. Mean duration (ms) of accented vs. unaccented moras and vowels in reiterant speech

No. of moras	Position of moras	No. of tokens	Accented moras (SD)	Accented vowels (SD)	No. of tokens	Unaccented moras (SD)	Unaccented vowels (SD)
2	1st	31	149.5 (29.4)	100.9 (6.5)	31	138.3 (29.0)	96.6 (11.9)
3	1st	18	139.0 (27.6)	* 87.8 (10.9)	28	138.3 (23.7)	* 96.3 (10.2)
3	2nd	15	178.3 (11.4)	103.1 (11.5)	28	185.6 (13.0)	110.3 (11.4)
4	3rd	15	181.4 (12.1)	104.4 (7.8)	43	176.0 (16.6)	98.8 (14.3)
5	4th	29	180.0 (21.8)	* 99.0 (14.8)	16	184.1 (20.1)	* 110.5 (13.4)

* indicates significant differences (ANOVA) at the level of $p<0.05$ between the two corresponding accented and unaccented durations. Those without * showed no significant difference.

2.5.3 *Effect of accent on final moras*

As has been pointed out, Figures 2-4 and 2-5 tell us that the utterance-final mora of a word/phrase or sentence without an accentual fall tends to increase in duration, whereas its counterpart with an accentual fall does not. The final position in a word/phrase or sentence in this experiment coincides with the

utterance-final, or prepausal position.

One-way ANOVAs were conducted in order to compare the durations of final moras of words/phrases "With an accentual fall" with those "Without an accentual fall" for each of two- to five-mora groups in Table 2-1. They showed highly significant differences between each pair of the groups with the same number of moras ($p<0.01$).

Next, a repeated measures ANOVA was conducted on durations of moras of each group with position as a fixed factor. The asterisks *** and * in the table indicate that the duration of the final mora is significantly different from all the other moras in the same group at the levels of $p<0.0001$ and $p<0.05$, respectively. As is obvious from Table 2-1, when test words/phrases consist of three to five moras and have no accentual fall, the final moras are significantly longer than any other mora in the same group of words/phrases. When test items have an accentual fall, by contrast, they show no significant lengthening in the final mora.

In this dissertation, 'final lengthening' is defined as durational increase observed in the final syllables/moras/segments that leads to a significant durational difference from their preceding counterparts. Accordingly, it can be said that the words/phrases with no accentual fall undergo final lengthening in utterance final position, whereas those with an accentual fall do not.

When the test item is composed of six moras, the final mora shows significant increase in duration even if they have an accentual fall (Table 2-1). However, the difference between the final and penultimate moras is only 8.7ms and statistically significant at the level of $p<0.05$. Thus, it can be said that the final lengthening observed in six-mora reiterant speech with an accentual fall is different in degree from that found in reiterant

speech of three to five moras without an accentual fall. The temporal structure of six-mora words is discussed in greater detail later in 2.5.5.

The reiterant speech for the seven-mora word *namida | na'gara | ni* also displays the phenomenon of a final mora being significantly longer than any other component moras. It may be noted that this item is the only adverbial containing two particles (/nagara/ and /ni/) in addition to the noun *na'mida*. It is suggested that this word has a temporal pattern similar to the six-mora *namidana'gara* except for the final particle, which seems to be subject to final lengthening. The final lengthening observed in the monomoraic particle is suggestive of a tendency for a final monomoraic suffix to be lengthened in prepausal position, which is discussed in detail later in Chapter 6.

The effect of an accentual fall on the duration of final moras is not restricted to words or phrases. Table 2-2 shows that the reiterant utterances for *naga sireru* and *manobisita* produce sentence-final moras significantly longer than any other preceding mora ($p<0.0001$). The reiterant speech for *maga no'bita*, by contrast, exhibits no significant lengthening in the duration of the sentence-final mora. Furthermore, the sentence-final moras of the two sentences without an accentual fall are significantly longer than that of the sentence *maga no'bita* with an accentual fall ($p<0.0001$).

It can be concluded that in reiterant speech for isolated words/phrases and sentences ranging from two to five moras, the final moras in utterances without an accentual fall are lengthened by a much larger degree than their counterparts with an accentual fall, showing highly significant differences. In addition, the final moras of utterances with no accentual pitch drop are significantly

longer than their preceding moras, while those with an accentual fall are not. Thus, the results suggest that accent has a crucial influence on the occurrence of utterance-final lengthening in words, phrases and sentences in isolation.

The next question to be addressed is which portions of syllables (consonants or vowels) of the test items are lengthened when they have no accentual fall. Figures 2-6 to 2-9 present mean duration of consonants and vowels in each position of two- to five-mora reiterant speech, respectively. The error bar in the figures shows a standard error above or below the mean. Test words/phrases are classified by the location of accent nuclei in the figures. For instance, "1" indicates words/phrases with an accent nuclei on the first mora, while "0" represents those with no accent nuclei.

The first finding from the figures is that all of the initial consonants are much shorter than all of the following moras in the test words/phrases of each number of moras, whereas the following consonants exhibit neither a large increase nor decrease in duration.

A two-way RM-ANOVA was run on the duration of consonants of the test items with each number of moras, respectively, with accent pattern and position as two fixed factors. It showed a highly significant effect of position. There was no significant effect of accent pattern nor interaction between the factors. Multiple Scheffé comparisons indicated the first consonants were significantly longer than all following consonants ($p<0.001$). The results make it clear that the presence/absence of an accentual fall has no effect on the duration of consonants and that the consonants gain no extra length in the final position of an utterance.

The final vowels, by contrast, show a sharp increase in

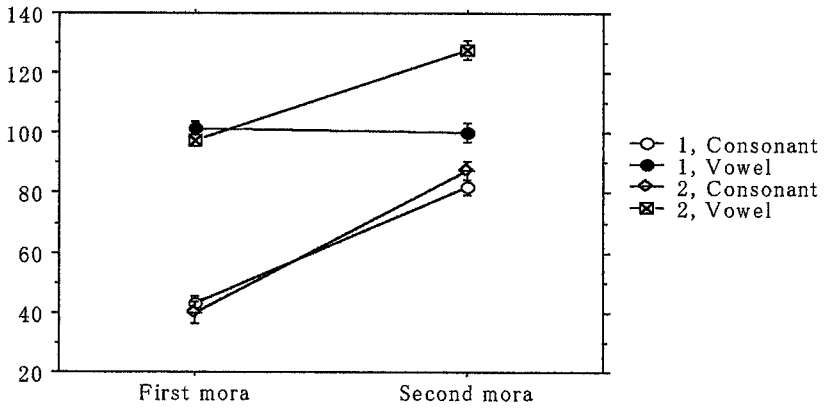


Fig. 2-6. Effect of position and accent on duration (ms) of consonants and vowels (two-mora words)

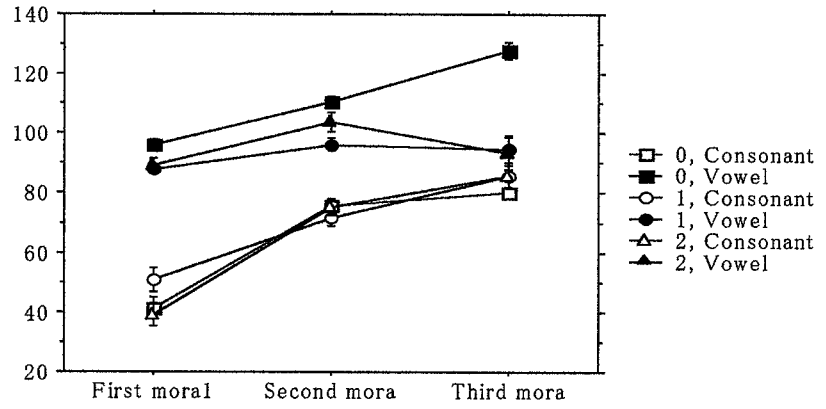


Fig. 2-7. Effect of position and accent on duration (ms) of consonants and vowels (three-mora words)

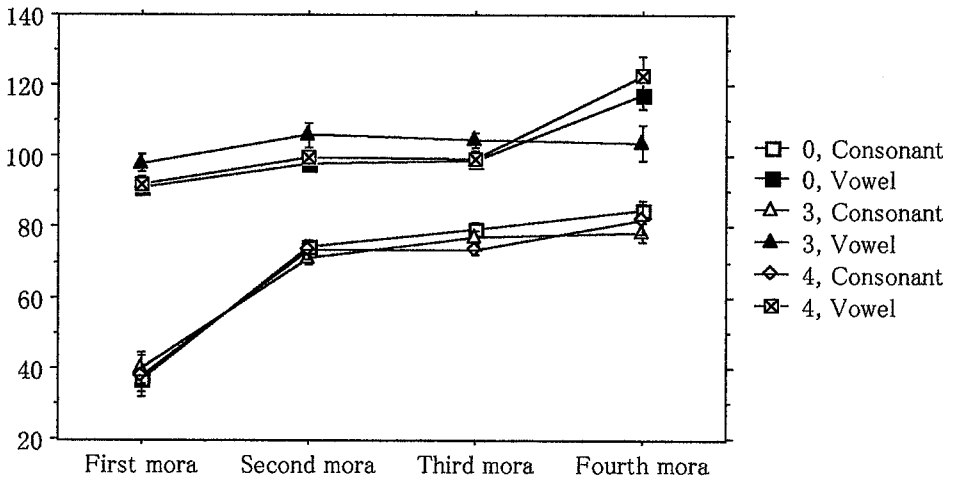


Fig. 2-8. Effect of position and accent on duration (ms) of consonants and vowels (four-mora words)

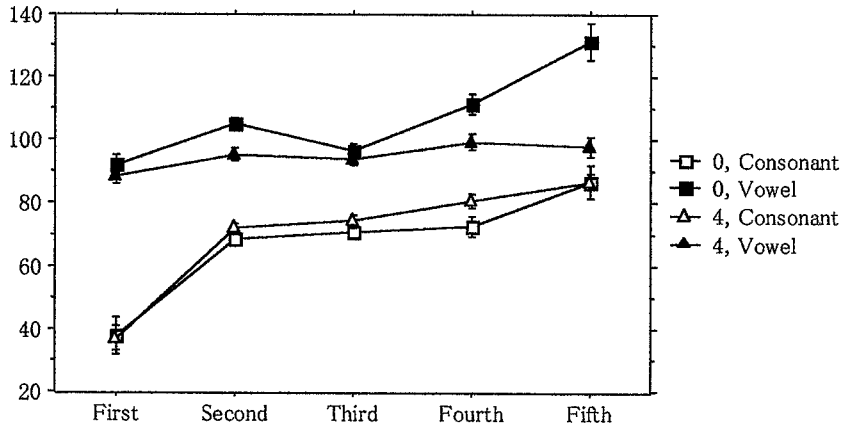


Fig. 2-9. Effect of position and accent on duration (ms) of consonants and vowels (five-mora words)

duration utterance-finally, both in finally-accented words/phrases and those without accent nuclei. A two-way RM-ANOVA with position and accent pattern as two fixed factors indicated significant effects of both factors and the interaction between them for each set of test items with two, three, and five moras, respectively ($p<0.0001$). Only in four-mora words/phrases, the effect of position and the interaction between position and accent pattern were significant ($p=0.0008$). Multiple Scheffé comparisons showed the final vowels were significantly longer than any other preceding vowel in the items with no accentual fall ($p<0.05$). In each set of items with two, three, and five moras, the final vowels of words/phrases with no accentual fall were significantly longer than any other vowel of those with an accentual fall ($p<0.0001$). In addition, in test words/phrases with accentual falls, no significant difference in duration between the constituent vowels were detected. Thus, it is quite evident that vowels are significantly lengthened only in the utterance-final position when words/phrases have no pitch fall induced by accent.

2.5.4 *Isochrony in moras*

Table 2-1 has demonstrated that there is a considerable variation in mean mora duration depending on position in words/phrases. It can be seen from the table, however, that the medial moras, that is, the moras other than the initial and final moras, tend to be of similar duration in reiterant speech for isolated six- or seven-mora words/phrases. There exists no significant difference between any pair of the medial moras in six- or seven-mora utterances. This finding is suggestive of an isochronous tendency in Japanese moras except for the initial and final moras in an utterance of six or more moras, which is roughly consistent with the finding from Experiment 1 (2.3.4).

The shortest initial mora and the longest final mora should be further discussed regarding isochronous tendency in moras. Lehiste (1977;1979) has shown that "an increment was needed that ranged from 30 to more than 100 ms" (1977:257) for subjects to perceive any interval longer than others, by employing four intervals of three basic durations (300, 400, and 500 ms). However, she has noted that when the four intervals were of equal length "the listeners tended to hear the first interval as longer than the others" (*Ibid.*) and the fourth [final] interval as shorter.

Lehiste's findings imply that the shortest initial mora and the longest final mora observed in the current study may not be perceived as short or as long as their acoustic data indicate. Instead, it is possible that they may be perceived to be more or less similar in duration to medial moras. Further, it can be hypothesized that an utterance made up of a shorter initial mora and a longer final mora plus medial moras of average durations may be perceived as more isochronous than one composed of moras of exactly same length.

However, the basic intervals in Lehiste's experiment that replicated English stresses were much longer than average mora intervals of Japanese. In addition, there may exist a cross-linguistic difference in the perception of intervals between English and Japanese listeners. Future investigation is required to examine the above assumption of the relation between the actual duration and perception of Japanese moras.

2.5.5 *Bimoraic unit*

The next question that we have to ask is whether any other durational variation is consistently observed in the component moras of test items. In two- to four-mora words, the variation in mora duration can be accounted for by the tendency for the initial mora to be consistently shortest and for the final mora to be longest in words without accentual falls. This timing pattern of four-mora words observed in this study is almost consistent with the data of Satô (1977a,b; 1994) who analyzed 32 four-mora words without accent nuclei uttered in isolation by one speaker. Moreover, his data indicate that the mean duration of the second moras is slightly longer than that of the third moras, which is not found in this study. He calls the timing pattern "a zigzag pattern", that is, alternation of shorter and longer moras, and assumes that "the bimoraic concatenation generates another rhythmic structure in mora sequences" in Japanese (1994: 85).

However, the current experimental results suggest the possibility that the presence of two pairs of shorter and longer moras may result from the shortest initial and longest final moras in four-mora words without an accentual fall.

In five mora words or phrases in Table 2-1, however, the test items show a considerable difference in mean duration between the

third and the fourth moras. The difference between them is statistically significant at the level of 0.05 (Fisher's PLSDs). In addition, it is evident from Fig. 2-4C that the five-mora reiterant speech without an accentual fall (*sumiyakiba*) shows a repetition of alternate shorter and longer moras except for the final mora, though there is no significant difference between the second and third moras.

Also, in reiterant speech for the three five-mora sentences, the mean mora duration alternately increases and decreases as a function of position, as can be seen in Fig. 2-5. Table 2-2 indicates that the fourth mora of *nagasireru* is significantly longer than any other preceding moras and shorter than the following mora (Fisher's PLSD, $p<0.05$). In addition, the mean duration of the second moras is slightly longer than that of the third moras in all the three sentences, though there exists no significant difference.

Table 2-4. Effect of position on mora duration (ms) in reiterant speech for isolated 6-mora words

Word replaced by 'ma'	No. of tokens	1st mora	2nd mora	3rd mora	4th mora	5th mora	6th mora	RM-ANOVA
namidana'gara	14	118.7 ***	164.1	165	166.1	168.6	168.1	F(5,65)= 25.50
SD		17.2	22	16.2	18.4	19.8	18.3	p<0.0001
sigotoba'kari	13	123.5 ***	170.1	167.4	167.6	169.3	181.7	F(5,60)= 16.53
SD		24.1	17.1	16.8	7.9	9.4	27.9	p<0.0001
sumiyaki'gama	14	125.9 ***	170.4	163.3 **	176.8	171.1	184.7	F(5,65)= 19.90
SD		21.2	16	15.2	14.9	16.7	20.2	p<0.0001

*** and * indicate significant differences at the levels of $p<0.001$ and $p<0.05$, respectively, from all other mora durations by post hoc tests.

When a reiterant word/phrase is as long as six moras, by contrast, there is no significant difference among the medial four

moras as mentioned above. However, according to Figs. 2-4A, 2-4C, and 2-4D, mora durations in the reiterant speech for *sumiyaki'gama* shows an alternation of shorter and longer moras, while *namidana'gara* or *sigotoba'kari* replaced by 'ma' exhibits a rather stable duration in medial moras. Table 2-4 presents mora duration averaged across speakers for each of the three six-mora items. Only the compound word *sumiyaki'gama* exhibits significant differences between the third and fourth moras, and fifth and sixth moras (Fisher's PLSD, $p<0.05$). In addition, the mean duration of the second moras is slightly longer than that of the third mora. Consequently, alternation of shorter and longer moras is repeated three times within the word.

Although all three of the six-mora compound words have the same number of moras and the same accent pattern (accent is located at the fourth mora), their morphological structures have a considerable difference. The word *sumiyaki'gama* can be divided into three bimoraic units by its morphological boundaries, whereas the other two cannot (*namida*| *na'gara*, *si*| *goto*| *ba'kari*, *sumi*| *yaki'*| *gama*). This suggests an influence of morphological structure on the occurrence of the timing pattern of alternate shorter and longer moras in reiterant speech for the six-mora words/phrases.

Finally, let us discuss the observed alternation of shorter and longer moras. When an utterance is composed of simple repetitions of the syllable 'ma' as in Experiment 2, the effect of segmental variation on mora duration is assumed to be excluded from the measured data. It has also been found that accent has a significant influence only on the duration of utterance-final moras, not on the duration of accented moras. If Japanese had no other

timing than the mora-timing model, then, the durations of all of the medial moras in reiterant speech should have been of almost equal durations with small irregular fluctuations. Thus, such a regular alternation of shorter and longer moras as observed in some of the reiterant speech implies the grouping of two moras, a bimoraic foot, asserted by researchers mostly based on morphological and/or phonological data (Poser, 1990; Satô, 1994). In addition, the results suggest that the size of utterance and morphological boundaries may be relevant to the occurrence of the bimoraic unit.

However, this experiment also makes it clear that the alternation of shorter and longer moras does not occur consistently in a statistically significant way. Thus, further investigation is required to examine the occurrences and lack thereof of the phenomenon and the factors that contribute to its occurrence.

2.6 Summary

All of the findings from the above experiments can be summarized as follows:

(1) The mean mora duration of isolated words/phrases decreases until the word/phrase size exceeds about four moras, whereas the same words/phrases embedded in a carrier sentence tend to maintain constant mora mean duration. The segmental shortening in isolated words/phrases is not restricted to the final mora, but rather all of the component segments shrink accordingly as the mora count increases up to five moras. This finding indicates that the size of utterance is one of the crucial factors influencing the duration of component moras/segments.

(2) Mean mora durations begin to gradually decrease or fluctuate

in a small range in utterances longer than four or five moras. This leads to the linear relationship between utterance length in moras and in real time. In other words, the results suggest mora timing in an utterance composed of more than about four or five moras.

(3) Reiterant speech which substituted isolated words, phrases, and sentences with repetitions of 'ma' makes it clear that accent has no consistent effect on the duration of an accented mora. It also reveals the effect of position such that the first mora is consistently the shortest of all component moras.

(4) A strong effect of the presence or absence of an accentual fall on the duration of an utterance-final mora is consistently observed in isolated reiterant speech of three to five moras. In addition, the utterance-final mora is significantly longer than all other component moras in words/phrases/sentences without an accentual fall. When they have an accentual fall, by contrast, the utterance-final mora shows no significant difference with the preceding moras.

Analyses of segmental duration clarify the observation that consonants are consistently shortest in utterance-initial position, whereas vowels are significantly longest in the utterance-final position of test items with no accentual fall. That is, the durational increase in utterance-final position is supposed to be realized in the utterance-final vowel.

(5) Alternation of shorter and longer moras is observed in some reiterant speech composed of five and six moras. There is a possibility that a morphological boundary may affect the occurrence of the bimoraic unit.

[1] The subjects who participated in this investigation totaled 13. They were all native speakers of Tokyo Japanese who were born in Tokyo and had lived there at least until they became about 20 years old. The dialect is loosely identified as standard or common Japanese.
[2] The term 'elasticity' is used here to refer to the varied amount of stretchable or compressible duration specific to each segment in context (cf. Port et al., 1980; Campbell, 1992a).
[3] Even if a phrase contains an initially- or medially-accented word, it sometimes happens that the phrase that contain the word does not have an accentual fall as in the case of *hana-no iro* ("the color of a flower").
[4] Since all factors were varied within speakers, a repeated measures analysis of variance (RM-ANOVA) is used in order to factor out interspeaker variation throughout this research whenever tokens are repeatedly obtained from several speakers under varying conditions.
[5] All statistical analyses in this thesis were conducted by using a SAS StatView.
[6] A Fisher's PLSD is used as a post-hoc test (multiple comparison) when the numbers of samples are the same, whereas a Scheffé is employed when they are not.

CHAPTER 3

Effect of accent on duration in isolated words and phrases

3.1 Introduction

In the preceding chapter we have observed a strong effect of the presence/absence of an accentual pitch fall on final vowel duration in reiterant speech that replaced real words, phrases or sentences with repetition of nonsense 'ma'. Although the results showed a highly significant main effect, there is a possibility that the experimental design that forced subjects to produce unnatural utterances may have brought about biased results. Thus, the data obtained by this method should be reinforced by data using real words, phrases, and sentences. Experiment 3 explores the effect of the presence/absence of an accentual fall on the duration of final vowels by employing sets of real words that have the same segmental structure but are different in accent pattern.

Second, the accent-related effect on the duration of final vowels is examined in isolated phrases. To be more specific, Experiment 4 tests which of the two accents, the preceding or the phrase-final, affects the duration of utterance-final vowels in phrases uttered in isolation, and attempts to clarify the crucial factor involved in the effect of accent on duration specific to the

utterance-final position of isolated words and phrases.

3.2 Experiment 3

The purpose of this experiment is to examine the effect of accent on segmental duration of real word utterances and clarify a strong effect of the presence or absence of an accentual fall on duration restricted to the final vowel of isolated words.

3.2.1 *Subjects*

The subjects were five middle-aged females, all native speakers of Tokyo (standard) Japanese.

3.2.2 *Materials*

Test words were 14 sets of 31 real Japanese words listed in Appendix 3-1 ranging from two to four moras in length. The words in each set were the same in segmental make-ups but different in accent pattern. In addition to pairs of test words with or without an accentual fall, there were several sets of three words that had different accent patterns. These sets were prepared to examine whether the position of accent nucleus has anything to do with the duration of final vowels. In *kisetu*-set, for instance, accents are placed at the first and second moras in *ki'setu* and *kise'tu*, respectively, whereas *kisetu* has no accentual fall. Most of the test words were constructed by only CV syllables. In order to complement the number of test words, three pairs of four-mora words that contain moraic consonants (N) and/or long vowels were added. (e.g. /seNko:/)

Each test word was written in the Japanese orthography (*kanji* and *kana),* one word per card.

3.2.3 *Procedure*

All 31 test words were mixed with 15 dummy words/phrases/sentences. Each subject was presented a test word (or a dummy item) on a card one at a time in a random order and asked to read the word naturally at a tempo comfortable to herself. Every time the subject finished reading all the test items, the cards were randomized and presented to the subject, for a total of seven different randomizations.

Productions were recorded by a Sony DAT (TCD-D10) in a sound proof room and digitized at a sampling rate of 11 kHz, and waveforms, wideband spectrograms, and pitch contours were made with a Kay Multi-Speech Model 3700. As a rule, the tokens from the second to the seventh reading were measured. Only when a token was misread or wrong in accent assignment, was the token discarded and the first utterance used instead. The total number of tokens analyzed was (31 sentences × 6 repetitions × 5 subjects=) 930.

Measurements were obtained by using manual cursors for duration of (1) every segment that was measurable in all the test words; (2) the VOTs for voiceless plosives; (3) the entire utterance.

For determining segmental boundaries, the same criteria used in Experiment 1 were used except for some criteria added for this experiment. When test words began with voiceless plosive /k/ as in /kakeru/ and /kisecu[1]/, measurements were made from the beginning of the VOT. When a vowel was devoiced or nasalized as in the case of /i/ or /o/ in /icigo/, the whole length of the syllable or a sequence of segments was measured.

The beginning or terminal point of a voiceless fricative/affricate such as /h/, /s/ or /c/ was determined by the same criteria as those for Experiment 1. The consonantal period was in

most cases easily identified by imperiodical waveforms of noise. When the beginning or terminal point of imperiodical waveforms was not distinct, measurement was done from or until the point where formants of a vowel were visible on the spectrogram, whether the frication had ceased or not. The beginning and terminal points of a VOT succeeding a voiceless stop were determined by the same criteria as those for fricatives/affricates.

In order to explore in depth the extent of the final lengthening in Japanese, the end of each utterance was defined to be the point at which the periodic waveform of the vowels ended. Since visible energy of formants depends partly on the setting of the "palette" of spectrogram, and partly on the intensity of the sound, there is a possibility that a vowel with lower f_0 and smaller energy may end up earlier on spectrogram. Thus, the period of periodical waveforms of a prepausal vowel was measured instead of that of visible energy of formants, considering that the former reflects glottal oscillation more faithfully than the latter. Vocal fry register pulses at utterance-final position were included only in the case that they occurred within 15ms from the end of the preceding glottal-pulse striation.

All the segmental durations were repeatedly measured to ascertain the accuracy of measurements and the consistency of criteria.

In addition to the duration measurement, the mean F0, and the maximum and minimum F0 of the final vowel were analyzed by visual inspection. When a F0 dropped abruptly, however, a voice sometimes changed into a creaky sound toward the end of an utterance. In this case, no periodic waveform was observed in the final vowel, and thus no F0 measurement was made.

3.3 Results & discussion

3.3.1 *Effect of accent on segmental duration*

Figures 3-1 to 3-14 plot the mean durations of segments or syllables for each set of test words. (The error bar in the figures shows a standard error above or below the mean.) As is obvious from the figures, corresponding segmental durations do not vary to a large degree except for the final vowels among test words composed of the same segments, suggesting little effect of accent on the duration of accented segments or syllables. Still, there were several segments/syllables whose duration indicated a significant effect of accent. The asterisks ** and * at the bottom of each figure indicate significant differences at the levels of $p<0.0001$ and $p<0.05$, respectively, by RM-ANOVAs and/or Fisher's PLSDs. Segments without asterisks show no significant difference related to differing accent patterns.

Several durational tendencies are consistently observed across the sets. First, accented V syllables (moras) that occur word-initially are significantly longer than their unaccented counterparts in /atari/ and /icigo/. The initial vowel /a/ in *a'tusa* and *a'sa* are also non-significantly longer than their counterparts in *atusa,* and *asa*, respectively. Second, accented vowels in medially-accented *kake'ru, kise'tu*, and *senkyo'ku*, are all shorter than their counterparts without accent in a statistically significant way (RM-ANOVAs, $p<0.05$; Fisher's PLSDs, $p<0.0001$). The second syllable in medially-accented *iti'go* is also shorter than its counterpart in *itigo* or *i'tigo* with no statistic significance.

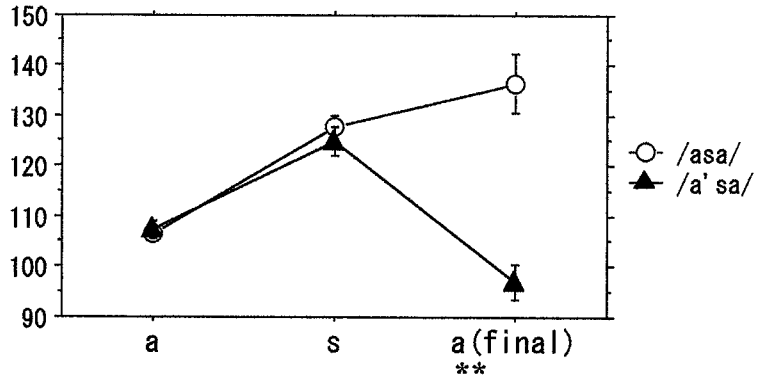


Fig. 3-1. Mean segmental duration (ms) of /asa/ and /a'sa/

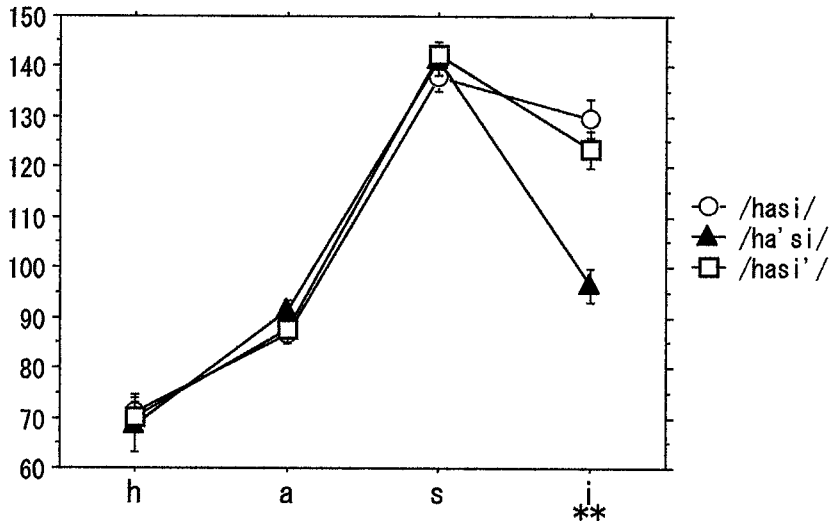


Fig. 3-2. Mean segmental duration (ms) of /hasi/, /ha'si/, and /hasi'/

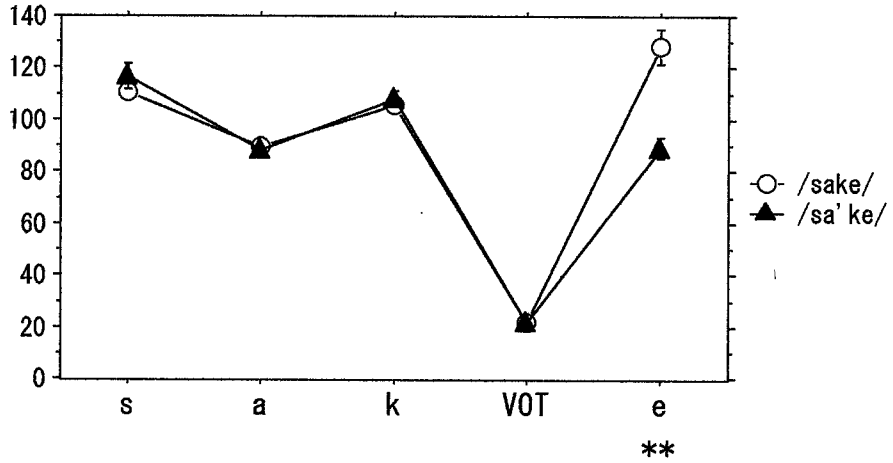


Fig. 3-3. Mean segmental duration (ms) of /sake/ and /sa'ke/

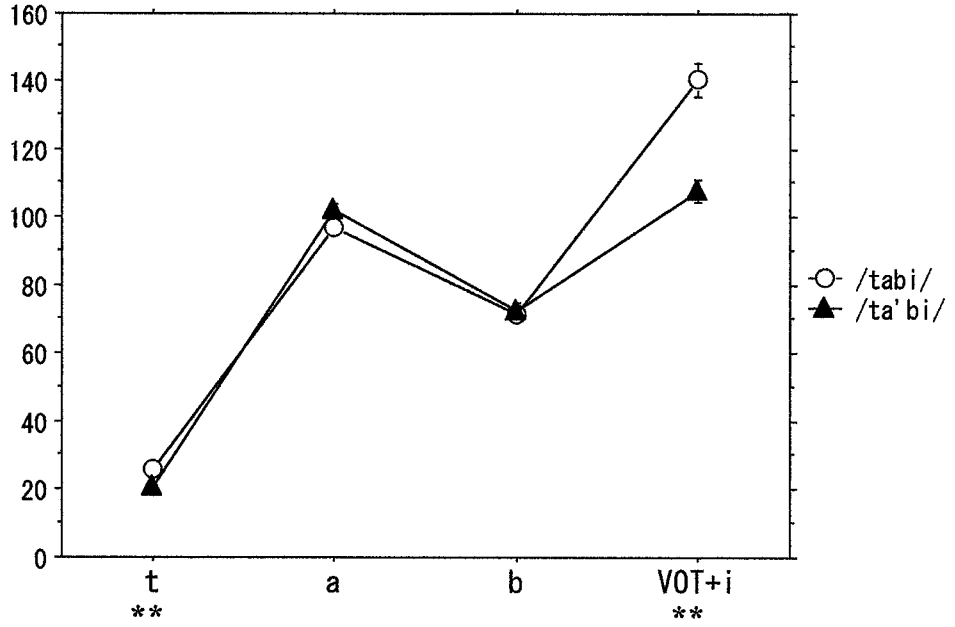


Fig.3-4. Mean segmental duration (ms) of /tabi/ and /ta'bi/

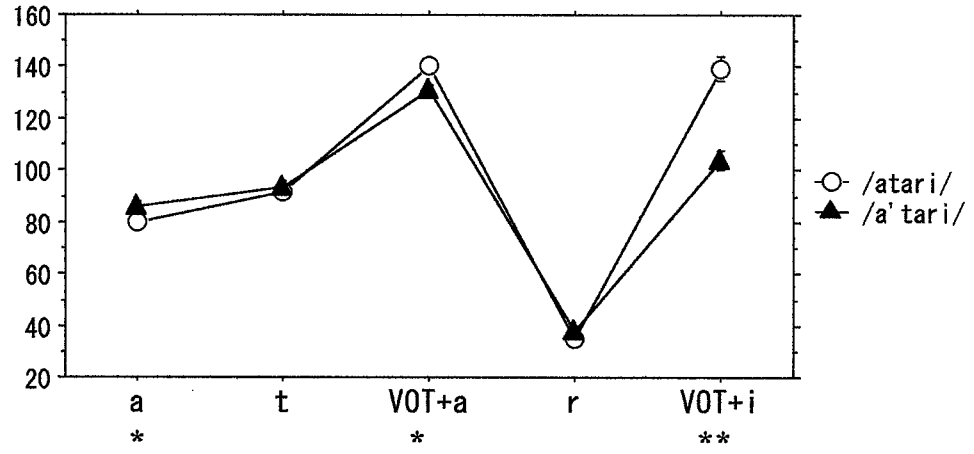


Fig. 3-5. Mean segmental duration (ms) of /atari/ and /a'tari/

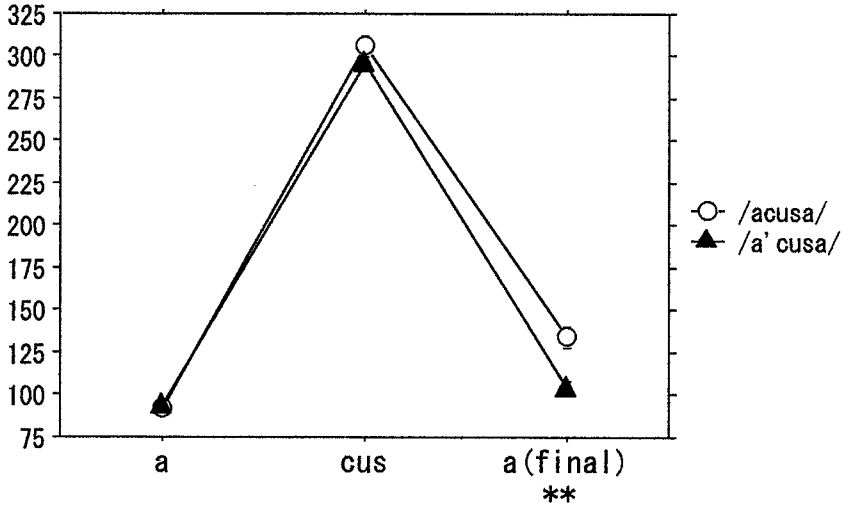


Fig. 3-6. Mean segmental duration (ms) of /acusa/ and /a'cusa/

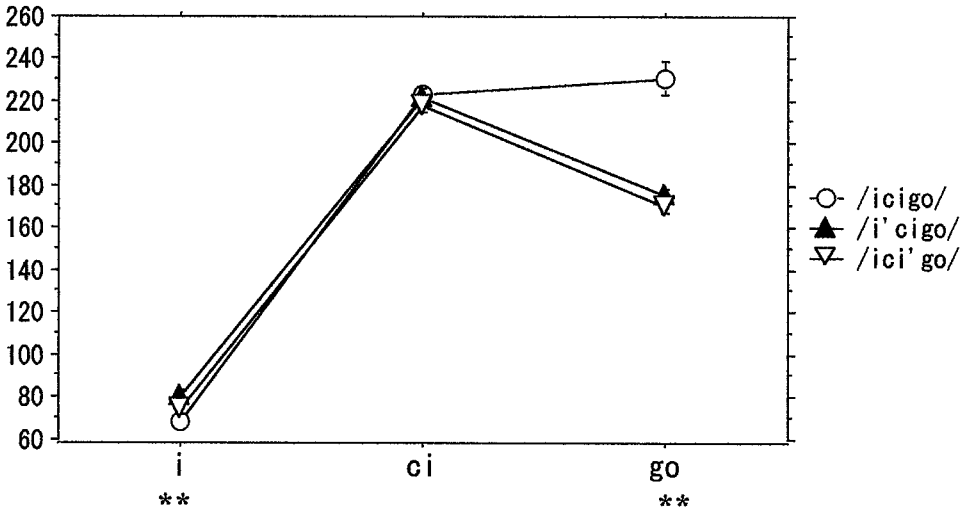


Fig. 3-7. Mean segmental duration (ms) of /icigo/, /i'cigo/, and /ici'go/

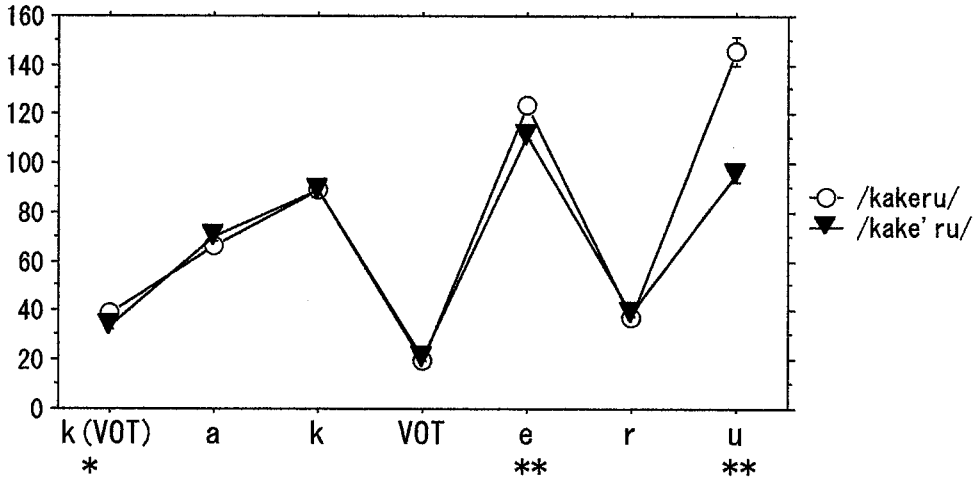


Fig. 3-8. Mean segmental duration (ms) of /kakeru/ and /kake'ru/

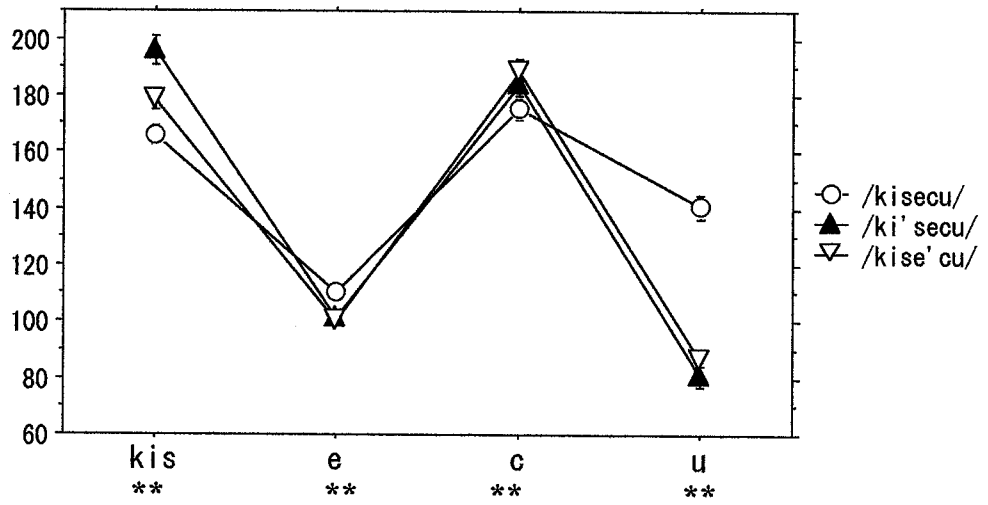


Fig. 3-9. Mean segmental duration (ms) of /kisecu/,/ki'secu/,/kise'cu/

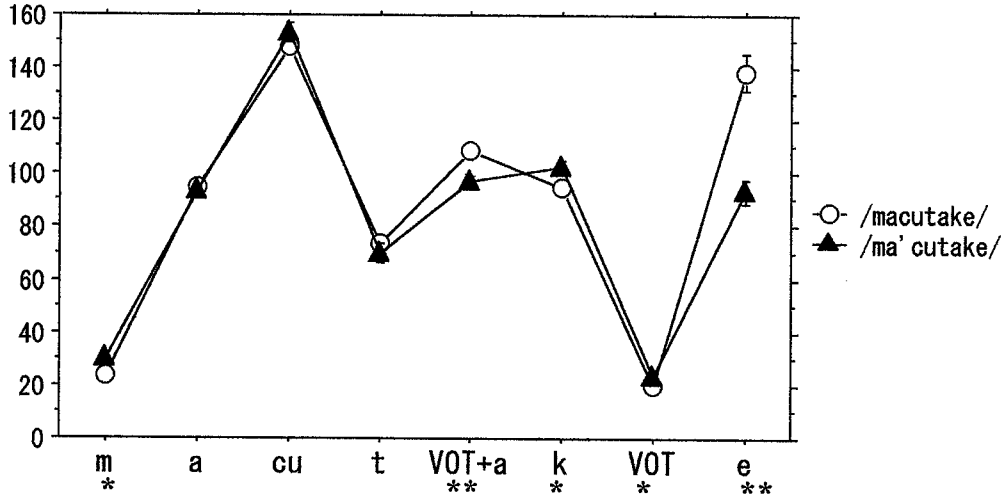


Fig. 3-10. Mean segmental duration (ms) of /macutake/ and /ma'cutake/

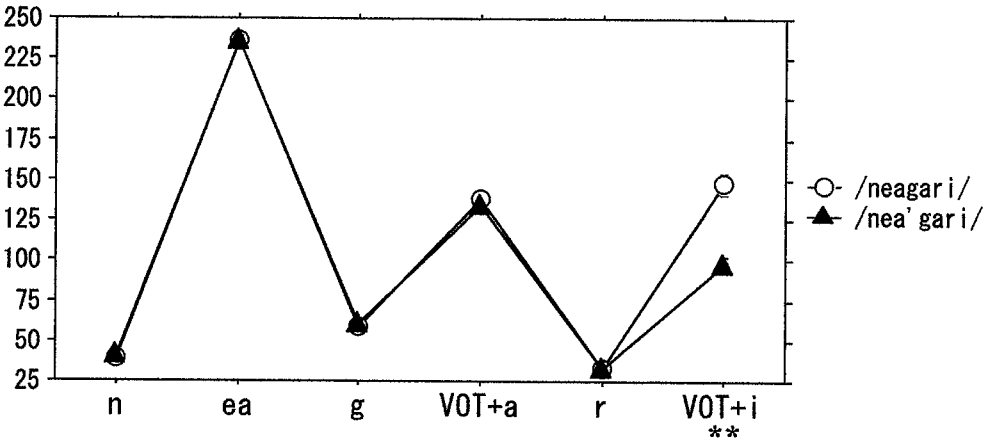


Fig. 3-11. Mean segmental duration (ms) of /neagari/ and /nea'gari/

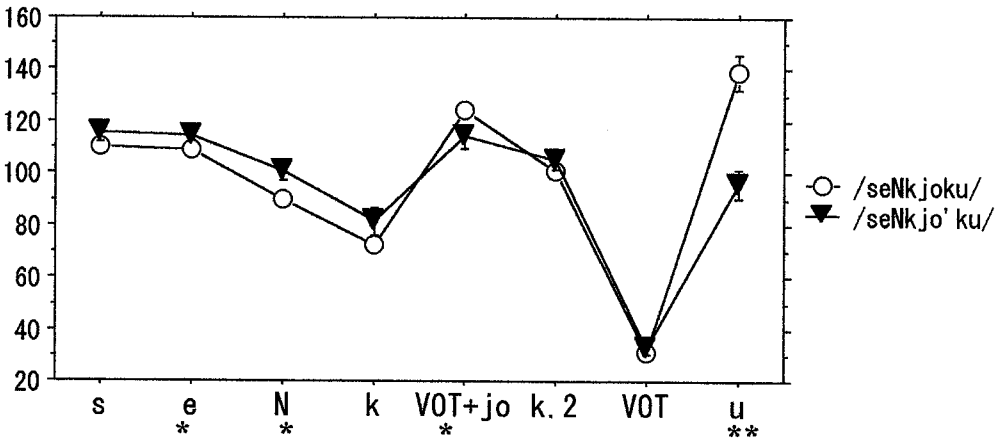


Fig. 3-12. Mean segmental duration (ms) of /seNkjoku/ and /seNkjo'ku/

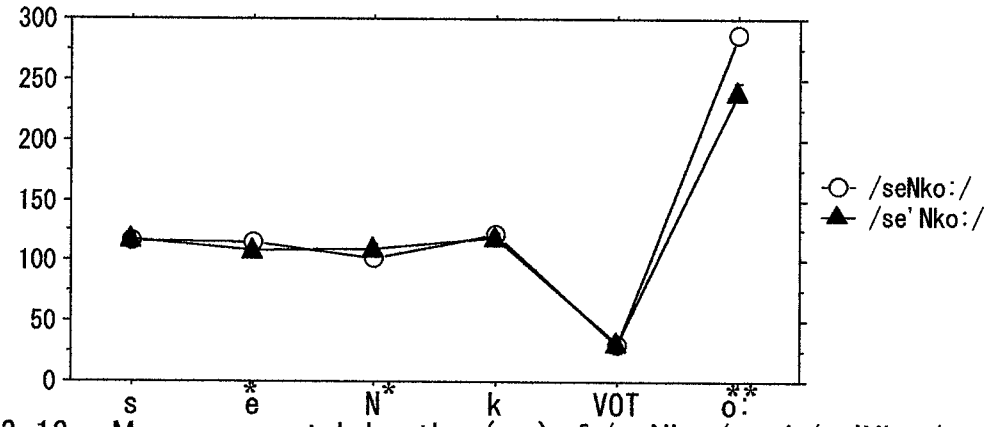


Fig. 3-13. Mean segmental duration (ms) of /seNko:/ and /se'Nko:/

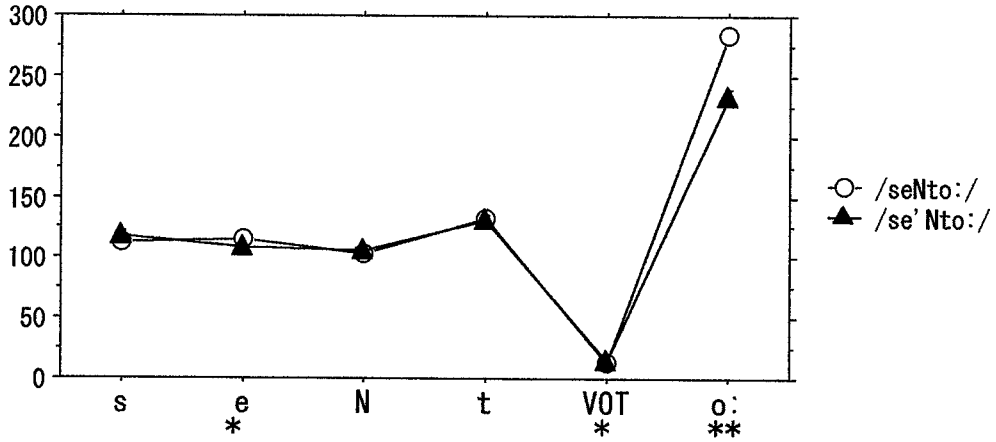


Fig. 3-14. Mean segmental duration (ms) of /seNto:/ and /se'Nto:/

When an initially-accented mora is composed of a consonant and a vowel, however, accent-related shortening is observed in accented vowels. Accented first vowels /e/ in *se'nkoo* and *se'ntoo* are significantly shorter than their counterparts with no accent. The accented vowel /a/ in *sa'ke* is also non-significantly shorter than its counterpart in *sake*. However, the accented /a/ in *ta'bi* is significantly longer than that in *tabi*. Accented initial consonant /m/ in *ma'tutake* is significantly longer than its counterpart without accent. The consonants /s/ in /sake/, /seNkjoku/, and /seNto:/ also gain an extra length non-significantly when the initial moras are accented. Thus, accented consonants tend to gain extra length in word (utterance) initial position, whereas the following vowels tend to be shortened. However, this phenomenon is not consistently and significantly observed. There is a possibility that segmental factors such as neighboring segments, intrinsic duration and elasticity of accented segments may partly contribute to the shortening or lengthening of segments that consist of accented moras. Further investigation to factor out these segmental variations is required in order to assert any consistent effect of accent on the segmental duration of accented moras.

Another finding from the figures is that in any pair of corresponding segments in the sets the durational difference is relatively small except for the final vowels. Most of the difference falls within about 10ms as far as the duration of a single segment is concerned. In the case of intrinsically long segments or a sequence of segments such as /kis/ and /c/ in /kisecu/ and VOT+a in /macutake/, discrepancy of over 15ms is observed.

3.3.2 *Effect of accent on word final vowels*

Table 3-1 presents all the mean durations of the final vowels of test

words averaged across speakers. The mean F0 (Hz) and range (semitone[2]) of the final vowel for each test word are also given. "Difference" in all the table of this thesis is obtained by subtracting the utterance-final vowel duration of a word with an accentual fall from its counterpart without an accentual fall.

As is obvious from the table, the final vowels of unaccented or finally-accented test words are considerably longer than those of initially- or medially-accented words. In addition, this phenomenon is consistently observed in all five types of Japanese vowels despite the variation in intrinsic duration or elasticity of the vowels, or irrespective of preceding consonants. When the long vowel /oo/ and the syllable /go/ in /icigo/ are excluded, the mean difference in duration between the final vowels of words with an accentual fall and those without is 40.5ms collapsed over sets, ranging from 27.1ms to 59.8ms. The mean ratio of the longer vowels to the shorter vowels is 1.43:1, ranging from 1.28:1 to 1.74:1. A one-way analysis of variance of repeated measures (RM-ANOVA) with accent pattern as a fixed factor shows a highly significant effect of accent pattern on the duration of final vowels for each set of test words ($p<0.0001$).

Three of the test sets (*hasi*-set, *itigo*-set, and *kisetu*-set) have three test words with distinctive accent patterns. However, as is shown in Table 3-2, there is no significant difference in the duration of the final vowels between initially- and medially-accented words, and between unaccented and finally-accented words. By contrast, the final vowel of initially-accented *ha'si* is highly significantly shorter than those of *hasi* and *hasi'*. The final vowels of *itigo* and *kisetu* with no accent nucleus are significantly longer than their counterparts of accented *i'tigo/iti'go* and *ki'setu/kise'tu*.

Table 3-1. Mean duration (ms), mean F0 (Hz), mean range (semitone), and SD (parentheses) of the final vowels for each set of test words

Test words	Vowel	SD	Difference	RM-ANOVA	F0	Range
asa'	136.5	32.2	39.4	F(1,29)=52.4	**237.6 (26.3)	**4.6(2.2)
a'sa	97.1	19.3		p<0.0001	172.4 (11.5)	9.6(4.1)
hasi	129.7	21.5	27.1	F(2,58)=44.7	**232.4(23.1)	*4.9(1.8)
hasi'	123.5	20.6	33.3	p<0.0001	236.5(29.8)	4.9(2.6)
ha'si	96.4	18.3			161.3 (14.0)	6.5(3.5)
sake	128.2	36.6	39.2	F(1,29)=44.9	**226.3 (13.7)	*4.9(1.8)
sa'ke	89.0	23.0		p<0.0001	164.8 (10.2)	7.8(4.4)
tabi'	140.6	27.3	32.9	F(1,29)=64.0	**236.7 (16.3)	**4.1(1.8)
ta'bi	107.7	18.4		p<0.0001	162.1 (12.1)	7.8(3.9)
atari	139.3	25.3	35.3	F(1,29)=107.7	**220.9 (12.3)	*3.9(1.4)
a'tari	104.0	21.3		p<0.0001	142.0 (14.7)	5.6(4.1)
atusa	134.2	33.1	29.9	F(1,29)=57.1	**223.9 (14.4)	*5.4(1.8)
a'tusa	104.3	22.7		p<0.0001	159.5(11.2)	7.6(3.9)
itigo	230.9	43.0		F(2,58)=51.1	**214.5 (13.0)	* 4.9(2.3)
i'tigo	175.4	16.9	55.5	p<0.0001	144.6 (15.9)	6.3(3.5)
iti'go	170.5	17.2	60.4		164.4 (18.0)	7.2(3.4)
kakeru	145.8	31.9	51.1	F(1,28)=81.8	**213.6 (15.3)	**4.7(2.4)
kake'ru	94.7	18.0		p<0.0001	153.6 (11.8)	7.4(3.1)
kisetu	140.7	22.4		F(2,58)=146.0	**214.1 (11.8)	4.7(1.7)
ki'setu	80.9	20.2	59.8	p<0.0001	137.6 (14.5)	4.9(3.7)
kise'tu	86.7	13.7	54.0		145.1 (12.1)	6.9(4.0)
matutake	138.4	38.7	45.1	F(1,29)=59.7	**204.0 (12.0)	*3.9(2.0)
ma'tutake	93.3	24.9		p<0.0001	138.7 (11.6)	8.0(5.0)
neagari	147.7	34.7	49.8	F(1,29)=154.2	**204.2 (15.0)	* 3.8(1.4)
nea'gari	97.9	27.8		p<0.0001	134.5 (17.4)	6.8(4.7)
senkyoku	139.0	36.8	43.3	F(1,29)=140.7	**210.9 (16.7)	4.1(2.0)
senkyo'ku	95.7	29.9		p<0.0001	150.7(11.2)	5.4(4.1)
senkoo	286.6	31.6	47.0	F(1,29)=39.0	**213.8 (16.1)	*4.3(1.4)
se'nkoo	239.6	40.0		p<0.0001	150.1(13.2)	6.4(5.7)
sentoo	284.4	33.3	51.3	F(1,29)=99.2	**214.2 (14.2)	*4.6(1.3)
se'ntoo	233.1	32.6		p<0.0001	150.4(12.0)	7.5(4.7)

** and * indicate statistic differences at the levels of p<0.0001 and p<0.05, respectively by RM-ANOVA.

Table 3-2. Results of post-hoc tests for duration (ms), F0 (Hz), and range (semitone) of final vowels for hasi-, itigo-, and kisetu-set.

Test words	Duration (ms)	F0 (Hz)	Range (semitone)
hasi-set	ha'si vs. hasi(') p<0.0001	ha'si vs. hasi(') p<0.0001	ha'si vs. hasi(') p<0.05
itigo-set	itigo vs. i'tigo/iti'go p<0.0001	itigo vs. i'tigo vs. iti'go p<0.0001	itigo vs. iti'go p<0.01
kisetu-set	kisetu vs. ki'setu/kise'tu p<0.0001	kisetu vs. ki'setu/kise'tu p<0.0001 ki'setu vs. kise'tu p<0.05	kise'tu vs. kisetu/ki'setu p<0.01

In isolated utterances, finally-accented words are supposed to have no accentual pitch fall since there are no following particles. Initially- or medially-accented words, on the other hand, have a F0 fall induced by accent nucleus.

Thus, these results indicate that the factor contributing to the significant difference in the duration of the final vowels is not the position of accent nucleus, that is, the individual accent pattern, but the presence or absence of an accentual fall within the test word utterances.

To sum up, the experimental results have made it clear that there is a marked effect of the presence or absence of an accentual fall on the duration of final vowels in isolated words. That is, the final vowel of a word with no accentual fall is consistently and significantly longer than that with an accentual fall. The difference is on average about 43ms and the ratio is 1.44:1.

3.3.3 *Effect of accent on final lengthening in isolated words*

The next question to be addressed is whether the duration increase

in the final vowels of words without an accentual fall can be considered to be 'lengthening' To put it another way, is the final vowel really longer than all other preceding vowels? In Experiment 2 in which segmental variation is factored out by using reiterant speech, 'final lengthening' is defined as "durational increase observed in the final syllable/moras/segments that leads to a significant durational difference from their preceding counterparts"(2.5.3).

In this experiment of real words, the segmental make-ups of the preceding syllables are greatly different from those of the final ones which makes it impossible to directly compare the obtained data. Thus, in order to grasp how long the average duration of medial vowels is, the mean duration of the vowels of all the second and third CV syllables[3] in three- or four-mora test words for Experiment 3 was calculated. The mean vowel duration is 104.6ms (SD 13.9), excluding devoiced vowels, VOTs for voiceless stops, and /o/ in /icigo/. The average duration of the final vowels of test words with or without an accentual fall is 95.9ms, and 136.4ms, respectively. Compared with the mean duration of medial vowels, the final vowels are assumed to be lengthened by about 30% in test words without an accentual fall, whereas no lengthening is assumed to take place in those with an accentual fall.

It is suggested from all the experimental results that the final vowel of an isolated word with no accentual fall undergoes final lengthening while its counterpart with an accentual fall does not.

3.3.4 *Interspeaker variation*

So far only repeated measures analyses of variance have been used to examine the effect of accent on duration. In repeated measures

analyses of variance, duration under one condition is compared with that under another within a single subject. That is, interspeaker variation caused by speaker-related factors such as speaking rate or speaking style are excluded from the statistic results.

However, it should be mentioned that there was a substantial variation among the speakers in the durational realization of the final vowels. Some speakers tended to drawl out the final vowel to a large degree, whereas some showed a tendency to shorten the final vowel. Figure 3-15 illustrates, for instance, the mean

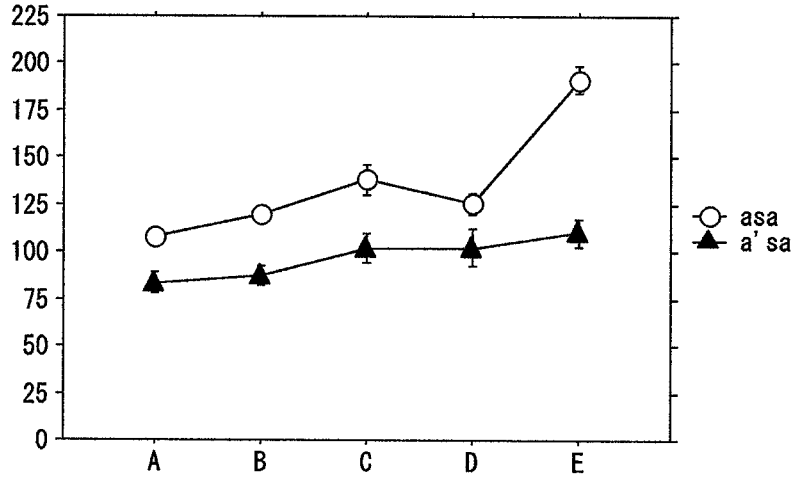


Fig. 3-15. Mean duration (ms) of final vowel in *asa* and *a'sa* for each speaker

duration of the final vowels in *a'sa* and *asa* for each speaker. The absolute duration of the final vowel /a/ exhibits a large variation among speakers. However, every speaker uttered the final /a/ in *asa* significantly longer than that in *a'sa*, and the difference in the vowel duration ranges from about 25ms to 80ms. The same tendency is consistently observed in a statistically significant way in all the other sets of test words. That is, despite the variation of speaking style or speaking rate, all the speakers exhibit the

tendency to lengthen the final vowel of a test word without an accentual fall while they tend to shorten its counterpart with an accentual fall.

3.3.5 *Correlation between F0 and duration in isolated words*

It follows from all the results that the presence or absence of an accentual fall has a consistent effect on the duration of final vowels in isolated words utterances. The question that we have to ask here is whether an accentual fall is really the factor most highly related to the length of final vowels. Is there any other factor that influences the duration of final vowels as well? As is clear from Table 3-1, the mean F0s of the final vowel in each set of words showed a highly significant effect of accent (RM-ANOVA, $p<0.0001$). The results are readily expected since accent nucleus is what causes F0 to drop. In addition, the presence/absence of an accentual fall indicated a significant effect on the mean range of the final vowels in twelve out of fourteen sets of test words at the level of $p<0.0001$ or $p<0.05$, while no significant effect was found in the other two sets.

What is the factor that contributes most to the durational differences in the final vowels? In the scattergram of Figure 3-16 the mean F0 of the final vowel for each test word is plotted against its duration. We can see positive correlation between duration and mean F0 of all the final vowels of test words. However, the group of initially- and medially-accented words is clearly separated from that of unaccented and finally-accented words.

The correlation coefficients of all the word-final vowel durations with their mean F0s and minimum F0s were 0.47 and 0.44 ($p<0.0001$), respectively. The coefficient between the duration and mean F0 for each set ranged from 0.41 for *sake*-set to

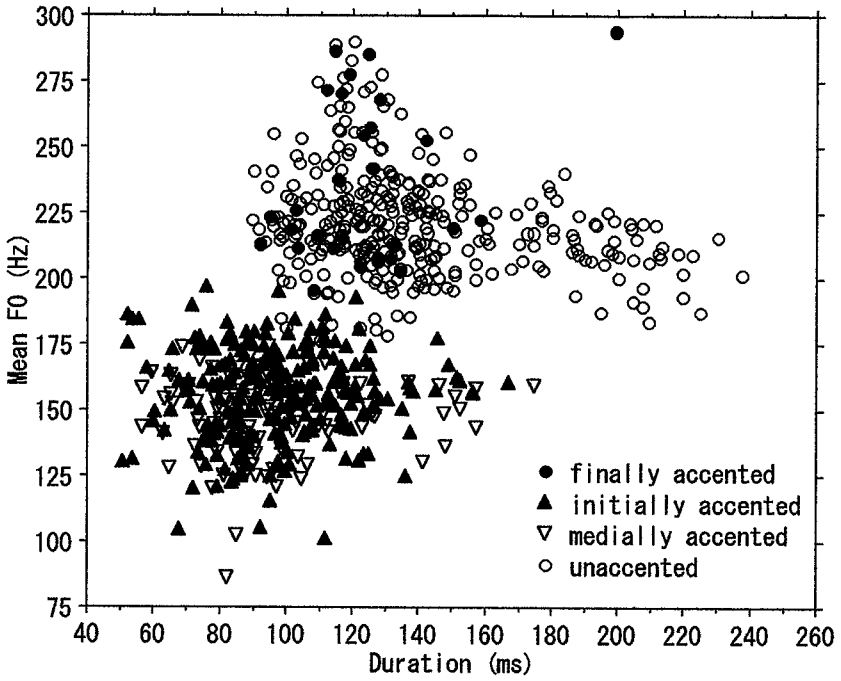


Fig. 3-16. Mean duration (ms) of final vowel for each test word plotted against mean F0

0.75 for *kisetu*-set ($p<0.05$). The duration and range of the final vowels showed a small negative correlation of -0.15. Thus, it can be assumed that the mean F0 and minimum F0, and the duration of word-final vowels are correlated to one another since accent is what triggers an F0's drop.

3.4 Experiment 4A

This experiment was carried out in order to examine whether the effect of accent on duration observed in the final position of isolated words is valid in isolated phrases. It also attempts to explore which accentual fall in a phrase, the preceding or the final, has a crucial effect on the duraion of utterance-final vowels in phrases uttered in isolation.

3.4.1 *Subjects*

The subjects were four middle-aged females, all native speakers of Tokyo (standard) Japanese. Two of them had participated in Experiment 3.

3.4.2 *Materials*

Test phrases were 18 sets of 39 real Japanese phrases listed in Appendix 3-2. The phrases are composed of "Noun 1 (N1)[4] + genitive particle *no* + Noun 2 (N2)", or "adjective + "Noun 2 (N2)". All the final nouns (N2) for each set have the same segmental structure. In all the test sets, the difference in accent pattern of N1 or N2 leads to the presence or absence of an accentual fall. In five out of the 18 sets, N2s are different in the presence/absence of an accentual fall (e.g. *zi-no sake* vs. *zi-no sa'ke*). In the other sets in which N2s are the same nouns without an accentual fall, N1s differ in the presence/absence of an accentual fall (e.g. *momo-no gara* vs. *he'bi-no gara*).

Each test phrase was written in the Japanese orthography (*kanji* and *kana),* one phrase per card.

3.4.3 *Procedure*

Experiment procedure was the same as that for Experiment 3. The tokens analyzed totaled (39 phrases × 6 repetitions × 4 subjects=) 936.

Measurements were obtained by using manual cursors for durations of (1) every segment measurable in all the final syllables of each test phrase; (2) the VOTs for voiceless plosives; (3) the entire utterance. In the sets whose phrases are made up of the same segments, all the component segments were measured as long as segmental boundaries were distinct enough to be determined.

The criteria for determining segmental boundaries were the same as those employed in Experiments 1 and 3 with some additional criteria for this experiment. When test words begin with voiceless plosive /t/ as in the sets of *toti-no sake* and *toti-no sa'ke*, measurements were made from the beginning of the VOT. Since vowel /u/ in /acusa/ was devoiced, the whole length of /cus/ was measured. All the segmental durations were repeatedly measured to ascertain the accuracy and consistency of the measurements.

In addition to the duration measurement, the mean F0, and the maximum and minimum F0 of the final vowel were measured by the same method as that employed in Experiment 3.

3.5 Results & discussion

3.5.1 *Effect of N2's accent on utterance-final vowels*

Table 3-3 presents the mean duration, mean F0, and mean range of utterance-final vowels for each set of test phrases whose final nouns (N2s) differ in accent[5]. That is, every set consists of a test phrase with no accentual fall in N2 and that with an accentual fall in N2.

As is evident from the table, the utterance-final vowel of a phrase without an accentual fall in N2 is considerably longer than its counterpart with an accentual fall. A one-way RM-ANOVA was run for each pair of phrases with the presence/absence of an accentual fall in N2 as a fixed factor, showing the effect highly significant ($p<0.0001$). The mean difference in vowel duration is 35.4ms, ranging from 26.8 to 40.7ms. The mean ratio of the longer vowel to the shorter one is 1.40:1, ranging from 1.29:1 to 1.46:1.

Let us compare these results with those in isolated word

Table 3-3. Mean duration (ms), mean F0 (Hz), mean range (semitone), and SD (parentheses) of the final vowels for each set of test phrases whose final nouns differ in accent

Test words	Vowel	SD	D	RM-ANOVA	F0	Range
zino sake	127.8	34.0		**F(1,23)=90.23	**211.3(33.4)	*3.3(2.4)
zino sa'ke	87.8	26.3	40.0	p<0.0001	167.6(21.5)	4.8(1.7)
totino sake	129.7	27.1		**F(1,23)=80.81	**214.3(27.8)	*3.0(1.7)
totino sa'ke	91.6	27.1	38.1	p<0.0001	163.2(19.6)	5.6(3.6)
kamino atusa	136.5	30.6		**F(1,23)=93.95	**211.5(31.9)	**3.8(1.5)
natuno a'tusa	95.8	22.0	40.7	p<0.0001	165.0(19.8)	6.5(3.1)
murano matutake	120.2	24.4		**F(1,23)=24.56	**212.6(31.0)	4.0(2.5)
murano ma'tutake	93.4	29.6	26.8	p<0.0001	149.7(21.0)	4.2(2.8)
o'okuno senkyoku	103.4	27.7		**F(1,23)=44.91	*178.9(23.6)	2.9(1.7)
ti'isana senkyo'ku	72.1	21.9	31.3	p<0.0001	140.5(33.4)	2.9(2.2)

** and * indicate statistic differences at the levels of p<0.0001 and p<0.05, respectively by RM-ANOVA.

utterances of Experiment 3 (3.3.2). In the data of Experiment 3, the durational difference between the utterance-final vowels of words with an accentual fall and those without is on average 40.5ms, i.e., 5.1ms longer than that observed in this experiment. However, the utterance length ranged from three to eight moras in the current experiment, whereas only two to four mora test words were used in Experiment 3. As has been discussed in Chapter 2, utterance length greatly affects component segment duration. Thus, it is suggested that the longer utterance length of the test phrases in Experiment 4A results in shorter segmental duration and thereby differences in the final vowel duration. In addition, this suggestion is supported by the fact that the ratio of the longer and shorter vowel duration in the word utterances (1.43:1) is almost the same as the ratio obtained from the phrase utterances in this experiment (1.40:1).

Figures 3-17, 3-18, and 3-19 illustrate the segment or syllable

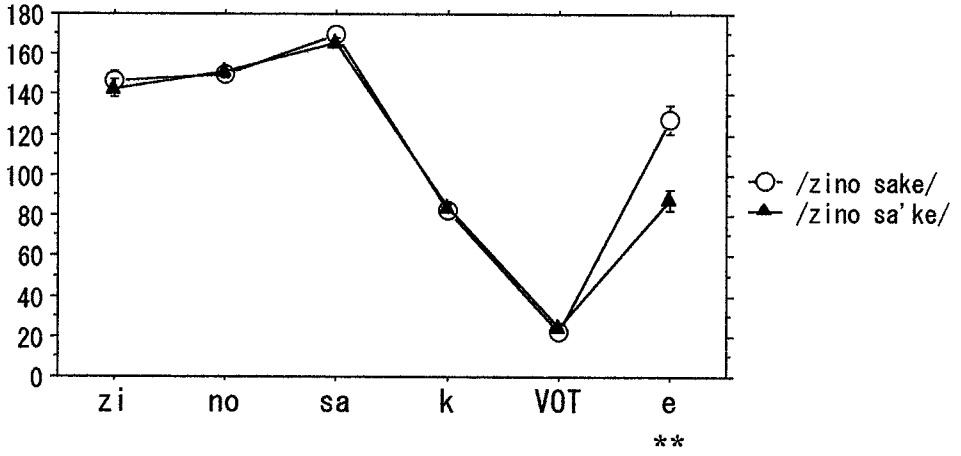


Fig. 3-17. Mean segment/syllable duration (ms) of /zino sake/ and /zino sa'ke/

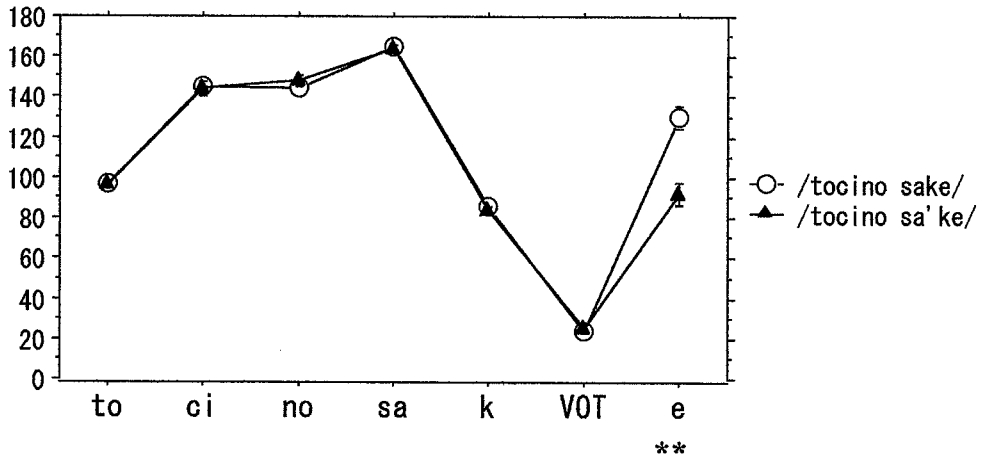


Fig. 3-18. Mean segment/syllable duration (ms) of /tocino sake/ and /tocino sa'ke/

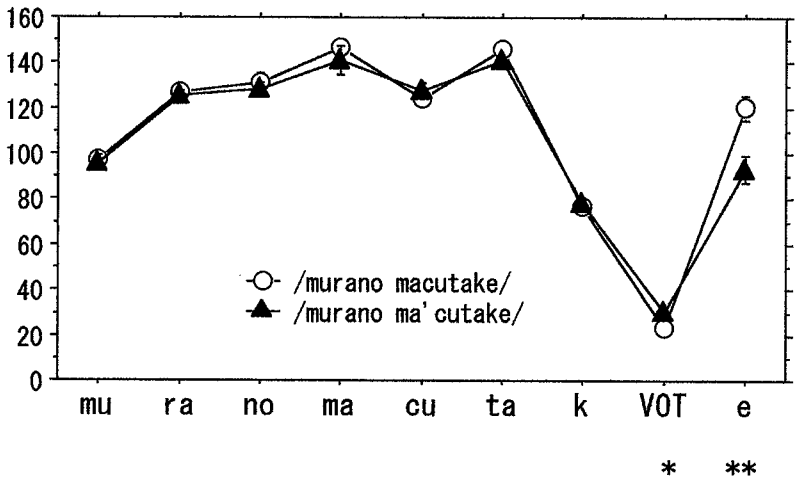


Fig. 3-19. Mean segment/syllable duration (ms) of /murano macutake/ and /murano ma'cutake/

duration of each set of test phrases. (The asterisks ** and * at the bottom of each figure indicate significant differences at the levels of $p<0.0001$ and $p<0.05$, respectively, by RM-ANOVAs. The error bar in the figures shows a standard error above or below the mean.)

As is obvious from the figures, the duration does not differ greatly between each pair of phrases that are made up of the same segment. The only exception is utterance-final vowels whose duration significantly differs depending on the presence/absence of an accentual fall in N2 of each pair of test phrases. It should be added that the final VOT in /murano macutake/ also shows a slight but significant difference of 6.8ms at the $p<0.05$ level.

It can be concluded from the above results that the effect of accent on duration is consistently and significantly observed between the presence/absence of an accentual fall in the final noun phrase (N2) and the duration of the utterance-final vowels in isolated phrases. In other words, it has been clarified that the utterance-final vowels in phrases with no accentual fall in N2s are substantially longer than their counterparts with an accentual fall in N2. The difference is on average 35.4ms, and the ratio of the longer vowel to the shorter one is 1.40:1.

3.5.2 *Effect of N1's accent on utterance-final vowels*

Table 3-4 presents the mean duration, mean F0, and mean range of utterance-final vowels for each set of test phrases whose preceding nouns (N1s) differ in accent. That is, every set consists of one phrase with no accentual fall and one or two phrases with an accentual fall in N1.

In seven out of thirteen sets, RM-ANOVAs showed a significant effect of the presence/absence of an accentual fall in the preceding nouns at the level of $p<0.05$. The absolute difference ranges from

Table 3-4. Mean duration (ms), mean F0 (Hz), mean range (semitone), and SD (parentheses) of the utterance-final vowels for each set of test phrases whose preceding nouns differ in accent

Test words	Vowel	SD	Difference	RM-ANOVA	F0	Range
konona	150.4	29.6		F(1,23)=2.78	**224.7(28.5)	2.3(1.3)
ki'nona	142.3	23.5	8.1	p=0.11	166.3(22.1)	2.4(1.3)
kono to	139.5	32.8		*F(1,23)=4.88	**224.3(29.0)	3.2(1.7)
ki'no to	127.8	30.5	11.7	p=0.037	167.1(13.5)	3.2(1.8)
mino gara	139.6	22.3		F(1,23)=1.16	**210.5(34.1)	2.0(1.1)
ki'no gara	135.3	23.8	4.3	p=0.29	169.4(23.8)	1.9(1.7)
umeno mi	127.6	25.4		F(1,23)=0.074	**222.8(39.9)	3.0(1.4)
gu'mino mi	126.5	29.1	1.1	p=0.79	161.1(20.5)	2.7(1.7)
sakeno na	147.5	28.6		*F(1,23)=10.82	**220.3(33.7)	2.9(2.1)
sa'keno na	133.0	23.0	14.5	p=0.0032	163.0(20.7)	2.0(1.3)
momono gara	139.0	20.8		*F(1,23)=5.03	**212.2(36.7)	*2.6(1.7)
he'bino gara	131.0	18.7	8	p=0.035	157.7(25.4)	1.8(1.4)
uekino na	142.7	31.3		*F(2,46)=3.86	**208.7(36.3)	2.8(1.8)
sa'kushano na	124.8	29.5	17.9	p=0.028	164.2(19.9)	2.3(2.2)
hana'yano na	135.7	22.9	7		163.3(20.4)	2.4(2.1)
siborino gara	138.5	26.3		F(2,46)=1.32	**202.1(41.8)	2.3(1.4)
o'tibano gara	133.0	23.0	5.5	p=0.28	159.3(16.5)	2.3(1.9)
higa'sano gara	130.6	18.9	7.9		157.8(24.2)	2.2(1.6)
sakeno umami	122.8	25.7		*F(1,23)=8.85	**204.7(32.1)	2.4(1.2)
sa'keno umami	107.7	19.4	15.1	P=0.0068	175.6(39.4)	2.0(1.1)
kirino getabako	116.2	28.3		*F(1,23)=13.22	**195.6(29.7)	2.6(1.6)
ka'sino getabako	103.1	22.5	13.1	p=0.0014	163.5(26.7)	2.9(1.7)
tonarino ueki	112.4	27.4		F(2,46)=2.86	**205.9(42.0)	2.7(1.1)
ka'danno ueki	110.9	31.3	1.5	p=0.068	170.8(23.2)	3.1(1.3)
hana'yano ueki	103.7	24.2	8.7		176.9(30.9)	3.3(1.9)
zyoozaino kikime	118.1	20.8		*F(1,23)=4.59	**202.4(35.9)	*3.2(1.8)
syo'oyakuno kikime	108.5	20.6	9.6	p=0.043	171.1(28.6)	1.9(1.2)
yasaino neagari	114.9	24.8		F(1,23)=0.64	**197.3(47.7)	2.1(0.9)
i'hukuno neagari	119.0	22.4	-4.1	p=0.43	167.0(30.8)	2.3(1.3)

** and * indicate statistic differences at the levels of $p<0.0001$ and $p<0.05$, respectively, by RM-ANOVA.

9.6 to 17.9ms. Thus, the duration discrepancy in utterance-final vowels is greatly different in its degree from that observed in the isolated phrases whose N2's accent differs, or that observed in the isolated word utterances in Experiment 3.

In the other sets of Table 3-4, the utterance-final vowel shows no significant variation related to the presence or absence of an accentual fall in the preceding noun. One set undergoes shortening when N1 has no accentual fall: The utterance-final /i/ in *yasaino neagari* is 4.1ms shorter than that in *i'hukuno neagari*.

Thus, all the results indicate that the presence/absence of an accentual fall in the preceding noun (N1) does not affect the utterance-final vowel duration to the same degree as the presence/absence of an accentual fall in the final noun (N2) does. In other words, there is no correlation between N1's accent and the duration of utterance-final vowels equivalent to that observed between N2's accent and the duration of utterance-final vowels in isolated phrases.

If no equivalent correlation is observed in N1's accent and the duration of utterance-final vowels, it is also relevant to ask whether the utterance-final vowels undergo lengthening or not. Table 3-4 shows that the mean duration of utterance-final vowels in phrases with no accentual fall in the N2s is 120.0ms, with a range of 103.1ms to 150.4ms. The mean duration of the final vowels in phrases whose N2s have no accentual fall in Table 3-3 is 123.5ms, ranging from 103.4ms to 136.5ms. Thus, all the absolute durations of utterance-final vowels fall within almost the same range when the test phrases have no accentual fall in the N2s. These findings suggest that the phrases in Table 3-4 whose N2s have no accentual fall undergo final lengthening of a similar

amount to their counterparts in Table 3-3. The results also suggest that the presence/absence of an accentual fall in N2 functions as a decisive factor in the occurrence or lack thereof of utterance-final lengthening in isolated phrases.

3.5.3 *Utterance-final lengthening vs. word-final lengthening*

The next question to be addressed herein is whether the effect of accent on vowel duration is exclusively observed in utterance-final position. Does the word-final vowel undergo any lengthening or shortening related to the accentual fall of the word? Figures 3-20 and 3-21 illustrate the durational variation in /sa(')keno na/ and /sa(')keno umami/. Each pair of the phrases consists of the same segments and the only difference in the phrases is the presence or absence of an accentual fall in the preceding noun /sake/.

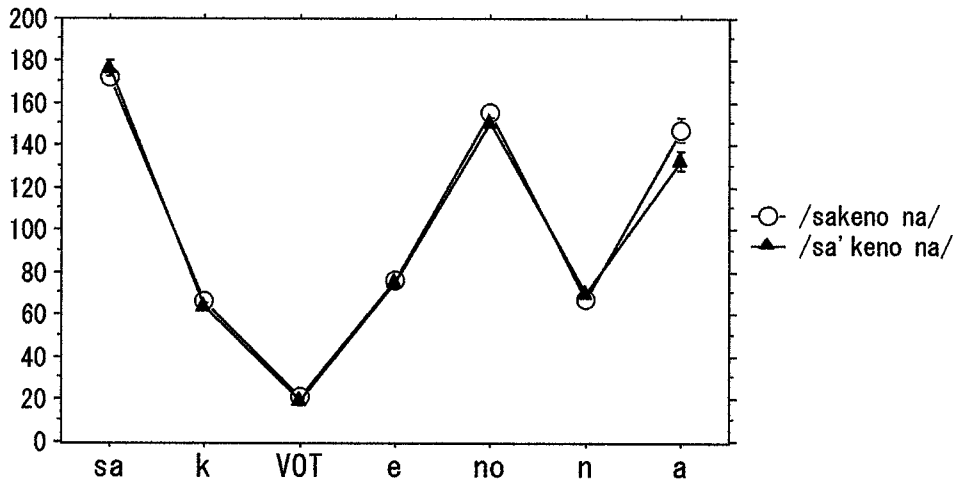


Fig. 3-20. Mean duration (ms) of segments/syllables of /sakeno na/ and /sa'keno na/

It is noticeable that no large segmental increase or decrease is caused by the presence/absence of an accentual fall in /sake/ except for the utterance-final vowel: The utterance-final vowels in *sakeno na* and *sakeno umami* are 11.7ms and 15.1ms longer than *sa'keno*

na and *sa'keno umami*, respectively. The difference was significant at the level of $p<0.05$ by RM-ANOVA. Thus, the results indicate that the effect of accent on duration is lost in the final position of the preceding noun.

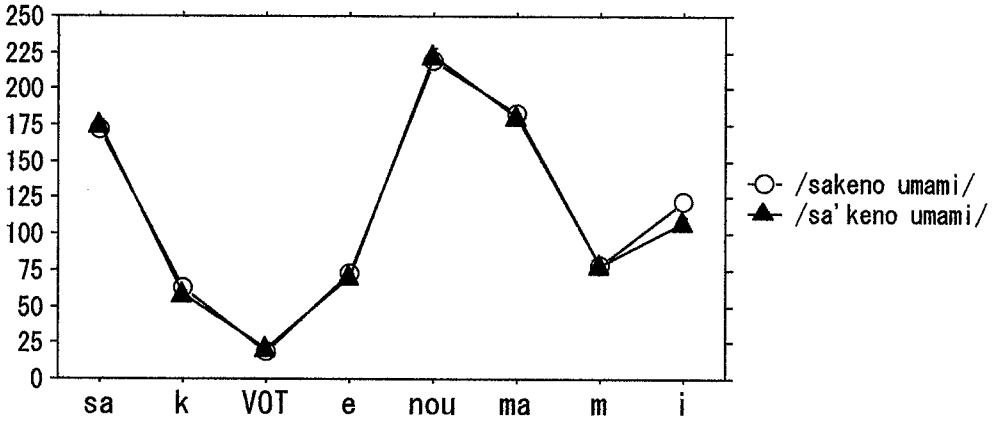


Fig. 3-21. Mean duration (ms) of segments/syllables of /sakeno umami/ and /sa'keno umami/

All these findings suggest that the accent-related effect is specific to utterance-final vowels, or prepausal vowels in isolated word and phrase utterances.

3.6 Experiment 4B

This experiment was carried out to obtain data that complement those from the previous experiment (Experiment 4A). Although Experiment 4A has clarified several points, the findings obtained from it have raised a couple of new questions. First, the word- or utterance-final syllables tested in the previous experiment were mostly /ke/ except for /sa/ and /ku/ (Table 3-3, Figs. 3-20, 3-21). It is necessary to examine whether the final syllable made up of the other consonant and vowel brings about the same results. Second,

it is necessary to find out whether the effect of accent on final lengthening is observed at the major lexical boundary between adjective phrase "N1 + no" and final noun N2. Although the pair of /sakenona/ shows no significant difference in the genitive *no*[6] (Fig. 3-20), a few more test phrases should be added to affirm the result.

3.6.1 *Subjects*

The subjects were five middle-aged females, all native speakers of Tokyo (standard) Japanese. Three of them had participated in Experiment 4A.

3.6.2 *Materials*

Test phrases were two sets of eight real Japanese phrases listed in Appendix 3-3. The phrases are composed of "Noun 1 (N1) + genitive particle *no* + Noun 2 (N2)", or "adjective + "Noun 2 (N2)". Each pair of the phrases in each set has the same segmental structure but is different in the presence/absence of an accentual fall: One pair differs in N1's accent, the other in N2's accent. These pairs were prepared in order to compare the effect of accent on duration in utterance-final position with that in phrase-final position.

Each test phrase was written in the Japanese orthography (*kanji* and *kana),* one phrase per card.

3.6.3 *Procedure*

The test phrases were mixed with 20 other test phrases and sentences that will be used for other experiments. The procedure of the experiment was the same as that in Experiments 3 and 4A. The total number of tokens analyzed was (8 phrases × 6

repetitions × 5 subjects=) 240.

Measurements were obtained by using manual cursors for durations of (1) every segment that was measurable in all the test phrases; (2) the VOTs for voiceless plosives.

The criteria for determining segmental boundaries were the same as those used in all the previous experiments of this study, except for some additional criteria for this experiment. The boundary between /g/ and preceding or following vowel was determined at the point in which formants of the vowel appeared.

In addition to the duration measurement, the mean F0, and the maximum and minimum F0 of the final vowel were analyzed in the same method as in the previous experiments.

3.7 Results & discussion

3.7.1 *Effect of segmental variation on utterance-final vowels*

Table 3-5 shows the mean duration, mean F0, and mean range of utterance-final vowels for each set of test phrases whose N2s or N1s differ in accent.

As can be seen from the table, the presence or absence of an accentual fall in /siro/ and /tabi/ affects the utterance-final vowel duration significantly when the two words are in the utterance-final position.

When /siro/ and /tabi/ comprise the preceding adjective phrases, by contrast, the accent difference results in no durational difference of a similar amount to that observed in the former case. In *siro*-set, the utterance-final vowel /e/ of *kabe* shows a significant duration difference of 10.0ms at the $p<0.05$ level when *siro* and *si'ro* comprise N1s. The degree of the difference is considered to be distinguishable from that observed when they are in the

utterance-final position.

Table 3-5. Mean duration (ms), mean F0 (Hz), mean range (semitone), and SD (parentheses) of the final vowels for each set of test phrases

Test words	Vowel	SD	Difference	RM-ANOVA	F0	Range
mi'gotona siro	143.7	21.9		**F(1,29)=62.7	*176.9(17.0)	**2.3(1.6)
mi'gotona si'ro	114.1	16.5	29.6	p<0.0001	167.0(17.8)	5.8(3.4)
sirono kabe	147.9	35.3		*F(1,29)=6.26	**203.5(26.4)	2.0(1.4)
si'rono kabe	137.9	24.7	10.0	p=0.018	170.7(21.1)	2.1(1.1)
naga'i tabi	139.9	35.7		**F(1,29)=40.93	**180.2(19.0)	*2.2(2.0)
naga'i ta'bi	103.7	22.0	36.2	p<0.0001	155.8(23.7)	4.2(2.9)
tabino na'gasa	107.8	29.2		F(1,29)=0.17	152.2(18.3)	4.1(3.3)
ta'bino na'gasa	106.4	32.1	1.4	p=0.68	147.7(13.8)	5.1(4.8)

** and * indicate statistic differences at the levels of p<0.0001 and p<0.05, respectively by RM-ANOVA.

The final syllables /ro/ and /bi/ of /siro/ and /tabi/ consist of consonants /r/ and /b/, and vowels /o/ and /i/, respectively. In Experiment 4A, the utterance-final syllables that showed a strong effect of accent on duration were composed of voiceless plosive /k/ and fricative /s/, and vowels /e/, /a/, /u/. Therefore, the results from this experiment, together with those from Experiment 4A have made it clear that the highly significant effect of an accentual fall on the utterance-final vowel duration is consistently observed despite the segmental variation in the syllable in the utterance-final position.

3.7.2 *Phrase-final lengthening vs. utterance-final lengthening*

Table 3-6 presents the mean duration of consonants and vowels of the final syllables in the preceding adjective (phrase) for each test phrase. (The target syllable in phrase-final position is underlined.)

One of the most obvious findings from the table is that the

difference in the duration of the phrase-final consonants and vowels marks a small minus value in six out of eight pairs, ranging from -6.3 to -0.4. That is, the target segment is slightly shortened when an accentual fall is absent from N1 or N2. It can also be seen that the absolute difference is as much as 6.3ms, which is far smaller than that observed in the utterance-final vowels of words differing in the presence/absence of an accentual fall, although in the pair of /sirono kabe/ the final vowel /o/ of the adjective phrase shows a significant difference at the level of $p<0.05$.

Table 3-6. Mean segmental duration (ms) of the underlined phrase-final syllables for each set of test phrases

Test words	Consonant	SD	Difference	Vowel	SD	Difference
mi'goto na siro	48.9	7.7		100.4	13.9	
mi'goto na si'ro	50.9	8.2	-2	101.6	15.0	-1.2
siro no kabe	58.4	10.3		81.5	9.1	**
si'ro no kabe	55.9	11.9	2.5	87.8	9.0	-6.3
na ga'i tabi	54.4	9.4		196.1	28.2	
na ga'i ta'bi	54.8	9.7	-0.4	198.6	25.6	-2.5
tabi no na'gasa	52.2	8.7		82.4	9.2	
ta'bi no na'gasa	52.1	10.7	0.1	83.4	11.8	-1

** and * indicate statistic differences at the levels of $p<0.0001$ and $p<0.05$, respectively by RM-ANOVA.

Figures 3-22 to 3-25 show the mean segmental duration of each test phrase. Fig. 3-22 indicates the effect of an accentual fall in the utterance-final noun (N2) exclusively observed in the utterance-final vowel including VOT (VOT+o). Fig. 3-23, by contrast, shows fluctuation of segmental duration presumably due to the presence/absence of an accentual fall in N1. However, small duration difference between the two phrases is not restricted to the segments composed of N1, but spreads to the end of the utterance.

Another finding from Figure 3-23 is that the segments of one phrase alternatively show the longer and shorter duration than those of the other phrase, suggesting temporal compensation within a unit of a mora or across several moras.

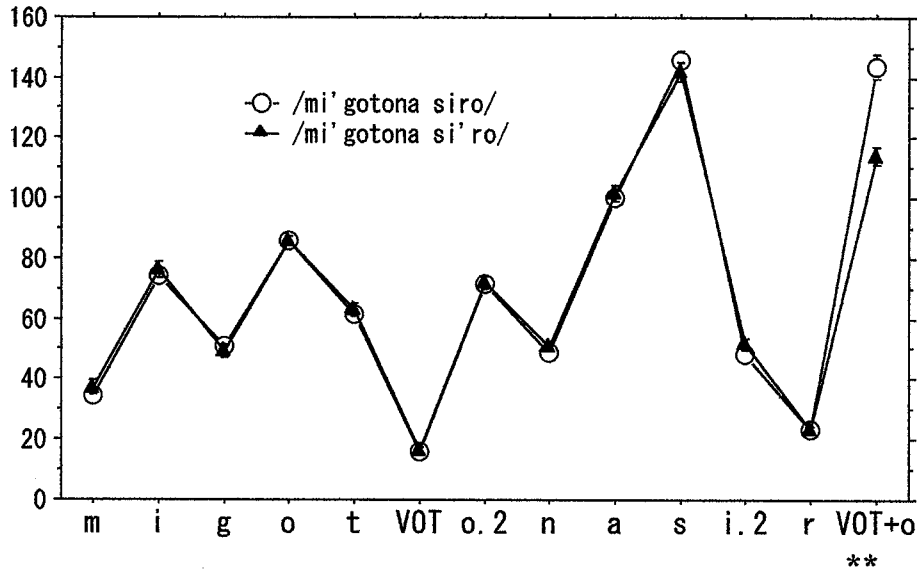


Fig. 3-22. Mean segmental duration (ms) of /mi'gotona siro/ and /mi'gotona si'ro/

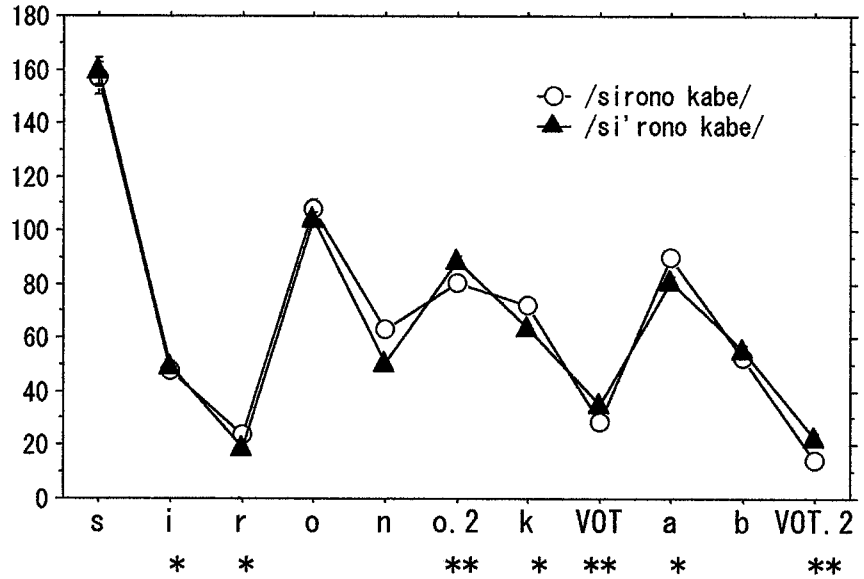


Fig. 3-23. Mean segmental duration (ms) of /sirono kabe/ and /si'rono kabe/

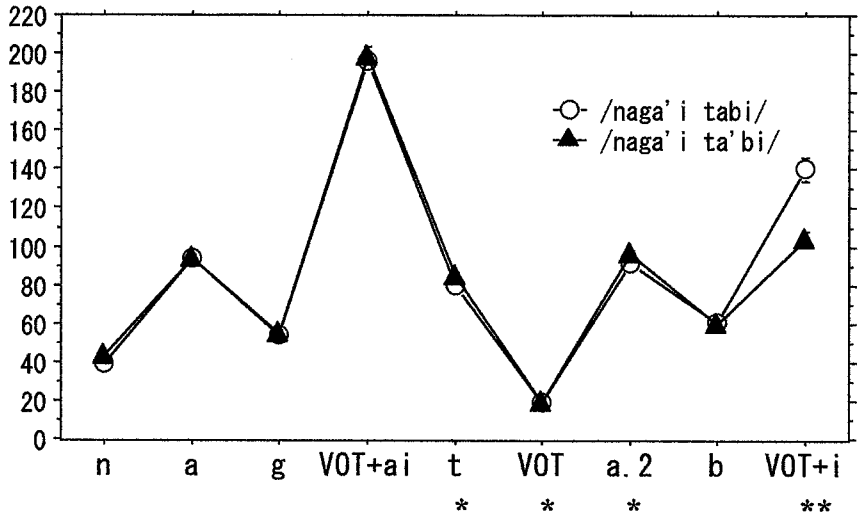


Fig. 3-24. Mean segmental duration (ms) of /naga'i tabi/ and /naga'i ta'bi/

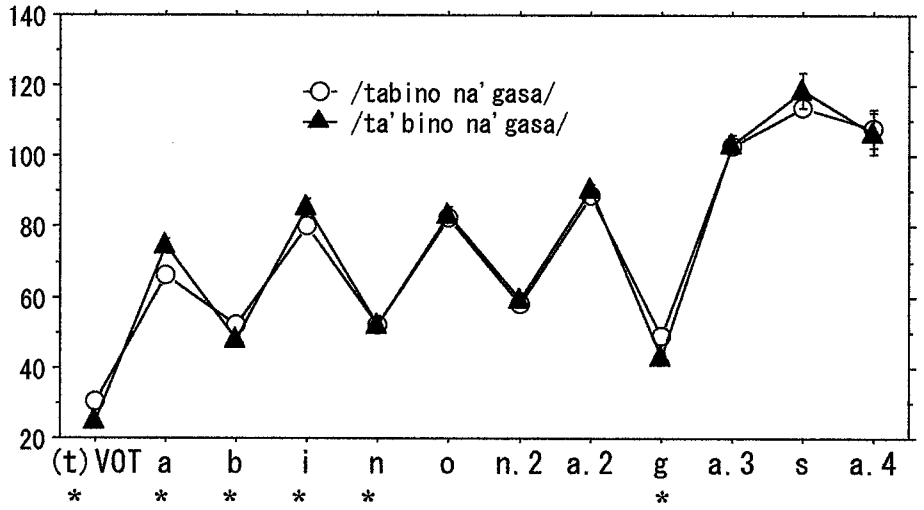


Fig. 3-25. Mean segment duration (ms) of
/tabino na'gasa/ & /ta'bino na'gasa/

Figure 3-24 also shows a marked effect of the presence/absence of an accentual fall in N2 on the duration of the utterance-final vowel /i/ including VOT. Fig. 3-25, by contrast, shows a small difference between the two phrases that differ in N1's accent as in the case of /sirono kabe/. In the same way as Fig. 3-23, the segments of one phrase alternately become longer or shorter than those of the other phrase. However, the two pairs of test phrases

share no other tendency in durational variation. This suggests that the segmental variation observed in Figs. 23 and 25 is concerned with segmental properties related to its accentuation.

Consequently, the results indicate no consistent difference related to the presence/absence of an accentual fall in N1 is observed in the final vowel of the preceding adjective phrase. Thus, it can be safely said that the consistent and highly significant duration difference of about 40% is restrictively observed between the utterance-final vowel of a phrase with no accentual fall in N2 and its counterpart with an accentual fall in N2. In other words, all the experimental results have demonstrated a strong effect of the presence/absence of an accentual fall in the utterance-final word on the duration of the utterance-final vowel. It should be added that the utterance-final position discussed here is identical with prepausal position.

3.7.3 *Correlation between F0 and duration in isolated phrases*

The last problem to be discussed is whether there is any other factor than an accentual fall that contributes to the durational variation of utterance-final vowels in isolated phrases. Unlike that in isolated words, the significant difference in mean F0 of utterance-final vowels does not always lead to a significant difference in the duration of the utterance final vowels in isolated phrases, as seen from Table 3-4. Still, all but one set in the table show slightly longer utterance-final vowels of phrases with no accentual fall in N1 than those with an accentual fall in N1. In addition, all the sets in the table showed a highly significant difference in mean F0 of the utterance-final vowels for each set.

Correlation coefficients of vowel duration with its mean and minimum F0, and semitone for all the tokens of isolated phrases

were 0.44, 0.47, and -0.22, respectively, suggesting a general tendency for an utterance-final vowel to shrink as its F0 falls ($p<0.0001$).

In the sets where N2s have accentual contrast in the presence/absence of an F0 fall, correlation coefficients of all the vowel durations with their mean and minimum F0s were 0.57 and 0.59 ($p<0.0001$). These results are easily explainable by the fact that the accentual fall in the utterance-final word causes the mean and minimum F0s of the vowel to fall utterance-finally. That is, the accentual fall in the utterance-final word, and the mean and minimum F0 of the utterance-final vowel are closely related to one another.

In sum, it is found that the F0 values of utterance-final vowels is related to their duration, suggesting a general tendency for the utterance-final vowel to become shorter as its F0 gets lower.

3.8 Summary

All the findings from these experiments can be summarized as follows:

(1) A marked effect of accent on duration is consistently observed in the utterance-final vowels of real words uttered in isolation. The final vowel of a word without an accentual fall is significantly longer than that with an accentual fall. The difference is on average about 41ms, and the mean ratio between the longer and shorter vowels is 1.43:1.

(2) The consistent effect of accent on segmental duration is restricted to the utterance-final position. No durational difference of a similar amount to that in the utterance-final position is observed in any other segments constituting the

words.

(3) In phrasal utterances, a strong effect of accent on duration is consistently observed in utterance-final position. That is, when an utterance-final word has no accentual fall the utterance-final vowel is consistently and significantly longer than the same vowel produced in a phrase in which an utterance-final word has an accentual fall. The difference is on average 35.4ms, and the ratio of the longer vowel to the shorter one is 1.40:1.

(4) The substantial lengthening of the final vowel is specific to utterance-final position. In word-final or phrase-final position with no following pause, the word- or phrase-final vowel shows neither lengthening nor shortening related to an accentual fall in the preceding noun phrase to the same degree as that found in utterance-final position.

---

[1] The presentation of each Japanese phoneme is in accordance with the list presented by Uemura(1989). Accordingly, the consonant of *ti* and *tu* is represented by /c/.

[2] The range was converted into semitones by dividing (logF0max-logF0min) by {(log880-log440)/12}.

[3] Though the difference is not statistically significant in Experiment 2, the first vowel is consistently shorter than the following vowels, which is in conformity with the previous studies (Sagisaka et al., 1984). Thus, the mean duration is calculated from the medial syllables only.

[4] In test phrases for Experiment 4A and 4B, the preceding and the final nouns are referred to as "Noun 1" and "Noun 2" and abbreviated as "N1" and "N2", respectively.

[5] What "Difference" represents in Table 3-3 is almost the same as that in Table 3-1. "Difference" in the table is obtained by subtracting the utterance-final vowel duration of a phrase with an accentual fall in N2 from its counterpart without an accentual fall.

[6] In some of the tokens /sakeno na/ by some speakers, the vowel /o/ of *no* was so nasalized that the segmentation between /n/ and /o/

was impossible. Thus, the whole length of /no/ was presented when averaging across speakers.

CHAPTER 4

Effect of accent on duration in declarative sentences

4.1 Introduction

The foregoing chapters have calarified a strong effect of an accentual fall on the duration of utterance-final vowels in isolated words and phrases. The utterance-final vowel of an utterance whose final word has no accentual fall is consistently longer than that of an utterance whose final word has an accentual fall. The difference in duration is highly statistically significant, and the former is on average more than 40% longer than the latter.

The next question we have to ask is whether the effect of an accentual fall on the duration of utterance-final vowels is observed in the utterances of sentences. Experiment 5 examines whether a similar effect is consistently observed between the presence/absence of an accentual fall and the duration of a sentence-final vowel in short declarative sentences. Specifically, it addresses which accentual fall, NP's or VP's, has a significant effect on the duration of phrase- or sentence-final vowels. The experiment also pursues what effect the accentual fall of a NP and a VP has on the duration of constituent segments other than the sentence-final vowels. Eventually, it aims to reveal that the accent-related effect is

restricted to sentence-final position.

The succeeding experiment, Experiment 6, further addresses the question whether the effect of an accentual fall in the sentence-final word on the duration of sentence-final vowel is specific to utterance-final position or whether it is observed in any prepausal position. In other words, the effect of position (sentence-final or non-sentence-final) is tested by analyzing utterances of declarative sentences that contain a pause before a major phrasal boundary.

4.2 Experiment 5

4.2.1 *Subjects*

The subjects were five middle-aged females, all native speakers of Tokyo (standard) Japanese.

4.2.2 *Materials*

Test sentences were seven sets of 18 real Japanese declarative sentences listed in Appendix 4-1. In A-Group each set of test sentences are composed of a noun phrase (NP) and a verb phrase (VP). Each pair of verb phrases is the same in segmental make-up but different in accent pattern. All the verbs or verb phrases are composed of CV syllables. Also, in *hareru*-set and *hareta*-set, the noun phrases serving as the subjects of the test sentences are different in accent pattern, producing four different combinations of accent patterns.

In B-Group each set consists of a pair of test sentences whose segmental structures are the same. The phrase-final nouns of the preceding NPs that serve as the subject or the object in each pair of sentences have different accent patterns: one has an accentual fall, while the other does not.

Each test sentence was written in the Japanese orthography (*kanji* and *kana),* one sentence per card. The total of 18 test sentences were then mixed with 20 other test sentences or phrases that will be used for other experiments.

4.2.3 *Procedure*

Each subject was presented a test sentence on a card one at a time in a random order and asked to read the sentence naturally at a tempo comfortable to herself. In addition, each subject was instructed to read each test sentence without placing a focus or emphasis on any particular constituent of the test sentence. Every time the subject finished reading all the sentences, the cards were randomized and presented to the subject, for a total of seven different randomizations.

Productions were recorded and digitized at a sampling rate of 11 kHz, and waveforms, wideband spectrograms, and pitch contours were made with a Kay Multi-Speech Model 3700. As a rule, the tokens from the second to the seventh reading were measured. Only when a token was misread or wrong in accent assignment, was the token discarded and the first utterance used instead. The total number of tokens analyzed was (18 sentences × 6 repetitions × 5 subjects=) 540.

Measurements were obtained by using manual cursors for durations of (1) every segment of a verb (phrase) that was measurable; (2) the VOTs for voiceless plosives; (3) the entire utterance. In each pair of sentences that have the same segmental structure, duration of every segment or a sequence of segments that constitute the sentences was measured as long as segmental boundaries were distinct enough to be determined.

The criteria for determining segmental boundaries were the same as those used in all the previous experiments of this study.

The end of each utterance was defined to be the point at which periodic waveform of the vowel ended as in Experiments 1 to 4. It should be mentioned here that the prepausal duration obtained by this criterion is not exactly the same as segmental duration with no following pause. To be more specific, since the intensity most likely decreases towards the end of a sentential utterance, there is a possibility that the perceived duration of the utterance-final vowel is shorter than its measured duration. All the segmental durations were repeatedly measured to ascertain the accuracy and consistency of the measurements.

In addition to the duration measurement, the mean F0, the maximum and minimum F0 of the sentence-final vowel were measured by the same method as that employed in Experiments 3 and 4. In test sentences of Group B, the mean F0, the maximum and minimum F0 of the vowels of the particles /ga/ and /no/ were also measured.

4.3 Results & discussion

4.3.1 *Effect of accent on the duration of sentence-final vowels*

Figures 4-1, 4-2, 4-3, and 4-4 illustrate the mean duration of segments that constitute each pair of test sentences. (The error bar shows a standard error above or below the mean. The asterisks ** and * at the bottom of each figure indicate significant differences at the levels of $p<0.0001$ and $p<0.05$, respectively, by RM-ANOVAs.) The two sentences in each figure consist of the same segments; the only difference between them is the presence/absence of an accentual fall in verbs or verb phrases. As is clear from the figures, no large durational difference is observed between each pair of component segments except for sentence-final vowels. The durational

difference in vowels in sentence-final position ranges from 30 to 50ms.

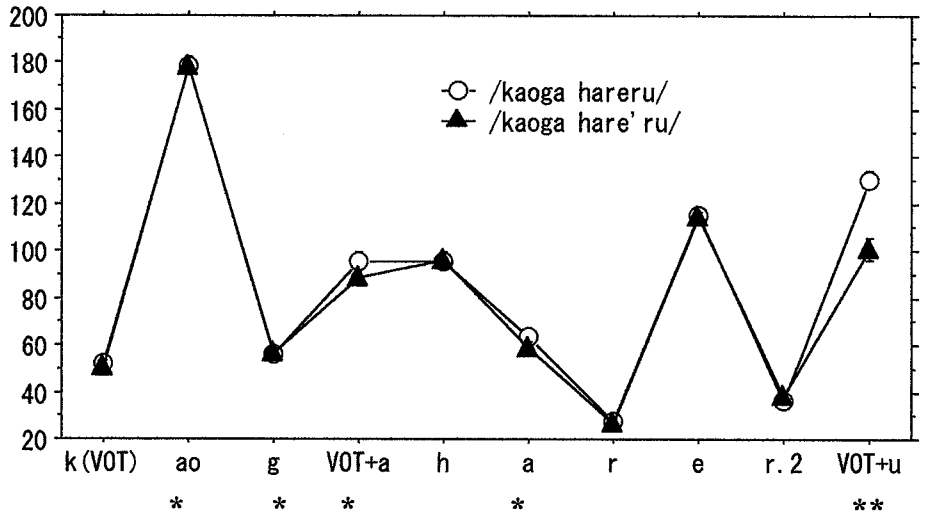


Fig. 4-1. Mean segmental duration (ms) of /kaoga hareru/ and /kaoga hare'ru/

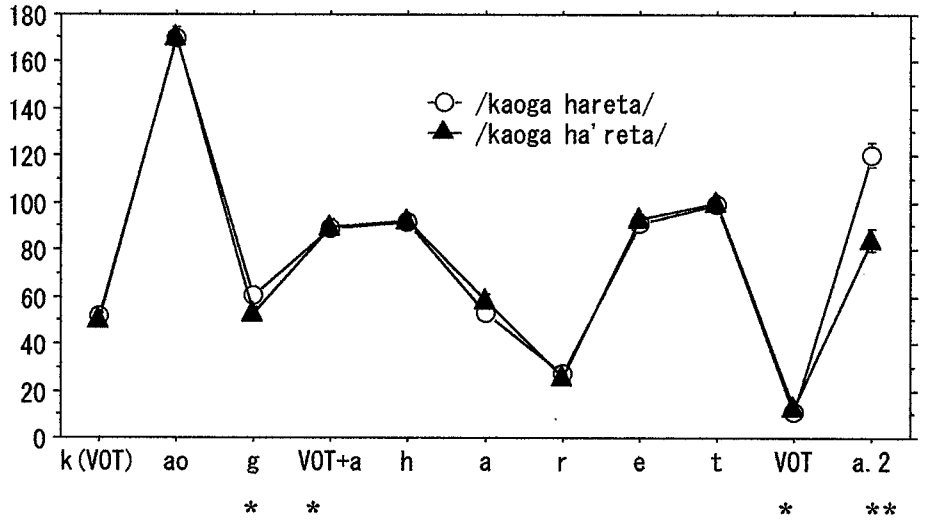


Fig. 4-2. Mean segmental duration (ms) of /kaoga hareta/ and /kaoga ha'reta/

Table 4-1 presents the mean duration and mean f_0 of sentence-final vowels for each test sentence. "Accent type" in the table is classified into four types (*a, a', b,* and *b'*), and is expressed, for example, as 1 + 0, where each number represents the number of

accentual falls in the preceding NP and the final VP, respectively. Thus, "1" means that there is an accentual fall in a NP or VP while "0" indicates the absence of an accentual fall. "D" in the table represents the difference in duration of each sentence-final vowel from the shortest vowel for each set. The results of RM-ANOVAs for each set are also given.

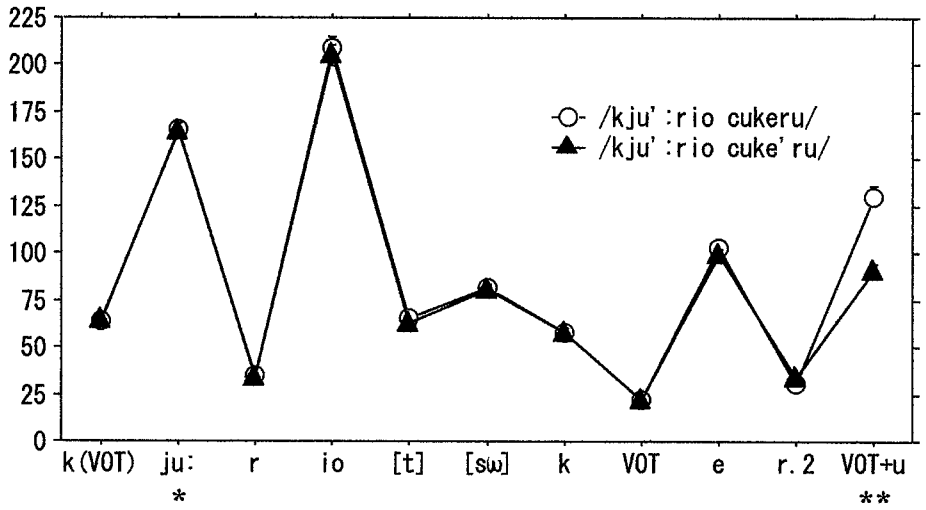


Fig. 4-3. Mean segmental duration (ms) of /kju':rio cukeru/ & /kju':rio cuke'ru/

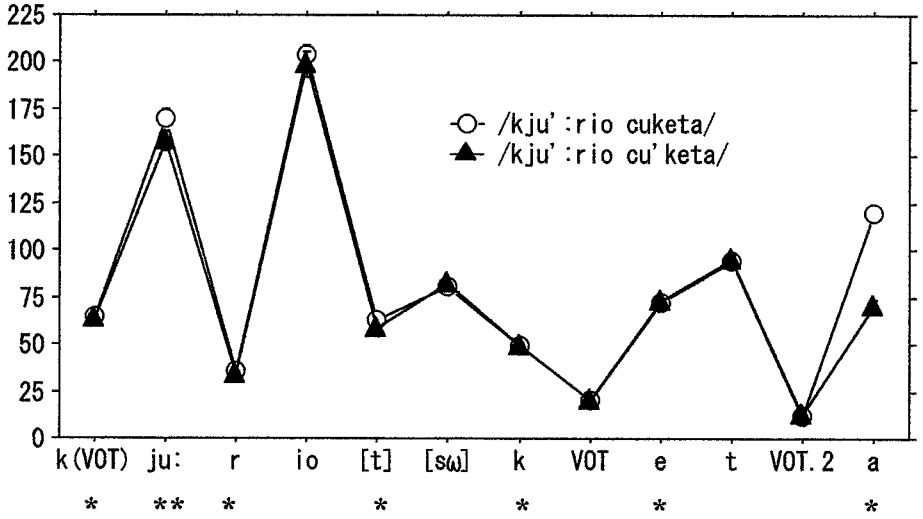


Fig. 4-4. Mean segmental duration (ms) of /kju':rio cuketa/ & /kju':rio cu'keta/

Table 4-1. Duration (ms) and mean F0 (Hz) of sentence-final vowel of each test sentence.

Test sentence	Accent type	Vowel	SD	D	RM-ANOVA	MeanF0
kao ga hareru	a. 0 + 0	129.9	22.2	34.6	F(1,29)=59.48	203.7(28.9)
ka'ta ga hareru	a'. 1 + 0	124.8	27.8	29.5	p<0.0001	175.8(25.2)
kao ga hare'ru	b. 0 + 1	100.6	26.4		a, a' > b, b'	156.3(20.4)
so'ra ga hare'ru	b'. 1 + 1	95.3	26.9			147.9(27.0)
kao ga hareta	a. 0 + 0	120.4	28.7	36.4	F(1,29)=130.85	206.1(29.8)
ka'ta ga hareta	a'. 1 + 0	113.9	30.2	29.9	p<0.0001	175.3(22.4)
kao ga ha'reta	b. 0 + 1	84.0	26.6		a, a' > b, b'	148.8(14.5)
so'ra ga ha'reta	b'. 1 + 1	85.5	26.7			143.4(25.5)
kyu'uri o tukeru	a'. 1 + 0	130.2	28.9	38.1	F(1,28)=71.57	174.3(19.8)
kyu'uri o tuke'ru	b'. 1 + 1	92.1	21.2		p<0.0001	149.1(28.1)
kyu'uri o tuketa	a'. 1 + 0	116.9	26.6	48.3	F(1,29)=110.0	177.6(20.8)
kyu'uri o tu'keta	b'. 1 + 1	68.6	19.0		p<0.0001	141.2(22.6)

We can see from the table that the sentence-final vowel tends to get longer when the verb (phrase) has no accent nucleus, i.e., no accentual fall. When the verb (phrase) is initially- or medially-accented, that is, when it has an accentual fall, the final vowel tends to be shorter. The difference in duration for each set is on average 36.1ms, ranging from 29.5 to 48.3ms. The ratio of the longer vowel duration to the shorter one is on average 1.43:1, ranging from 1.31:1 to 1.70:1.

Two-way RM-ANOVAs were run on the duration of sentence-final vowels for *hareru*-set and *hareta*-set, respectively, with the presence or absence of an accentual fall in a NP and that in a VP as two fixed factors. They showed a highly significant effect of the presence or absence of an accentual fall in a VP. There was no significant effect of the presence/absence of an accentual fall in a NP nor any significant interaction between the two factors.

In *tukeru*-set and *tuketa*-set, one-way RM-ANOVAs with the presence or absence of an accentual fall in a VP as a fixed factor showed the effect highly significant.

These results have revealed that a significant effect of accent on duration is specific to sentence-final position. In other words, the factor contributing to a highly significant difference in sentence-final vowel duration of short declarative sentences is the presence/absence of an accentual fall in the sentence-final verb phrase only. Whether there is an accentual fall or not in the preceding noun phrase does not significantly affect the duration of the sentence-final vowels.

The results have also made it clear that the decisive factor involved in the phenomenon is not actually the accent pattern, but the presence or absence of an accentual fall, since initially-accented verb phrases (e.g. *ha'reta, tu'keta*) do not differ in their effect on the duration of sentence-final vowels from medially-accented verbs (e.g. *hare'ru, tuke'ru*).

4.3.2 *Effect of accent on phrase-final vowels*

Table 4-2 shows the mean duration and mean F0 of vowels in phrase final position and sentence-final position, for each set of test sentences in Group B. Recall that each pair of sentences in Group B is exactly the same in their segmental structure. The only difference between the pair is the presence/absence of an accentual fall in the phrase-final nouns: One has no accent nucleus, the other is initially-accented.

As is obvious from the table, the presence/absence of an accentual fall in the preceding NPs does not result in any significant difference in vowel duration in either phrase-final or sentence-final position. The only significant difference is observed in the F0 values of the NP-final vowels for each set, apparently due to the

presence/absence of an accentual fall in the preceding nouns. The results also indicate that a statistically significant difference in F0 does not necessarily lead to durational difference. In other words, the effect of an accentual fall on the duration of final vowels is specific to sentence-final position. The sentence-final position is identical with utterance-final position and is a prepausal position in the tokens of the current experiment.

Table 4-2. Mean duration (ms), mean F0 (Hz) and SD (parentheses) of the underlined vowels for each set of test sentences

Test sentences	Phrase final position			Sentence final position		
	Vowel	SD	F0	Vowel	SD	F0
…atusani tomado'tta	67.8	12.9	**248.5(35.1)	85.7	33.4	140.9(23.3)
…a'tusani tomado'tta	67.4	15.6	179.0(19.3)	89.0	34.7	139.2(22.6)
sakega…oisiso'oda	81.5	10.9	**265.9(16.1)	120.0	29.7	142.6(27.3)
sa'kega…oisiso'oda	79.8	12.8	193.0(10.8)	120.1	33.2	140.3(28.9)
…siroga mi'eta	92.4	10.5	**229.8(26.6)	98.4	34.5	141.6(24.8)
…si'roga mi'eta	89.1	10.8	189.2(27.6)	101.7	37.1	140.2(23.3)

** and * indicate statistic differences at the levels of $p<0.0001$ and $p<0.05$, respectively, by RM-ANOVA.

Figures 4-5 and 4-6, for instance, demonstrate little or no accent-related difference in segmental duration between test sentences in Group B that are composed of the same segments. The vowel /i/ of /sirogamieta/ is the only segment in the test sentences in Group B that showed a significant difference in duration: The accented vowel /i/ is about 8.5ms longer than its unaccented counterpart[1].

However, the accent-related effect is restricted to this vowel, and the magnitude of the effect is far smaller than the one observed in the sentence-final vowels in test sentences with accentual contrast in their VPs.

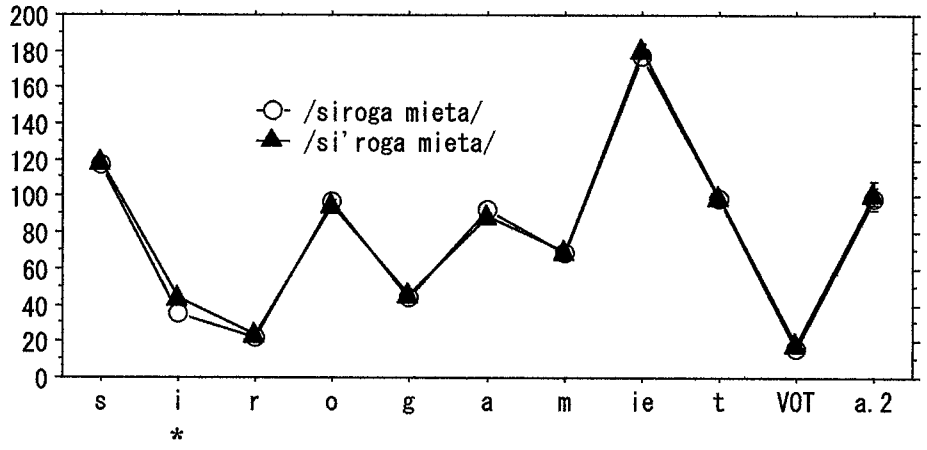


Fig. 4-5. Mean segmental duration (ms) of /siroga mieta/

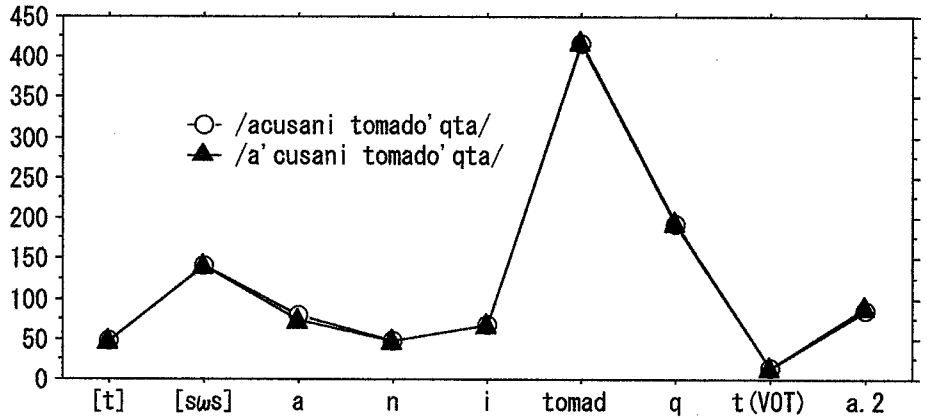


Fig. 4-6. Mean segmental duration (ms) of /acusani tomado'qta/ and /a'cusani tomado'qta/

Consequently, it can be said that the presence/absence of an accentual fall in the phrase-final position followed by no pause results in no consistent durational difference in the sentential utterances.

4.3.3 *Correlation between F0 and duration in sentence-final vowels*

In Figure 4-7, the durations of all sentence-final vowels in Group A are plotted against their mean F0s. Black squares represent the sentence-final vowels of initially- or medially-accented verbs/verb

phrases. White circles represent the sentence-final vowels of verbs or verb phrases without an accentual fall. We can see from the figure a positive correlation between vowel duration and mean F0. That is, there is a tendency for a sentence-final vowel to get longer as its F0 gets higher.

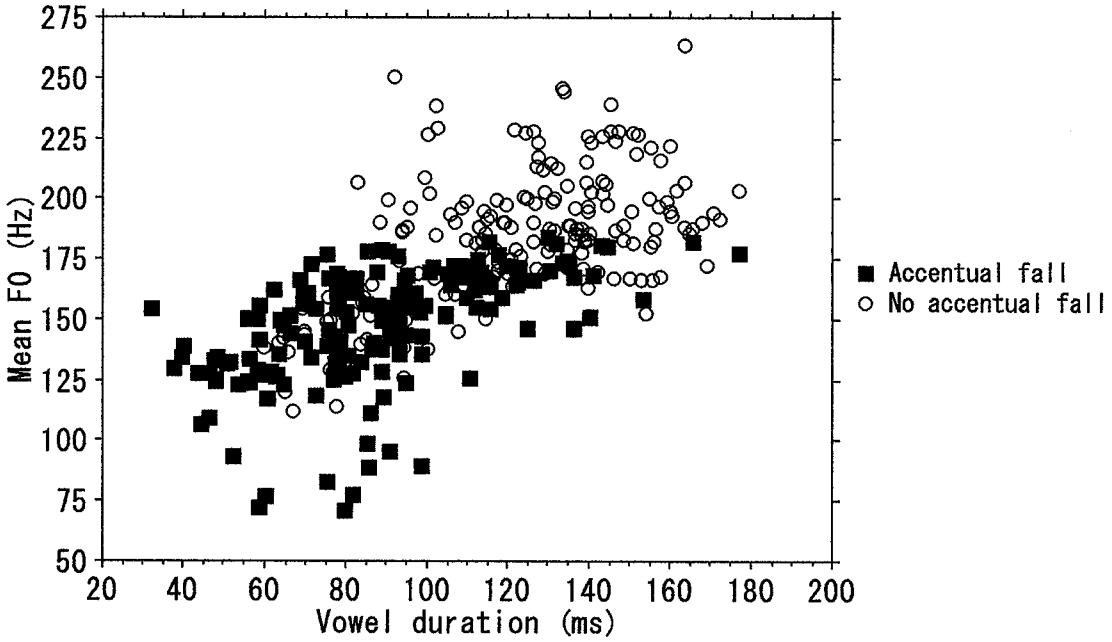


Fig. 4-7. Mean duration of sentence-final vowels (ms) plotted against their F0s (Hz)

It is also noticeable that when the sentence-final verb (phrase) has an accentual fall the final vowel tends to have shorter duration and lower F0 than it does when the sentence-final verb (phrase) has no accentual fall.

Correlation coefficients of the duration of a sentence-final vowel with its mean F0 were computed for all the tokens of Group A in Experiment 5. It was 0.69 collapsed over sets, ranging from 0.65 for *hareta*-set to 0.75 for *tuketa* set ($p<0.0001$). Considering the sample size of 350, it can be safely said that there is a strong correlation between the duration and F0 of sentence-final vowels.

These results are understandable since an accentual F0 fall in a sentence-final verb phrase causes an F0 drop towards the end of the sentence. That is, an accentual F0 fall in the sentence-final verb phrase and the mean F0 of the sentence-final vowel are closely related to each other.

In sum, the F0 values of sentence-final vowels were found to be closely related to their duration, suggesting a general tendency for the sentence-final vowel to become shorter as its F0 gets lower.

4.4 Experiment 6

4.4.1 *Subjects*

The subjects were six middle-aged females, all native speakers of Tokyo (standard) Japanese. Three of them had participated in Experiment 5.

4.4.2 *Materials*

Test sentences were seven sets of 26 real Japanese sentences listed in Appendix 4-2. In *atusa*-set, for instance, the first pair of test sentences has *atusa-ga* and *a'tusa-ga* at the phrase-final position of NPs followed by pauses. The second pair of the test sentences has *atusa-da* and *a'tusa-da* at the sentence-final position. Thus, in the phrase-final or sentence-final position, the pair of test sentences contains nouns that are the same in segmental make-up but different in accent pattern.

The phrase-final and sentence-final positions in this experiment are followed by pauses, that is, they are prepausal. However, the former is not in a sentence-final position, whereas the latter is. In this experiment, the two prepausal positions are referred to as sentence-final and non-sentence-final.

Each test sentence was written in the Japanese orthography (*kanji* and *kana),* one sentence per card. The total of 26 test sentences were then mixed with 20 other test sentences/phrases that will be used for other experiments.

In sentence-final position, the consonant preceding a prepausal vowel /a/ is /d/ of /da/, whereas it is /g/ of /ga/ in non-sentence-final position. Thus, it is possible that the change of the preceding consonant from /g/ to /d/ may affect the duration of the prepausal vowel. Nonsense words that were made up of the same two syllables as the ones in prepausal position for each set of test sentences were prepared. They were *saga* and *sada* for *atusa*-set, *siga* and *sida* for *hasi*-set, *kega and keda* for *sake*-set, and *roga* and *roda* for *siro*-set. Ten repetitions of each word written on a sheet in random order were elicited in order to examine whether the duration of the target vowel /a/ is affected by the segmental change of the preceding consonant from /g/ to /d/.

4.4.3 *Procedure*

Each subject was presented a test sentence on a card one at a time in a random order and asked to read the sentence naturally at a tempo comfortable to herself without placing any focus or emphasis on a particular constituent. She was also asked to read each sentence inserting a pause after each punctuation mark if she did not find it unnatural or difficult. All the subjects answered that the pause insertion was easy and natural to them. Every time the subject finished reading all the sentences, the cards were randomized and presented to the subject, for a total of seven different randomizations. Next, a list of nonsense words composed of a pair of prepausal syllables for each set was read by each subject to produce a total of ten tokens for each nonsense word.

Productions were recorded and digitized at a sampling rate of 11 kHz, and waveforms, wideband spectrograms, and pitch contours were made with a Kay Multi-Speech Model 3700. As a rule, the tokens from the second to the seventh reading were measured. Only when any of the tokens was misread or wrong in accent assignment, was the token discarded and the first utterance was used instead. The total number of tokens analyzed was [(26 sentences × 6 repetitions × 6 subjects) + (8 words × 10 repetitions × 6 subjects)=] 1416.

Measurements were obtained by using manual cursors for durations of (1) every segment of a prepausal noun phrase that was measurable; (2) the VOTs for voiceless plosives; (3) the entire length of each PBU; (4) the duration of prepausal vowels /a/ in tokens of nonsense words. In each pair of sentences that had the same segmental content, duration of every segment or a sequence of segments was measured as far as segmental boundaries were distinct enough to be determined.

The criteria for determining segmental boundaries and the end of each PBU were the same for the previous experiments (Experiments 1 to 5).

All the segmental durations were repeatedly measured to ascertain the accuracy and consistency of the measurements.

In addition to the duration measurement, the mean F0, the maximum and minimum F0 of a prepausal vowel and the penultimate vowel, and the max F0 of a pause-bounded-unit (PBU) were measured by the same method as that employed in Experiments 3 to 5.

4.5 Results & discussion

4.5.1 *Effect of accent and position on the duration of prepausal vowels*

Figure 4-8 shows the mean duration of prepausal vowels /a/, /e/, or /o/ for each accent pattern and position for all the test sets. Sentence-final-position and non-sentence-final position are abbreviated into "SF" and "NonSF", respectively, while "A" and "NoA" represent an accentual fall and no accentual fall, respectively. Accordingly, what the abbreviations of X bars in the figures and discussions hereafter represent is the four combinations of accent (the presence or absence of an accentual fall) and position (sentence-final or non-sentence-final position) as follows:

NonSF-NoA: non-sentence-final position, with no accentual fall

NonSF-A: non-sentence-final position, with an accentual fall

SF-NoA: sentence-final position, with no accentual fall

SF-A: sentence-final position, with an accentual fall

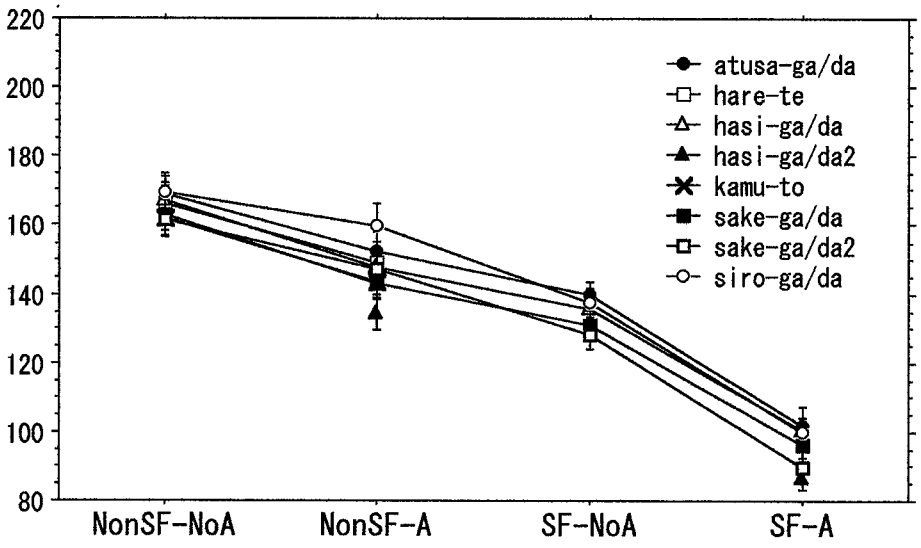


Fig. 4-8. Effect of accent and position on duration of prepausal vowels (ms)

The most obvious feature of Fig. 4-8 is that all the durations of

prepausal vowels for the eight sets show a similar durational variation as a function of the accent and position of the final noun phrases. In both non-sentence-final and sentence-final positions, prepausal vowels of the noun phrase with no accentual fall are longer than their counterparts with an accentual fall, suggesting the effect of an accentual fall on vowel duration. In addition, it is evident from the figure that the accent-related durational difference in NonSF is much smaller than that in SF.

Another obvious finding from the figure is that the prepausal vowels in non-sentence-final position are longer than those in sentence-final position whether or not the prepausal noun phrase has an accentual fall. This suggests the effect of position (sentence-final vs. non-sentence-final) on the duration of prepausal vowels.

Table 4-3 presents the mean duration of prepausal vowels for each test sentence and the results of RM-ANOVAs and post hoc tests. The two numbers combined by plus "+" in "Accent type" each represents the number of accentual falls in the preceding phrase and in the final (prepausal) noun phrase, respectively, for each test sentence. Whether the prepausal position is sentence-final (SF) or non-sentence-final (NonSF) is also given.

It can be seen from the table that, in sentence-final position, the presence or absence of an accentual fall in the final noun phrase has resulted in a large difference in the duration of prepausal vowels; the difference is on average 39.5ms, ranging from 34.6ms to 49.2ms. The ratio of the longer vowel to the shorter one for each set is on average 1.42:1, ranging from 1.36:1 to 1.57:1. (Appendix 4-3 presents each ratio of prepausal vowels.) This ratio of accent-related lengthening is almost the same with those in isolated words, phrases, and short sentences.

4-3. Mean duration (ms) of prepausal vowels of test phrases with different accent patterns. "P * A" represents an interaction between the effects of position and accent.

Target item	Position	Accent type	Vowel duration	SD	RM-ANOVA & post hoc test
1) atusaga	NonSF	*a.* 0 + 0	**168.9**	30.4	Accent: F(1,35)=49.54, p<0.0001
2) a'tusaga	NonSF	*b.* 0 + 1	**152.3**	34.4	Position: F(1,35)=79.76, p<0.0001
3) atusada	SF	*a.* 0 + 0	**139.8**	23.3	P * A: F(1,35)=24.88, p<0.0001
4) a'tusada	SF	*b.* 0 + 1	**101.6**	33.1	1) > 2) > 3) > 4) p<0.05
1) hasiga	NonSF	*a'.* 1 + 0	**167.0**	30.6	Accent: F(1,35)=55.40, p<0.0001
2) ha'siga	NonSF	*b'.* 1 + 1	**147.5**	26.9	Position: F(1,35)=181.36, p<0.0001
3) hasi'ga	NonSF	*b'* 1 + 1	**134.3**	22.6	P * A: F(1,35)=8.57, p=0.006
4) hasida	SF	*a'.* 1 + 0	**135.7**	31.4	2), 3), 4): non-sig.
5) ha'sida	SF	*b'.* 1 + 1	**100.5**	22.2	5), 6): non-sig.
6) hasi'da	SF	*b'* 1 + 1	**86.5**	16.5	1) > 2)/3)/4) > 5)/6)
1) sakega	NonSF	*a.* 0 + 0	**162.9**	26.0	Accent: F(1,35)=79.29, p<0.0001
2) sa'kega	NonSF	*b.* 0 + 1	**143.1**	19.9	Position:F(1,35)=185.86, p<0.0001
3) sakeda	SF	*a.* 0 + 0	**130.9**	26.7	P * A: F(1,35)=8.07, p=0.0075
4) sa'keda	SF	*b.* 0 + 1	**96.3**	24.0	1) > 2) > 3) > 4) p<0.01
1) sakega2	NonSF	*a'.* 1 + 0	**161.7**	32.0	Accent: F(1.35)=59.44, p<0.0001
2) sa'kega2	NonSF	*b'.* 1 + 1	**147.0**	34.5	Position: F(1,35)=91.10, p<0.0001
3) sakeda2	SF	*a'.* 1 + 0	**128.3**	25.3	P * A: F(1.35)=14.54, p=0.0005
4) sa'keda2	SF	*b'.* 1 + 1	**90.0**	28.3	1) > 2) > 3) > 4) p<0.01
1) siroga	NonSF	*a'.* 1 + 0	**169.4**	33.1	Accent: F(1,35)=62.77, p<0.0001
2) si'roga	NonSF	*b'.* 1 + 1	**159.8**	38.0	Position: F(1,35)=121.83, p<0.0001
3) siroda	SF	*a'.* 1 + 0	**137.5**	25.9	P * A: F(1,35)=24.76, p<0.0001
4) si'roda	SF	*b'.* 1 + 1	**100.1**	19.1	1) > 2) > 3) > 4) p<0.05
1) kamuto	NonSF	*a.* 0 + 0	**162.1**	31.1	F(1,35)=27.03
2) ka'muto	NonSF	*b.* 0 + 1	**143.4**	30.0	p<0.0001
1) harete	NonSF	*a.* 0 + 0	**166.1**	31.8	F(1,34)=15.23
2) ha'rete	NonSF	*b.* 0 + 1	**149.3**	35.4	p=0.0004

In non-sentence-final position, on the other hand, the accent-

related difference is on average 18.1ms, ranging from 9.6ms to 32.7ms. The ratio of the vowel length in target phrases with no accentual fall to its counterpart with an accentual fall is on average 1.13:1, ranging from 1.06:1 to 1.24:1. It turns out that in non-sentence-final position, i.e., in non-utterance-final position, the accent-related durational difference in prepausal vowel is much smaller than that in utterance-final position.

When a comparison is made between prepausal vowels with the same accent pattern in non-sentence-final and sentence-final positions, the ratio of position-related lengthening is obtained. The ratio is on average 1.24:1 for the target phrase with no accentual fall, ranging from 1.21:1 to 1.26:1; in the case that the target phrase has an accentual fall, the lengthening averages 1.54:1, ranging from 1.47:1 to 1.63:1. The figures reveal that the degree of the position-related lengthening in NonSF-A (54%) is a little greater than that of the accented-related lengthening (42%).

The maximum ratio of lengthening resulted from the effects of accent and position is obtained for each set by dividing the prepausal vowel duration in NonSF-NoA by that in SF-A. The average ratio is as large as 1.74:1, ranging from 1.66:1 to 1.93:1.

Two-way RM-ANOVAs were run on the duration of prepausal vowels with accent (the presence/absence of an accentual fall in the prepausal noun phrase) and position (sentence-final/non-sentence-final position) as two fixed factors. They showed highly significant main effects of accent and position as well as their interaction in all the sets. The results make it clear that the duration of a prepausal vowel differs significantly depending on the accent and position of the vowel. In addition, the effects of accent and position interact with each other. By this, it is meant that the effect of accent is influenced by that of position and is weakened at the non-sentence-final position.

Similarly, the presence or absence of an accentual fall affects the effect of position, and the magnitude of the position-related effect is larger when a prepausal noun phrase has an accentual fall.

Post hoc tests in Table 4-3 indicated that, in *atusa*-, *sake*-, *sake*2-, and *siro*-sets, the duration order of prepausal vowels from the longest to the shortest is 1) > 2) > 3) > 4) and the difference between any neighboring durations is significant at the $p<0.05$ level. In other words, as can be seen from Fig. 4-8, the prepausal duration in NonSF-NoA is significantly longer than that in NonSF-A, which is significantly longer than that in SF-NoA, which then significantly differs from that in SF-A (NonSF-NoA > NonSF-A > SF-NonA > SF-A).

If we consider the sentence-final vowel of a noun phrase with no accentual fall (SF-NoA) to be lengthened compared with its counterpart with an accentual fall (SF-A), it can be safely said that the prepausal vowels in NonSF-A are "lengthened" compared with its counterpart in SF-A.

In *hasi*-set, the prepausal vowels of initially-accented phrase (*ha'siga* & *ha'sida*) and finally-accented phrase (*hasi'ga* & *hasi'da*) do not differ significantly both in non-sentence-final and sentence-final positions. The result indicates that the initially- and finally-accented words are similar in their durational patterns of prepausal vowels.

4.5.2 *Effect of preceding consonants on prepausal vowels*

The previous section has made a comparison between the durations of prepausal vowels in sentence-final and non-sentence-final positions. Although the difference is statistically significant, there is a possibility that the change of the preceding vowel from /g/ to /d/ may contribute to the position-related difference in duration.

Table 4-4 shows the mean duration of prepausal vowels in

nonsense word utterances composed of the last two syllables before the pause for each test sentence. RM-ANOVAs with the differing preceding consonant as a fixed factor showed significant differences in the duration of the prepausal /a/ in *sake-* and *siro*-sets at the $p<0.05$ level. The absolute durational difference is 6.7ms and 5.5ms, respectively. The other two sets (*atusa*-set and *hasi*-set) indicate no significant difference, although the duration of vowel /a/ in /ga/ is longer than its counterpart in /da/ by 1.5ms and 2.8ms, respectively.

Table 4-4. Mean duration (ms) and SD of the prepausal vowel /a/ in each test word

Test set	Test word	/a/ in /ga/(ms)	SD	Test word	/a/ in /da/(ms)	SD	RM-ANOVA
atusa-set	saga	200.1	51.8	sada	198.6	50.9	F(1,59)=0.30 non-sig.
hasi-set	siga	196.3	44.2	sida	193.5	48.9	F(1,59)=0.95 non-sig.
sake-set	kega	180.8	38.9	keda	174.1	38.3	F(1,57)=6.37 p<0.05
siro-set	roga	191.5	44.2	roda	186.0	40.4	F(1,59)=5.27 p<0.05

In sum, it can be said the prepausal vowel /a/ preceded by /g/ tends to be a little longer than that preceded by /d/. This finding suggests that segmental change from /g/ to /d/ may partly contribute to the durational difference of the following /a/ between sentence-final and non-sentence-final positions. However, Table 4-4 indicates that the magnitude of the consonant-related effect on the duration of prepausal /a/ falls within a range of 1-7ms. Thus, it is evident that the position-related durational difference that ranges from 30ms to 60ms observed in the present experiment cannot be accounted for solely by the variation of the preceding consonants.

4.5.3 *Effect of position on the duration of penultimate vowels*

All the previous experiments (Experiments 1 to 5) show no consistent accent-related effect on segment duration except for the effect on utterance-final vowels. The effect of position, however, does not appear to be restricted to utterance-final vowels.

For instance, Figure 4-9 shows the mean segment duration of /no *sake*-ga/ and /no sake-da/ of *sono tugitate-no sake-ga/da* and *sono toretate-no sa'ke-ga/da* in *sake*-set. As has been discussed in 4.5.1,

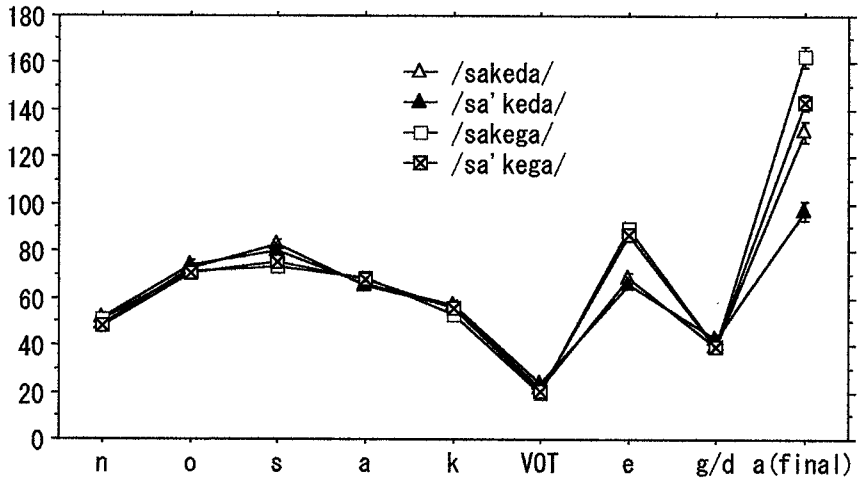


Fig. 4-9. Mean segmental duration (ms) of *no sakega/da* and *no sa'kega/da*

the four final (prepausal) vowels /a/ in Fig. 4-9 show a substantial variation in duration depending on their accent and position. Contrary to this, the duration of the preceding consonants /d/ and /g/ hardly differ among the four sentences. However, we can see from the figure that the vowel /e/ in /sake/ produced in non-sentence-final position has longer duration than that same vowel produced in sentence-final position. This suggests a position-related lengthening not only in the final (prepausal) vowel but also in the [2]penultimate vowel in *sake*-set. The same tendency is observed in the other test sets.

Figures 4-10 and 4-11 illustrate the mean duration of the preceding consonants and penultimate vowels depending on accent and position for all the test sentences. The consonants in Figure 4-10 show no consistent durational variation related to accent and position. The penultimate vowels in Figure 4-11, by contrast, are longer in non-sentence-final position than their counterparts in sentence-final position in all the test sets that have sentence-final and non-sentence-final contrasts.

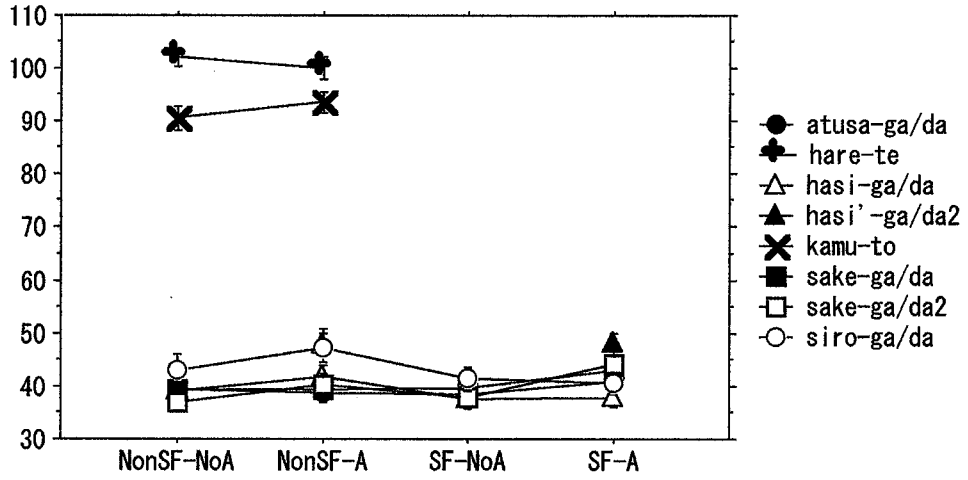


Fig. 4-10. Effect of accent and position on preceding consonants (ms)

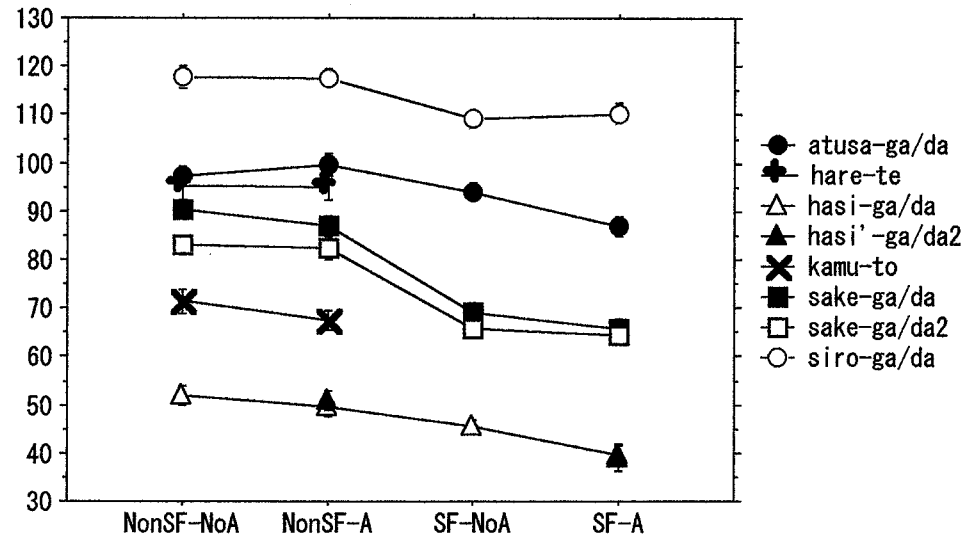


Fig. 4-11. Effect of accent and position on penultimate vowels

Table 4-5 demonstrates the effect of position (non-sentence-final vs. sentence-final) on the mean durations of penultimate vowels, preceding consonants, prepausal vowel, and pause-bounded units (PBUs). The durational difference between each pair of segments produced in non-sentence-final position and in sentence-final position

Table4-5. Effect of position on the component duration (ms) and SD (parentheses)

Target phrase	Penulti-mate vowel	D	Preceding conso-nant	D	Prepausal vowel	D	PBU	D
atusaga	97.2(11.4)	3.1	39.5(11.7)	1.0	168.9(30.4)	29.1	1206.8(89.7)	32.1
atusada	94.1(10.3)		38.5(8.6)		139.8(23.3)	**	1174.7(111.4)	*
a'tusaga	99.6(12.9)	12.5	38.9(11.1)	-1.9	152.3(34.4)	50.7	1184.6(77.9)	44.9
a'tusada	87.1(11.7)	**	40.8(9.2)		101.6(33.3)	**	1139.7(103.8)	*
hasiga	51.9(12.3)	6.5	38.9(12.9)	1.5	167.0(30.6)	31.3	1506.3(113.5)	45.0
hasida	45.4(8.8)	*	37.4(10.4)		135.7(31.4)	**	1461.3(97.5)	*
ha'siga	49.4(11.0)	9.8	41.7(13.3)	3.9	147.5(26.9)	47.0	1494.1(96.0)	83.7
ha'sida	39.6(13.1)	**	37.8(10.4)		100.5(22.2)	**	1410.4(100.4)	**
hasi'ga	50.8(9.1)	11.9	47.7(15.4)	-0.3	134.3(22.6)	47.8	1472.0(101.9)	75.3
hasi'da	38.9(12.7)	**	48.0(9.6)		86.5(16.5)	**	1396.7(143.7)	**
sakega	89.8(10.5)	21.0	39.3(11.0)	-3.3	162.9(26.0)	32.0	1440.1(103.9)	50.3
sakeda	68.8(12.2)	**	39.6(9.8)		130.9(26.7)	**	1389.8(106.8)	**
sa'kega	87.0(11.7)	21.6	39.3(7.5)	-3.8	143.1(19.9)	46.8	1379.3(113.2)	55.2
sa'keda	65.4(12.5)	**	43.1(9.9)	*	96.3(24.0)	**	1324.1(119.2)	**
sakega2	82.4(11.4)	17.0	37.3(12.4)	-1.0	161.7(32.0)	33.4	1535.2(119.4)	54.5
sakeda2	65.4(10.7)	**	38.3(11.2)		128.3(25.3)	**	1480.7(163.2)	**
sa'kega2	81.6(13.8)	17.2	40.8(13.3)	-3.4	147.0(34.5)	57.0	1521.4(132.6)	77.0
sa'keda2	64.4(9.9)	**	44.2(10.0)	*	90.0(28.3)	**	1444.5(159.9)	**
siroga	117.8(13.3)	8.7	43.1(17.6)	1.8	169.4(33.1)	31.9	1496.4(94.7)	48.7
siroda	109.1(11.0)	**	41.3(13.5)		137.5(25.9)	**	1447.7(113.9)	**
si'roga	117.5(10.7)	7.3	47.1(16.3)	6.6	159.8(38.0)	59.7	1473.9(103.0)	57.6
si'roda	110.2(13.3)	*	40.5(11.9)	*	100.1(19.1)	**	1416.3(117.1)	**

* and **indicate significant effects of position on duration at the levels of $p<0.05$ and $p<0.0001$, respectively, by one-way RM-ANOVAs. D signifies the durational difference between the two durations.

is also given. As is obvious from the table, position has a lengthening effect on penultimate vowels for each pair of sentences, whereas it has no consistent effect on the duration of preceding consonants. That is, the penultimate vowels in non-sentence-final

position are consistently longer than their counterparts in sentence-final position. The durational increase is on average 12.4ms, ranging from 3.1ms to 21.6ms.

Two-way RM-ANOVAs were run on the duration of penultimate vowels with accent and position as two fixed factors. In *sake*2-, and *siro*-sets, only the effect of position was significant ($p<0.0001$). In *sake*-set and *hasi*-set, the effects of position ($p<0.0001$) and accent ($p<0.01$) were significant. In *atusa*-set, the effect of position ($p<0.01$) and interaction between the effects of position and accent ($p<0.001$) were significant.

These results indicate that the effect of position on penultimate vowel duration is significant across the five test sets. That is, penultimate vowels are consistently longer in NonSF than their counterparts in SF. No other segments that constitute the test sentences showed a significant effect of position shared by all the test sets.

Another finding from Table 4-5 is that the length of a PBU is longer when it is followed by another PBU than when it constitutes the final part of a sentence. In other words, the preceding PBU is consistently and significantly longer than the following PBU. The difference ranges from 32.1ms to 83.7ms. However, the durational difference between the two PBUs for each pair of test sentences almost equals the sum of the durational differences in penultimate and prepausal vowels.

This suggests that the durational increase in the preceding PBU is manifested mostly in the prepausal vowels and the penultimate vowels. In the pairs *ha'siga/da* and *hasi'ga/da*, however, the difference in the duration of PBUs is about 20ms longer than the total difference observed in penultimate vowels, preceding consonants, and prepausal vowels. Fig. 4-12 shows durational

variation of /nohasi'ga/ and /nohasi'da/ in *hasi*-set. It is noticeable that not only prepausal and penultimate vowels but also /s/ in /hasi/ is lengthened by 10-15ms at the non-sentence-final position, suggesting its contribution to the position-related lengthening in the preceding PBU. However, this lengthening effect of position on the penultimate consonants is specific to the pairs of *ha'siga/da* and *hasi'ga/da*, suggesting segment-specific durational feature.

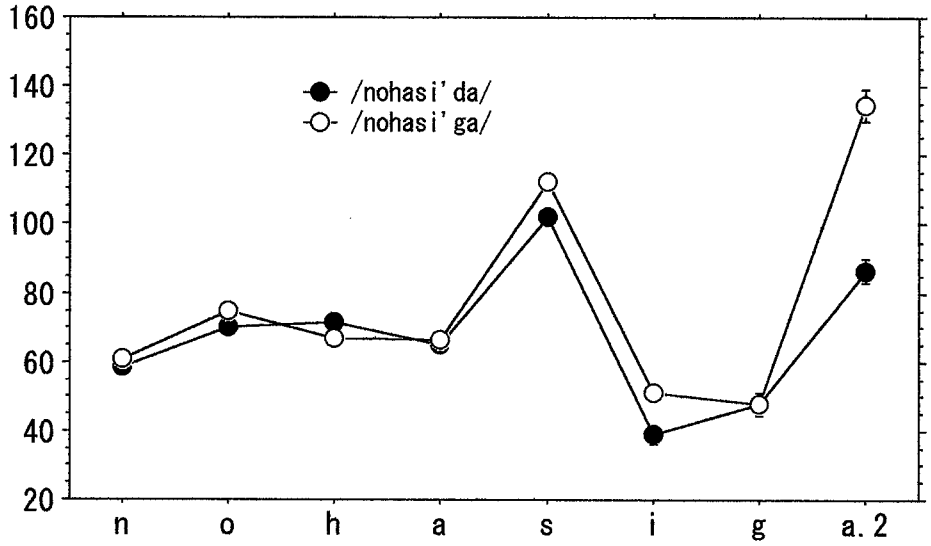


Fig. 4-12 Mean segmental duration (ms) of *(no)hasi'ga* and *(no)hasi'da*

In sum, the above results have made it clear that the position-related increase in duration is consistently observed in the penultimate vowels of all the test sets. This finding suggests that the lengthening in penultimate vowels partly contributes to the position-related durational increase in non-sentence-final position. In other words, position-related lengthening consistently observed in non-sentence-final position is manifested not only in the prepausal vowel but also in the penultimate vowels.

In English, final lengthening has been repeatedly reported at

various levels of prosodic boundaries (cf. Klatt, 1975; Edwards & Beckman, 1988). The most reliable domain of the lengthening has been claimed to be the rhyme of a syllable immediately preceding prosodic boundaries (Wightman et al. 1992). A question arises as to why in Japanese penultimate vowels consistently cooperate with the final vowels in the realization of position-related lengthening. It can be assumed that the penultimate vowels make a little contribution to the lengthening since a Japanese syllable does not usually have a coda except for a heavy syllable. The next question raised from this explanation is why the preceding consonant does not undergo this position-related lengthening in non-sentence-final position. As has been pointed out in 2.3.3, a greater elasticity in vowels than in consonants has been observed in Japanese (cf. Campbell 1999). It is assumed, therefore, that lengthening takes place both in the final vowel and the penultimate vowel to realize a substantial amount of prepausal lengthening in non-sentence-final position.

4.5.4 *Effect of accent and position on the pitch of prepausal vowels*

Figure 4-13 shows the mean F0 of prepausal vowels for each accent and position in each test set. It clearly indicates the effect of an accentual fall on the mean F0s in NonSF-A and SF-A, which are lower than those in NonSF-NoA and SF-NoA, respectively. However, the magnitude of the accent-related effect appears to be stronger in sentence-final position than in non-sentence-final position.

Two-way RM-ANOVAs were run on the mean F0s of prepausal vowels with the presence or absence of an accentual fall and sentence-final or non-sentence-final positions as two fixed factors. They showed significant main effects of accent and position, and a significant interaction between the two factors ($p<0.01$). In *atusa-*

set and *sake*-set, only the main effects of accent and position were significant ($p<0.0001$).

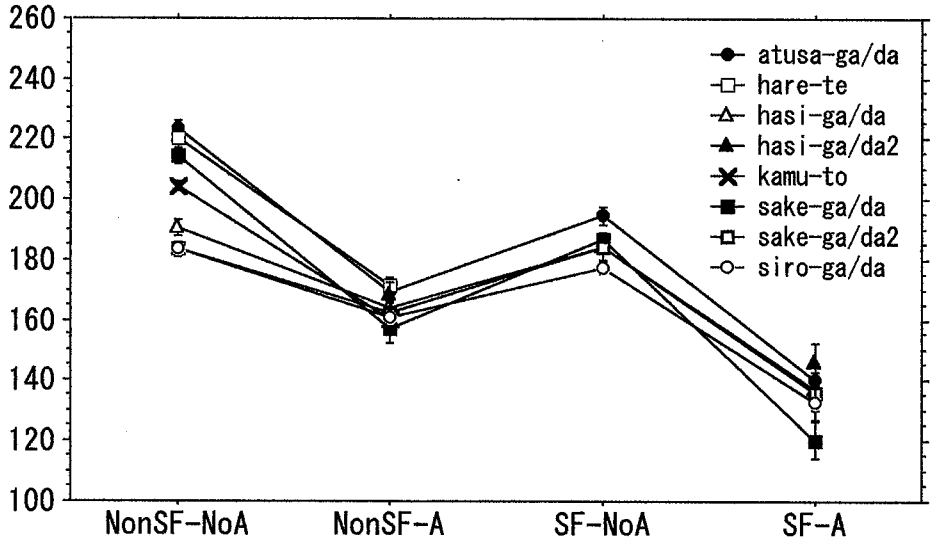


Fig. 4-13. Effect of accent and position on mean F0 (Hz) of prepausal vowels

The results indicate that prepausal vowels of NPs with an accentual fall have a significantly lower F0 than those with no accentual fall in both sentence-final and non-sentence-final positions. The prepausal vowels also have a significantly higher F0 in non-sentence-final position than in sentence-final position, whether or not the prepausal phrase has an accentual fall.

Figure 4-14 presents the amount of pitch drop in semitones between the penultimate vowels and the final (prepausal) vowels for each set. It shows that when prepausal NPs have an accentual fall, a sharp pitch drop takes place between the word-final vowels and the following particles, whereas a much smaller pitch drop is observed when the prepausal NP has no accentual fall. In addition, the magnitude of the pitch drop is much larger in SF-A than that in NonSF-A.

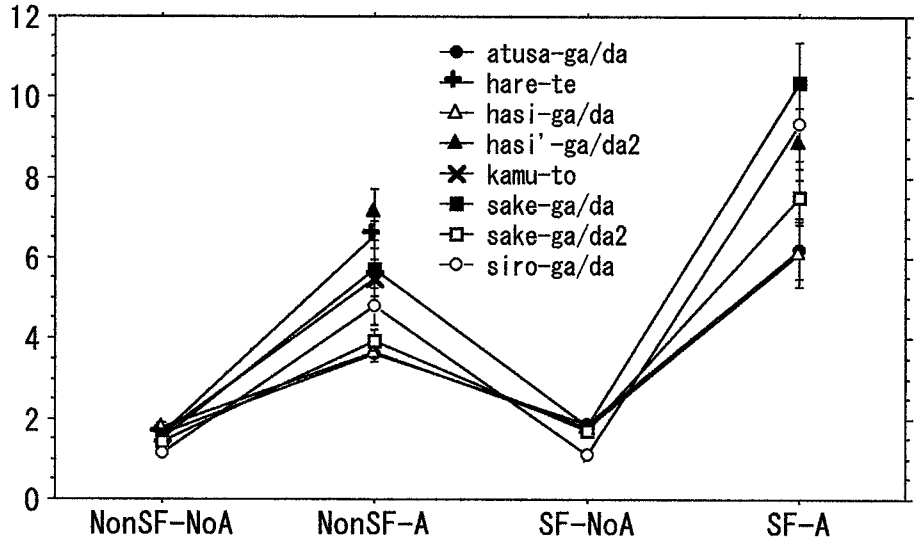


Fig. 4-14. Pitch drop (semitone) between penultimate and final vowels

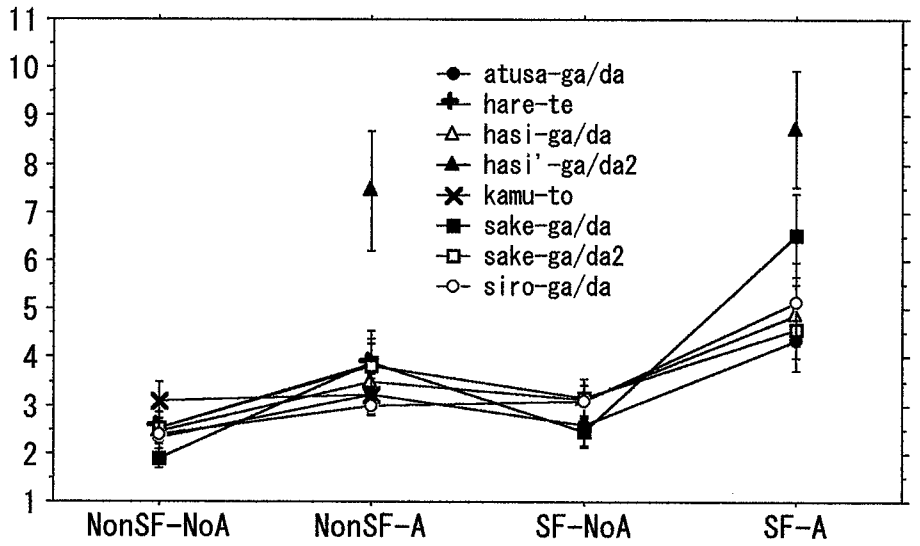


Fig. 4-15. Mean range (semitone) of the prepausal vowels

Figure 4-15 shows the amount of pitch change in semitones within the prepausal vowel for each accent type and position. One obvious finding from the figure is that in the sentence-final position, when the phrase has an accentual fall, pitch further drops within the prepausal vowel, whereas in non-sentence-final position it is not the

case. That is, the magnitude of the pitch drop within the prepausal vowels in NonSF-A does not differ greatly from those in SF-NoA or NonSF-NoA. This finding indicates that the accentual fall in the prepausal NP is mostly realized in the prepausal noun in non-sentence-final position, whereas it is realized in the noun and the following particle in the sentence-final position. It is also clearly shown that, in *hasi'ga* and *hasi'da* composed of finally-accented nouns suffixed with particles, a far larger pitch drop is realized within the following particle both in sentence-final and non-sentence-final positions.

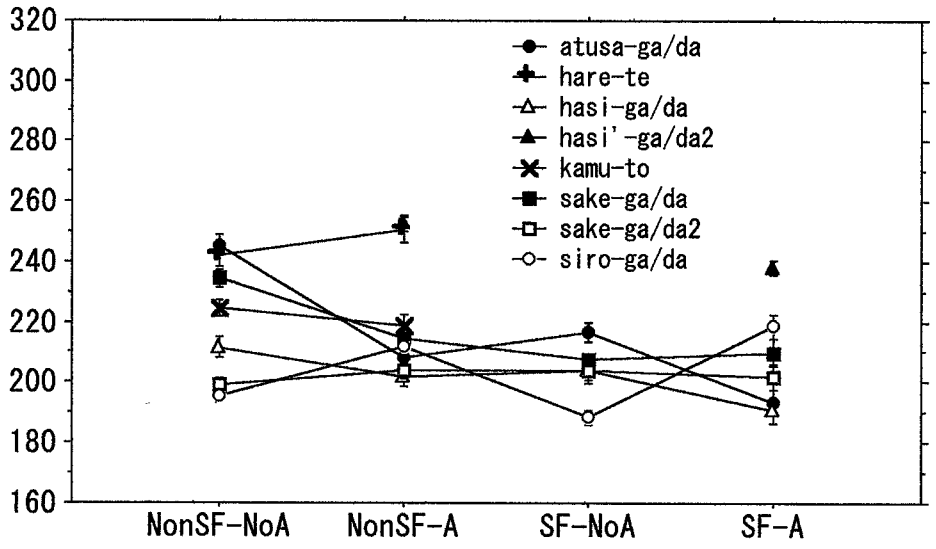


Fig. 4-16. Mean F0 (Hz) of penultimate vowels for each test sentence

Another figure, Figure 4-16, illustrates the mean F0 of penultimate vowels for each test sentence. That is, it shows the F0 of word-final vowels for each set. It is noticeable that there is a large variation in mean F0 across the test sets. In addition, no consistent difference in the F0s is observed depending on accent and position. To be more specific, the mean F0s of penultimate vowels in SF-A are not necessarily lower than those in SF-NonA and NonSF-A.

These findings suggest that the lowest mean F0 consistently

observed in Figure 4-13 in SF-A is realized not between the accented mora and the next mora but between the word-final mora and the following particle. In addition, Fig. 4-15 tells us that the large pitch drop within the vowel of the prepausal particle in SF-A also appears to greatly contribute to the realization of the lowest F0s. Moreover, these findings suggest that there is a strong tendency to realize a similar target F0 in prepausal positions except for NonSF-NoA regardless of a wide variation in the F0 of the penultimate vowels.

To sum up, it can be said that the magnitude and the domain of pitch drop induced by an accent nucleus in a prepausal phrase show a great difference between sentence-final position and non-sentence-final position. In non-sentence-final position, the pitch drop is mostly realized between an accented mora and the next mora, as has been generally believed to be a function of an accent nucleus. In sentence-final position, however, a relatively large pitch drop between the following unaccented mora and the sentence-final particle is consistently observed. Moreover, it has been revealed that pitch further drops sharply within the sentence-final mora.

4.5.5 *Correlation between F0 and duration in prepausal vowels*

In isolated words, phrases, and short sentences, a striking effect of the presence/absence of an accentual fall on the duration of utterance-final vowels has been consistently observed. In addition, a positive correlation has been shown between the mean F0 and duration of the utterance-final vowels.

In the present experiment which has revealed an effect of position, the relation between the presence/absence of an accentual fall, the mean F0 and duration of prepausal vowels appears to be more complicated. It is noted from Figures 4-8 and 4-13 that the

prepausal vowels in NonSF-A are lengthened substantially despite their low F0s.

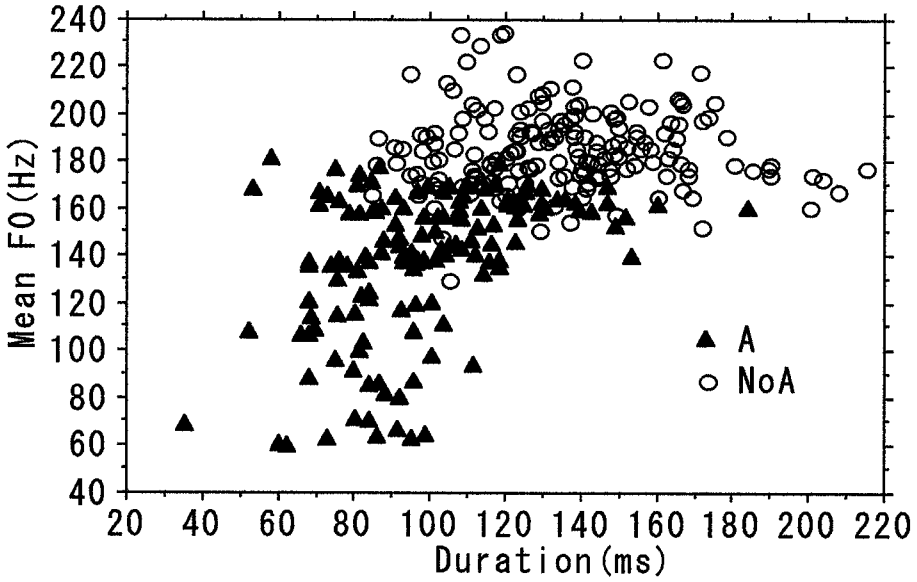


Fig. 4-17. Durations of prepausal vowels plotted against their F0s in sentence-final position

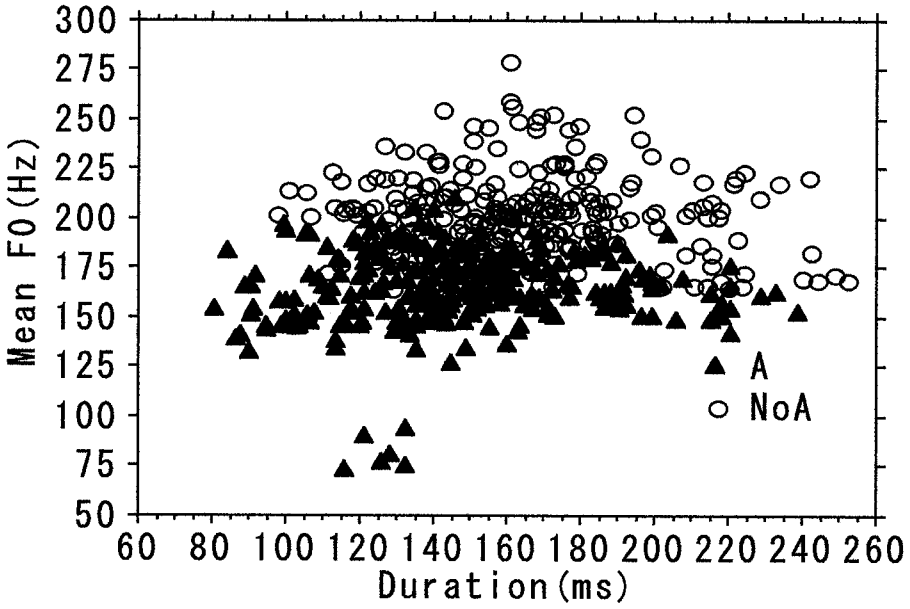


Fig. 4-18. Durations of prepausal vowels plotted against their F0s in non-sentence-final position

In Figures 4-17 and 4-18, all the mean F0s of the prepausal vowels from Experiment 6 are plotted against their durations in sentence-final position and in non-sentence-final position, respectively. In sentence-final position, a positive correlation is observed between the F0 and duration of prepausal vowels (Fig.4-17). In non-sentence-final position, however, the scatterplot shows no marked correlation between duration and mean F0 (Fig. 4-18).

Correlation coefficients of all the prepausal vowel durations with their mean F0s were 0.53 in sentence-final position ($p<0.0001$), and 0.19 in non-sentence-final position ($p<0.0001$). In sentence-final position the coefficients fall within the similar range to those observed in utterance-final position of isolated words, phrases, and sentences with no intervening pauses. The results from Experiments 3 to 5 showed coefficients ranging from 0.44 to 0.69 between the duration and F0 of utterance-final vowels.

Sentence-final position in the present experiment can be considered to be utterance-final, while non-sentence-final position is not identical with utterance-final position. Thus, the above results have clarified that a correlation between the F0 and duration of the final vowel is specific to utterance-final position. The utterance-final position is consistent with the word-final position in isolated word utterances or sentence-final position in isolated sentential utterances.

In non-sentence-final position, the presence or absence of an accentual fall in a NP at prepausal position leads to a significant and consistent durational difference in the prepausal vowel. However, the correlation coefficient (0.19) of the durations of prepausal vowels with their mean F0s hardly shows a correlation given the large sample number of 519. Thus, it can be safely said that the durational variation observed in the prepausal vowels in non-

sentence-final position is not related to their F0s.

In addition, the above durational increase in NonSF-A can be considered to be "lengthening" in a sense that it is significantly longer than its counterpart in sentence-final position (SF-A). This finding suggests an intentional position-related lengthening of prepausal vowels specific to non-sentence-final position.

In sum, all the results have clarified that the strong correlation between duration and F0 of final vowels is restricted to utterance-final position, the point at which a speaker's production of connected speech finishes. When an utterance does not finish, i.e., in non-utterance-final position, speakers significantly and consistently lengthen the prepausal and penultimate vowels regardless of the low F0s of the prepausal vowels. This suggests that there is an intentional position-related lengthening in non-utterance-final position. Finally, the results are suggestive of different duration control mechanisms for prepausal vowels in utterance-final and non-utterance-final positions.

4.6 Summary

Findings from these experiments can be summarized as follows:

(1) A strong effect of the presence or absence of an accentual fall in sentence-final VPs on the duration of sentence-final vowels is consistently observed in sentences with no intervening pauses. The effect is specific to sentence-final position, since the presence/absence of an accentual fall in the preceding NPs does not affect the duration of phrase-final and sentence-final vowels.

(2) Sentence utterances with pauses at major phrasal boundaries have shown the effects of accent (the presence/absence of an accentual fall) and position (sentence-final or non-sentence-final

position) as well as their interactions on the duration of prepausal vowels. In both non-sentence-final and sentence-final positions, prepausal vowels of the noun phrase with no accentual fall are significantly longer than their counterparts with an accentual fall. The ratio of the longer vowel to the shorter one is on average 1.42:1 in sentence-final position, whereas it stays as low as 1.13:1 in non-sentence-final position. In addition, the prepausal vowels in non-sentence-final position are significantly longer than those in sentence-final position. This position-related lengthening shows the ratio of 1.54:1 between the longer vowel and the shorter one when the prepausal phrase has an accentual fall. When it has no accentual fall, the ratio is on average 1.24:1. When the effects of accent and position get together, the ratio of the longest vowel to the shortest one for each set reaches 1.74:1 averaged across sets.

(3) A marked correlation between duration and F0 of final vowels are restricted to sentence-final (utterance-final) position. In non-utterance-final position, speakers significantly and consistently lengthen the prepausal and penultimate vowels regardless of the low F0s of the prepausal vowels. This suggests an intentional position-related lengthening in non-utterance-final position. In addition, the results are suggestive of different duration control mechanisms of the prepausal vowels in utterance-final and non-utterance-final positions.

---

[1] In Figs. 3-13 and 3-14 of Chapter 3, accented vowels preceded by fricative /s/ are significantly shorter than its unaccented counterparts.

[2] The term "penultimate" has two meanings: One is "the second from

the end of the word", and the other "the second from the end of anything". In this thesis, the term is used to refer to the position immediately preceding utterance-final or prepausal position.

CHAPTER 5

Prosodic control in declarative sentences

5.1 Introduction

Chapter 4 presents the data that clearly show the strong effect of accent and position on prepausal vowels in the reading of declarative sentences. The data have clarified the wide range of durational variation in prepausal position depending on accent and position. In addition, the prepausal vowel duration does not vary irregularly among the test sets and/or speakers, but show a similar durational pattern depending on accent and position.

From all these findings, this chapter makes a proposal for a descriptive prosodic control model of prepausal vowels in declarative sentences. The model attempts to predict a relative duration and pitch of any prepausal vowel in sentence-final position and in non-sentence-final position by only two factors: the presence/absence of an accentual fall in the preceding (non-prepausal) phrase and that in the final (prepausal) phrase.

5.2 Prosodic control of sentence-final vowels in short declarative sentences

As is evident from Experiment 4, in sentence-final position of the

short declarative test sentences, there is a highly significant difference in duration between the final vowel of a verb phrase with no accentual fall and that with an accentual fall. The experimental results have also made it clear that the factor that contributes to this consistent durational difference is the presence/absence of an accentual fall in the final verb phrase only (4.3.1). Whether there is an accentual fall or not in the preceding phrase does not significantly affect the duration of the sentence-final vowels (4.3.1; 4.3.2). Thus, the duration of final vowels can be differentiated into shorter and longer ones by the presence/absence of an accentual fall in the sentence-final verb phrase.

The mean F0s of sentence-final vowels seem to be more complicated. The data from Experiment 5 show that the mean F0 of a sentence-final vowel is a little over 200 Hz when neither the preceding noun phrase nor the following verb phrase has an accentual fall (e.g. *kaoga hareru, kaoga hareta*). When only the preceding noun has an accentual F0 fall, it stays around 175 Hz (e.g. *ka'taga hareru, kyu'urio tukeru*). When the final verb phrases has an accentual fall, the mean F0 gets even lower, ranging from 141.2 Hz to 156.3 Hz, whether or not the preceding noun phrase has an accentual fall. The absolute values of F0 may vary greatly among speakers (e.g. a male or female). However, it is assumed that relative pitch height is likely to show a similar pattern.

Figures. 5-1 to 5-4 present samples of the pitch contours and waveforms of test sentences of *hareta*-set with accentual contrasts in NPs and/or VPs. They present tokens of the four test sentences all uttered by Speaker B, and the displayed duration is the same among the four figures, thus direct comparison of duration on the figures is possible. The F0 domain surrounded by two lines is the period of each sentence-final vowel /a/.

Fig. 5-1. Waveforms and F0 contour of *kao ga hareta*

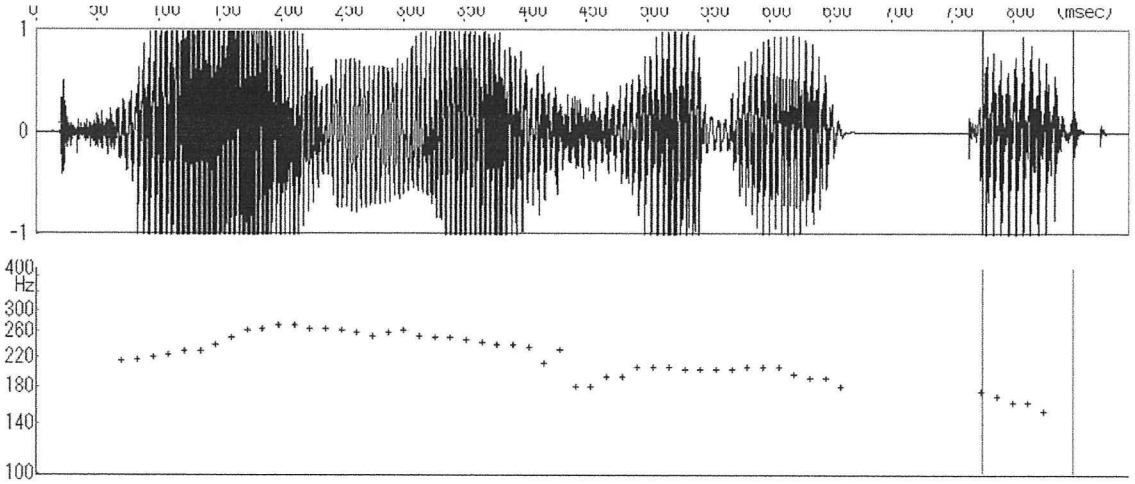


Fig. 5-2. Waveforms and F0 contour of *ka'ta ga hareta*

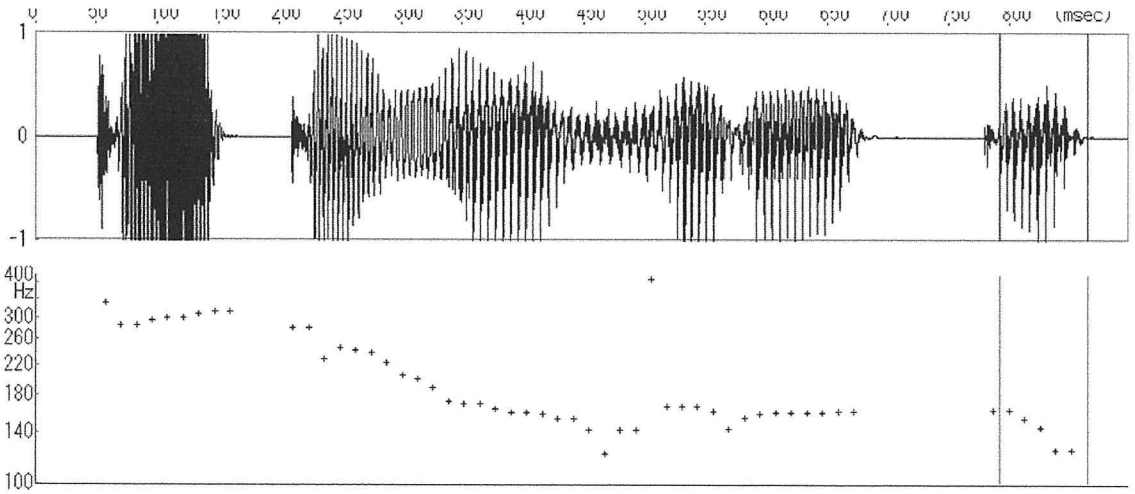


Fig. 5-3. Waveforms and F0 contour of *kao ga ha'reta*

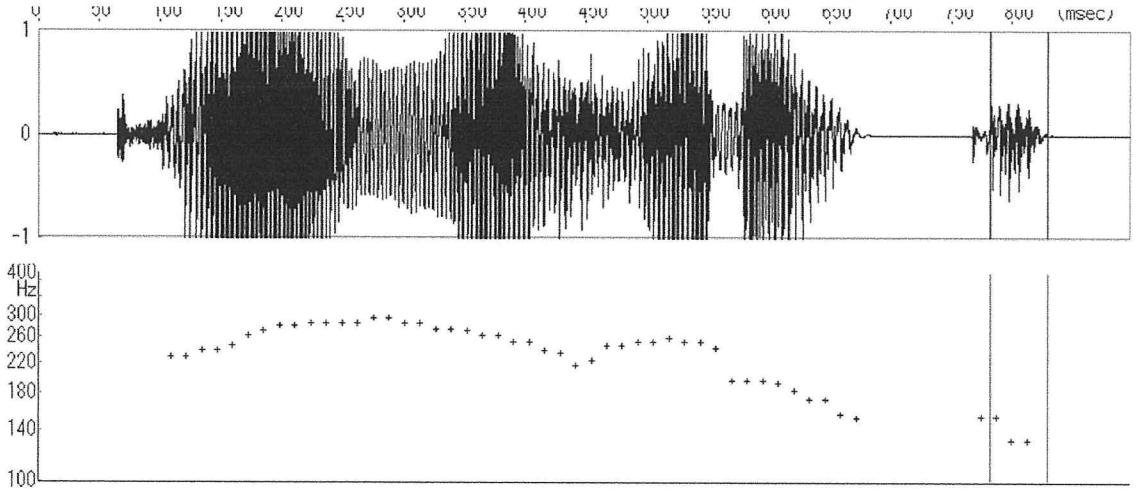
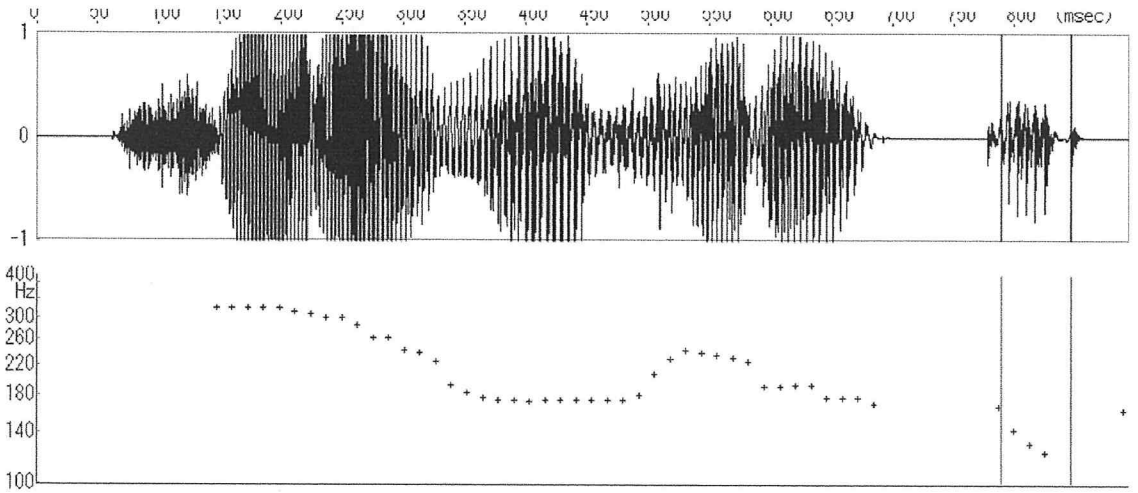


Fig. 5-4. Waveforms and F0 contour of *so'ra ga ha'reta*



Figures. 5-5 to 5-8 schematically represent how the pitch contour and duration of final vowels are realized in the four accent types *a, a', b, b'* (See Table 4-1). When a test sentence has no accentual fall as

in Fig. 5-5 (Fig. 5-1), the final vowel is relatively high-pitched and long in duration. When only the preceding noun has an accentual fall as shown in Fig. 5-6 (Fig. 5-2), the F0 in the sentence-final position gets lower than its counterpart in Fig. 5-5, but the duration is as long as that in Fig. 5-5. In Fig. 5-7 (Fig. 5-3) where only the final verb phrase has an accentual fall, the F0 gets even lower and the duration becomes shorter than those in Figs. 5-5 and 5-6. Finally, when both NP and VP have accentual falls as in Fig. 5-8 (Fig. 5-4), almost the same F0 and duration of the sentence-final vowel as those in Fig. 5-7 are realized.

Fig. 5-5. *a.* 0 + 0

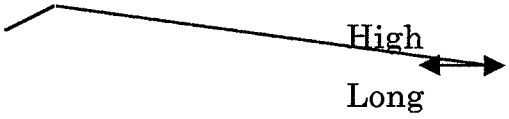


Fig. 5-7. *b.* 0 + 1

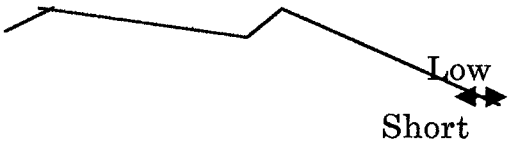


Fig. 5-6. *a'.* 1 + 0

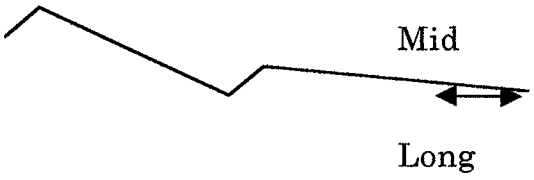
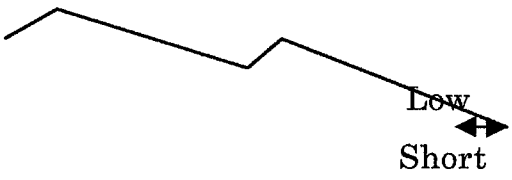


Fig. 5-8. *b'.* 1 + 1



From all the above findings, a proposal is made to distinguish the duration of sentence-final vowels into two relative levels (long vs. short), and their mean F0 into three relative levels (high, mid, and low). Based on this proposal, Table 5-1 presents the relative duration and pitch of sentence-final vowels that are predicted by the presence/absence of an accentual fall in the preceding and final phrases in a short declarative sentence. In order to examine

Table 5-1. Duration and pitch control of the final vowel in a short declarative sentence

Accent types	Accentual fall in the preceding noun phrase	Accentual fall in the final verb phrase	Duration of sentence-final vowel	Pitch of the sentence-final vowel
a	absent	absent	Long	High
a'	present	absent	Long	Mid
b	absent	present	Short	Low
b'	present	present	Short	Low

whether the differentiation in Table 5-1 is supported by statistically significant differences, ANOVAs were conducted on the durations of all the sentence-final vowels in Table 5-1 (i.e., those in Table 4-1), with accent type *a, a', b, b'* as a fixed factor. There was a highly significant effect of accent type [$F(3, 355)=52.61$, $p<0.0001$]. Multiple Scheffé comparisons indicated that there was no significant difference in vowel duration between accent types *a* and *a'*, and between *b* and *b'*, while a highly significant difference was detected between accent types *a* and *a'*, vs. b and *b'*($p<0.0001$). Thus, a clear distinction could be made between the vowel duration of *a* and *a'*, and that of *b* and *b'*. They are referred to as "Long" and "Short" in the table, respectively.

Another ANOVA was run on the mean F0 of all the sentence-final vowels in Table 5-1 (Table 4-1) with accent type as a fixed factor, showing a highly significant main effect [$F(3, 346)=94.43$, $p<0.0001$]. Post hoc tests (Scheffé) showed no significant difference between accent types *b* and *b'*. The other pairs of accent types indicated highly significant differences in mean F0 ($p<0.0001$). The results provides a support for distinguishing the F0s into three groups: those of *a*, those of *a'*, and those of *b* and *b'*. These are referred to as "High", "Mid", and "Low", in Table 5-1, respectively.

The above statistical analyses indicated that the two levels of

length and three levels of pitch proposed here were differentiated from one another by statistically highly significant differences ($p<0.0001$). Those that are treated as a single level of length or pitch, by contrast, showed no significant difference. Thus, the results provide a statistical support for the validity of this control model.

It should be added here that the sentences used to produce this model were all declarative sentences composed of a NP and a VP with no intervening pause. In addition, the tokens that consist of the data were obtained in the phonetic experiments where speakers were assumed to read each sentence with no focus or emphasis on a particular constituent of the sentence[1].

5.3 Prosodic control of prepausal vowels in declarative sentences with pauses

5.3.1 *Effect of test words in long declarative sentences*

As has been pointed out in 4.5.1, prepausal vowels of all the test sets in Experiment 6 shows no great difference in duration among the test sets under each condition of position and accent (cf. Figure 4-8).

ANOVAs were performed on the duration of prepausal vowels with the test sets as a fixed factor for each condition of position and accent (NonSF-NoA, NonSF-A, SF-NoA, SF-A). They showed no significant effect of the test sets. That is, there is no significant durational difference in the prepausal vowels of all the test sets for NonSF-NoA, NonSF-A, SF-NoA, and SF-A, respectively.

The results have also made it clear that the presence/absence of an accentual fall in the preceding phrase (non-prepausal phrase) in the test sentences for each set results in no significant difference. In *sake*-set and *sake*2-set, for instance, the preceding phrase of the

former set has no accentual fall, whereas the latter has an accentual fall. However, there is no significant difference in the duration of the corresponding prepausal vowels for the two sets; actually, the differences is less than 4ms (Table 4-3).

This finding suggests that the prepausal vowel duration of long declarative sentence with intervening pauses can be differentiated only by accent and position independent of the variation in segmental content of the final syllables /ga/, /te/, and /to/.

5.3.2 *Duration control of prepausal vowels in long sentences*

It turns out that, in the realization of prepausal vowel duration, the presence or absence of an accentual fall in the preceding phrase does not result in any significant difference. What plays a decisive role in the duration of prepausal vowels is whether or not prepausal noun phrases have an accentual fall. Thus, the relative length of prepausal vowels in sentential utterances with intervening pauses can be predicted by the presence or absence of an accentual fall in the prepausal noun phrase and by the sentence-final or non-sentence-final position.

In addition, Table 4-3 tells us that in each set the order of the vowels duration from the longest to the shortest is NonSF-NoA > NonSF-A > SF-NoA > SF-A except for *hasi*-set. In *hasi*-set, the vowel duration of NonSF-A and that of SF-NoA show no significant difference (4.5.1).

From all the above findings, the first proposal is made that the prepausal vowel duration in sentence-final position is differentiated into two relative levels, "Long" and "Short", which are the same levels proposed in the previous section (5.2) on the basis of the data from Experiment 5. Second, an additional proposal is made to differentiate the prepausal vowel length in non-sentence-final

position into two levels, "Extra-long" and "Long".

In order to examine the validity of the differentiation into each of the two durational levels, ANOVAs were conducted on the duration of all the prepausal vowels in non-sentence-final position and sentence-final position, respectively, with accent type *a, a', b, b'* presented in Table 4-3 as a fixed factor. There was a highly significant effect of accent type both in non-sentence-final position [$F(3, 523)=14.38$, $p<0.0001$] and sentence-final position [$F(3,344)=62.19$, $p<0.0001$]. In both positions Scheffé tests indicated that there was no significant difference in vowel duration between accent types *a* and *a'*, and between *b* and *b'*, while a highly significant difference was detected between accent types *a* and *a'*, vs. b and *b'* ($p<0.001$).

These distinctions of prepausal vowel duration lead to the following table which predicts the duration of a prepausal vowel in a declarative sentence with a pause before a major phrasal boundary.

Table 5-2. Duration control of prepausal vowels in long declarative sentences with intervening pauses

Accent type	Accentual fall in the preceding phrase	Accentual fall in the prepausal phrase	Duration of non-sentence-final vowel	Duration of sentence-final vowel
a	absent	absent	Extra-long	Long
a'	present	absent	Extra-long	Long
b	absent	present	Long	Short
b'	present	present	Long	Short

The questions may arise from the above table as to whether the levels represented as "Long" in non-sentence-final position and in sentence-final position can be considered almost similar duration or whether they show a statistic difference. Statistically, the two groups showed a significant difference [$F(1,454)=21.65$]. However, as stated above, when analyzed by set, there was no significant

difference between the vowel duration of NonSF-A and that of SF-NoA in *hasi*-set. In addition, the difference in vowel duration is on average 13.2ms, which is shorter than any other difference in the pairs of the durations in NonSF-NoA, NonSF-A, SF-NoA, and SF-A (cf. Fig. 4-8). Moreover, there is a small effect (1-7ms) of the preceding consonants /g/ and /d/ on the duration of prepausal vowels. Thus, there is a possibility that the difference in the two lengths is too small to be perceived even within the utterances of a single speaker. Therefore, it is proposed that the two groups are referred to as the same "Long" level in this control model.

5.3.3 *Pitch control of prepausal vowels in long sentences*

Let us now discuss the F0 realization in prepausal vowels in sentence-final and non-sentence-final positions. As has been seen in Fig. 4-13, the mean F0 of prepausal vowels does not show a large variation among the test sets except those in NonSF-NoA. ANOVAs were performed with the test sets as a fixed factor on the mean F0s in NonSF-NoA, NonSF-A, SF-NoA, and SF-A, respectively. In NonSF-A and SF-A, there was no significant effect of the test sets. In SF-NonA, there was a significant difference of 17.3 Hz between the mean F0 of *atusa*-set and that of *siro*-set at the $p<0.01$ level. No other pair of the test sets showed a significant difference. In NonSF-NoA, all the sets with accent type *a* had significantly higher F0s than those with accent type *a'*. That is, in NonSF-NoA, PBUs with no accentual fall in the preceding phrase tend to have higher mean F0 than those with an accentual fall in the preceding phrase.

However, the prepausal vowel of *kamuto*-set in NonSF-NoA showed a significant difference with its counterpart of *atusa*-set and *harete*-set, though it showed no difference with that of *sake*-set. Actually, as can be seen in Fig. 4-13, the F0 of *kamuto*-set in NonSF-

NoA is located around the middle between the higher groups (*atusa*-, *harete*-, and *sake*-sets) and lower groups (*hasi*-, *sake*2, and *siro*-sets). A little lower F0 observed in *kamuto*-set than those in the other sets with no accentual fall may be attributable to the segmental structure of the prepausal noun phrase. However, in differentiating the F0 values into several levels, the *kamuto*-set is included in the higher F0 group composed of those with no accentual fall in the preceding phrase on the ground that the F0 showed no significant difference with that of *sake*-set of the higher F0 group.

Another ANOVA was run to compare the mean F0s of the lower group in NonSF-NoA (*hasi*-, *sake*2, and *siro*-sets) with the mean F0s of all the test sets in SF-NoA. It showed no significant difference in mean F0 between the two groups. The result suggests that the relative level of F0s in SF-NoA can be differentiated into the same level of F0s of the lower group in NonSF-NoA.

It follows from all the above findings that the level of the mean F0s observed in the prepausal vowels in Experiment 6 can be differentiated by four levels and referred to as "Extra-High", "High", "Mid", and "Low" in this thesis.

In order to examine whether this differentiation is supported by significant differences between each pair of the levels, an ANOVA was run on mean F0 of sentence-final vowels with the four kinds of accent types *a, a', b, b'* as a fixed factor, showing highly significant main effects both in non-sentence-final position [$F(3,516)=252.91$, $p<0.0001$] and in sentence-final position [$F(3,317)=105.71$, $p<0.0001$]. Post hoc tests (Scheffé) showed no significant difference between accent types *b* and *b'* in non-sentence-final position ($p<0.0001$). The other pairs of accent types indicated highly significant differences in mean F0 in NonSF. At sentence-final position, there was no significant difference between *a* and *a'* and

between *b* and *b'*. Any other pair of accent types showed a significant effect on mean F0 ($p<0.0001$).

5.3.4 *Prosodic control of prepausal vowels in declarative sentences with intervening pauses*

Based on all the above findings and discussions, a prosodic control model is proposed in Table 5-3 by which relative duration and mean F0 of the prepausal vowel are predicted by the presence or absence of an accentual fall in the preceding phrase and the prepausal phrase in a PBU of a sentence. The utterances whose prepausal vowels are predicted by this table are the utterances of declarative sentences presumably without any intentional focus or emphasis by speakers.

Table 5-3. Prosodic control of prepausal vowels in declarative sentences with intervening pauses. A.T. represents accent type.

A. T.	Accentual fall in the preceding phrase	Accentual fall in the prepausal phrase	Duration and pitch of non-sentence-final vowel		Duration and pitch of sentence-final vowel	
a	absent	absent	Extra-Long	Extra-High	Long	High
a'	present	absent	Extra-Long	High	Long	High
b	absent	present	Long	Mid	Short	Low
b'	present	present	Long	Mid	Short	Low

It should also be recalled that what this control model predicts is a relative level of duration and pitch of prepausal vowels under a certain accent or position condition. A prepausal vowel with accent type *a* or *a'* in non-sentence-final position, for instance, is most likely to be relatively longer than that with accent type *b* or *b'* in the same non-sentence-final position. A prepausal vowel with accent type *a* or *a'* can also be predicted to be longer in non-sentence-final position than its counterpart with accent type *a* or *a'* in sentence-final

position.

The model is not designed to provide an absolute value of duration or F0 of any utterance produced by any random speaker. Needless to say, there is a wide variation in the height of sound produced by men and women. Moreover, there is a great interspeaker variation in the realization of duration and F0 of any segment that constitutes their speech. It is possible, for example, that the absolute F0 (Hz) of a prepausal vowel produced by a low-voiced man in non-utterance-final position may be lower than that same prepausal vowel produced by a high-voiced woman in sentence-final position. Thus, the relative levels presented in Table 5-3 is valid within the utterances produced by a single speaker, or the mean F0s or durations of utterances produced by a certain group of people.

Finally, it should be noted that this prosodic control model is the first model where the relation between duration and pitch is taken into consideration. No previous models have predicted both the duration and F0 of prepausal vowels in relation to the accentual fall. In addition, previous studies have reported relatively large rms error values in the duration in prepausal position by their duration control models. The consistent effect of accent on prepausal vowel duration clarified in this study may contribute to enhance the accuracy and naturalness of their synthesized speech.

Another contribution that may be made by this model is that it provides a sufficient and clear-cut explanation of sentence-final intonation of Japanese declarative sentences for learners of Japanese as their second language. As has been pointed out in 1.2.2, there exists a large discrepancy in the description of the sentence-final pitch realization in Japanese pedagogy, which has provided no satisfactory description of sentence-final tone in relation to the accent pattern of the sentence constituents.

5.4 Summary

Based on all the data obtained from Experiments 5 and 6, a prosodic control model is proposed to predict the relative duration and pitch of prepausal vowels by the presence or absence of an accentual fall in the preceding and prepausal phrases in pause-bounded units of sentence utterances.

---

[1] In the serious of experiments in this study, all the speakers confirmed that they read test items without any emphasis on a certain constituent of test items. Thus, it is assumed there was no intentional focus or emphasis placed on any particular element of their utterances. However, this does not totally deny a possibility that speakers unconsciously put a focus on some part of their utterances.

CHAPTER 6

Markedness of prepausal vowels in Japanese rhythmic structure

6.1 Introduction

The foregoing chapters have clarified consistent and significant effect of accent on duration of prepausal vowels. The final questions we have to ask are why the accent-related effect is specific to prepausal position and what this position-specific variability in duration means in Japanese.

First, we examine other phenomena that show a consistent temporal variation in Japanese such as lengthening of monomoraic words without particles or phrase-final lengthening, and demonstrate to what extent they exhibit durational variability under each specific condition.

Second, the extent of durational variability of these phenomena is compared with that of prepausal vowels in the present study. Showing no other phenomena exhibit such a wide durational variation of segments, markedness of prepausal position is pointed out. It is claimed that prepausal position is out of mora-timing constraint and that phonological contrast between long and short vowels weakens in this final position, which makes it possible for speakers to lengthen or shorten prepausal vowels as they like.

6.2 Lengthening of monomoraic nouns

6.2.1 *Introduction*

So far I have discussed a considerable increase in duration specifically observed in prepausal vowels in cases where the prepausal content words have no accentual fall compared with those that do. The results from Experiments 1 to 6 also indicate that no substantial lengthening related to accentuation occurs except for prepausal position, presumably due to the mora-timing constraint to keep the length of each mora roughly equal (1.1.2). In addition, as has been discussed in 1.1.1, it is assumed that Japanese moras are not allowed to be lengthened to be two moras because of a phonological contrast between short and long vowels. That is, if a short vowel doubles in length, it may lead to produce a different morpheme or word.

However, anecdotal observation suggests that Japanese one-mora words functioning as subjects or objects are considerably drawn out and frequently followed by pauses, when not suffixed with monomoraic case markers such as *ga* or *o*. For example, the sentences *me'ga akaiyo* "Your eyes are red." or *te'o aratta* "Did you wash your hands?", can be rendered as *me' akaiyo* or *te' aratta* in colloquial Japanese. In the latter renditions, the monomoraic nouns *me'* "eye" and *te'* "hand" are not suffixed with a subject case marker *ga* and an object case marker *o*, respectively. In these contexts, the monomoraic nouns seem to undergo substantial lengthening.

In the Osaka dialect of Japanese, it has been pointed out that such monomoraic lengthening is substantial whether or not monomoraic words are followed by particles. Sugitô (1975) provides phonetic evidence for the lengthening of monomoraic nouns uttered by five native speakers of the Osaka dialect by measuring

monomoraic nouns with or without particles in unspecified sentences. Her experimental results show that monomoraic nouns without particles are about one and a half times longer than those with particles when averaged across speakers. Moreover, in some of the utterances produced by the speakers even monomoraic nouns suffixed with particles undergo a similar amount of lengthening to those without particles.

Phonological analyses of this lengthening phenomenon is provided by Itô (1990) who points out that in standard Japanese monomoraic words are lengthened "in (i) 'schedule' compounding" such as *"getsu-kaa"* ("Monday and Tuesday"), "(ii) telephone number recitation" such as *"goo, kyuu"* ("five, nine"), and "(iii) verbal root reduplication" such as *"shii-shii"* ("while doing") (pp.225-226). Thus, she claims that monomoraic words are lengthened in their derived forms.

The question arises whether monomoraic lengthening is not restricted to a particular dialect of Japanese or to derived words, but is consistently observed in monomoraic nouns without particles. If the single mora is consistently lengthened, this would run counter to the mora-timing hypothesis of Japanese. The purpose of the present experiment is, therefore, to investigate the extent to which monomoraic lengthening can be found acoustically.

First, I set out to clarify the questions of whether monomoraic nouns are lengthened when not suffixed with particles in standard (Tokyo) Japanese and whether the lengthening is specific to nouns with one mora, that is, whether bimoraic nouns undergo the equivalent amount of lengthening even without particles. Second, the current experiment tests whether the lengthening occurs regardless of the accent pattern of neighboring words as well as in the absence of focus, both of which are factors that have been found to

influence segmental duration in a host of other languages (Sluijter & van Heuven, 1995; Turk & White, 1999; Cambier-Langeveld & Turk, 1999).

Third, the degree of the monomoraic lengthening is discussed in relation to the bimoraic constraint on the minimal size of prosodic words, phonological constraint of vowel length, and phrase-final lengthening, in Japanese.

6.2.2 *Experiment 7*

This experiment was designed to examine the assumption that monomoraic nouns are consistently lengthened when not suffixed with particles. First, a pair of sentences consisting of "monomoraic noun + *ga* + verb (phrase)" and "monomoraic noun + verb (phrase)" was prepared for each test sentence. The two sentences do not differ greatly in meaning but do differ in the presence/absence of the particle /ga/. Thus, in each pair of test sentences, the duration of the monomoraic nouns by themselves can be compared with that suffixed with /ga/ in order to acoustically examine the lengthening of monomoraic nouns.

Second, possible factors (vowel type, accent pattern of the following verb, grammatical function, and the location of focus) that may influence the lengthening of monomoraic nouns were examined to ascertain the extent to which the lengthening occurs regardless of these factors.

Materials

All the test sentences and phrases are listed in Appendix 6-1. The test sentences were 16 real Japanese sentences, all composed of only light syllables (CV and/or V syllables). First, as shown in (1) of Appendix 6-1, monomoraic nouns *ne'* "root" and *na'* "vegetable" were

chosen as basic target nouns because the boundary of the nasal /n/ with the adjacent vowels is distinct enough to be consistently identified and used as a measurement point.

In order to make a comparison between monomoraic nouns that are suffixed with /ga/ and those that are not, a pair of sentences with or without the case marker /ga/ was prepared. Thus, /ne/ and /nega/ were combined with verb phrase *nuketa* "was/is pulled out", whereas /na/ and /naga/ were combined with *nieta* "was/is boiled".

In order to examine the effect of the accent pattern of a verb phrase on monomoraic lengthening, *ne'* and *ne' ga* were also combined with the verb phrase *no'bita* "has grown" whose pitch accent differs from that of *nuketa*[1].

The test sentence with *ne'* as the object, *ne' o nuita* "I pulled out the root." and the same sequence minus a particle, *ne' nuita*, were prepared in order to investigate whether /ne/ is lengthened when not suffixed with the particle /o/; that is, whether the monomoraic lengthening occurs even when the one-mora noun functions as an object.

Second, negative sentences *ne'(ga) nuke na'katta* and *na'(ga) nie na'katta* in (2) of Appendix 6-1 were constructed by adding an auxiliary verb phrase *na'katta* (indicating negation) to the positive counterparts. In these negative sentences focus is usually considered to be placed on the predicate (Kôri, 1997), not on the subject. These test sentences were added in order to investigate whether monomoraic lengthening consistently occurs even if speakers place focus on constituents other than the monomoraic nouns.

Third, bimoraic nouns *ya'ne* and *i'ne,* whose second moras comprise the same segments as does the monomoraic noun *ne'* were set in the same frame sentences *(ga) nuketa* and *(ga) no'bita*,

respectively. These were prepared to see if the second moras of the bimoraic nouns undergo lengthening when not suffixed with particles.

In the test sentence *ne'(ga) nuketa*, the syllable that immediately follows /ne/ is the syllable /ga/ when suffixed with a particle, while it is /nu/ of *nuketa* when not suffixed with a particle. Thus, it is possible that the change of the neighboring syllable from /ga/ to /nu/ may affect the duration of the monomoraic noun /ne/. Ten repetitions of nonsense words /nega/ and /nenu/ written on a sheet in random order was prepared in order to examine whether the duration of the target syllable /ne/ is influenced by the segmental change of the following syllable from /ga/ to /nu/.

Informants and procedure

Informants were five middle-aged females, all native speakers of Tokyo (standard) Japanese. All informants were first asked if they uttered in their daily lives sentences in which monomoraic nouns were combined with verb phrases without case markers, such as *te' aratta* or *me' akaiyo*. They all confirmed that they did utter such sentences. Next, each informant was presented the test sentence with 20 other dummy sentences on a card one at a time in a random order and asked to read the sentence naturally at a tempo comfortable to herself as if she spoke it in her daily conversation. Every time the informant finished reading all the 16 sentences, the cards were randomized and presented to the informant, for a total of seven different randomizations. Third, a list of nonsense words *ne'ga* and *ne'nu* was read twice by each informant to produce a total of twenty tokens for *ne'ga* and *ne'nu*, respectively.

Productions were recorded and digitized at a sampling rate of 11 kHz, and waveforms and wideband spectrograms were made with a

Kay Multi-Speech Model 3700. As a rule, the tokens from the second to the seventh reading were measured. Only when any of the tokens was misread, was the token discarded and the first utterance was used instead. The total number of tokens analyzed was (16 sentences × 6 repetitions × 5 subjects=) 480.

Measurements were obtained by using manual cursors for durations of (1) the consonant /n/ and vowel /e, a/ of the target monomoraic nouns /ne/ and /na/, (2) the syllable /ga/, (3) the whole utterance, and (4) pauses, if any, immediately following the monomoraic nouns. The beginning of the nasal /n/ was taken at the start of the nasal murmur where the nasal formant at low frequency appears. The boundary between the nasal and the following vowel was defined at the point where the formants of the vowel appear. All the segmental durations were repeatedly measured to ascertain the accuracy and consistency of the measurements.

Second, pitch contours were extracted with all the utterances of the two negative sentences *ne'(ga) nuke na'katta* and *na'(ga) nie nákatta* to confirm the placement of focus on the predicates.

Third, the duration of /ne/ was measured in the twenty productions of /nega/ and /nenu/ by each speaker.

6.2.3 *Results and discussion*

Lengthening of monomoraic nouns

Table 6-1 shows the duration of target monomoraic nouns /ne/ and /na/ averaged across speakers for each test sentence. In the case of /ne/ suffixed with the case marker /o/, only the duration of the successive syllables /neo/ was measured since the boundary between /e/ and /o/ is not distinct enough to be objectively determined. The one-mora nouns without particles are further divided into those with pause insertion and those without.

Table 6-1. Mean duration (ms) and SD (in parentheses) of /ne/ or /na/, /ga/, and pause for each test sentence. N indicates number of tokens.

Test sentence	With particle		Without particle					
	/ne/ or /na/	/ga/	Without pause			With pause		
			N	/ne/or /na/	MLR	N	/ne/ or /na/	Pause
ne'(ga) nuketa	137.3 (19.2)	142.3 (12.8)	17	193.3 (26.2)	1.41:1	13	184.1 (23.1)	114.5 (31.3)
ne'(ga) no'bita	136.4 (20.5)	145.5 (11.3)	18	194.5 (36.2)	1.43:1	12	190.0 (15.6)	118.9 (28.0)
ne'(ga) nuke na'katta	139.2 (19.3)	130.8 (10.5)	20	196.4 (33.2)	1.41:1	10	182.6 (18.7)	121.3 (28.9)
ne'(o) nuita	287.7 (22.5)		19	207.1 (35.0)		11	186.2 (18.7)	118.2 (29.7)
na'(ga) nieta	140.7 (21.9)	153.4 (10.6)	16	204.6 (29.3)	1.45:1	14	200.1 (26.4)	126.2 (32.4)
na'(ga) nie na'katta	140.2 (24.3)	144.7 (10.3)	20	209.8 (40.3)	1.50:1	10	196.8 (30.1)	125.4 (37.9)
ya'ne	141.9 (12.3)	130.8 (12.9)	30	149.0 (9.5)	1.05:1	0		
i'ne	153.9 (13.5)	140.5 (10.9)	26	160.2 (16.7)	1.04:1	4	189.8 (8.9)	122.3 (11.0)

Compared with the length of the target one-mora nouns suffixed with /ga/, the sentence-initial monomoraic nouns /ne/ and /na/ were consistently lengthened by 56.0ms to 69.6ms when not suffixed with the particle. Furthermore, in 10 to 14 out of a total of 30 tokens of each test sentence, pauses were inserted immediately after the monomoraic nouns.

The ratio of the mean duration of /ne/ or /na/ without particles and pauses to that suffixed with /ga/ was calculated so as to represent the degree of monomoraic lengthening. This ratio is named "Monomoraic Lengthening Ratio" in the present study and

abbreviated as MLR. To be more precise, Monomoraic Lengthening Ratio (MLR) is defined as a ratio of the mean duration of monomoraic nouns followed neither by particles nor by pauses to that followed by the particle /ga/ for each pair of test sentences.

As seen in Table 6-1, MLR ranges from 1.41:1 to 1.50:1, which means that in utterances without particles and pauses the monomoraic nouns are lengthened by about 40% to 50%.

Repeated measures analyses of variance (RM-ANOVAs) were conducted in order to examine the effects of the possible factors on the monomoraic lengthening. Only the durations of monomoraic nouns followed neither by particles nor by pauses and those of their counterparts followed by /ga/ were included in the statistic analyses.

First, a two-way RM-ANOVA[2] was run on the durations of /ne/ in *ne'(ga) nuketa* and /na/ in *na'(ga) nieta* with the vowel quality (/e/ vs. /a/) and the presence/absence of a particle as two fixed factors. The effect of presence/absence of a particle was found to be highly significant [$F(1,31)=230.4$; $p<0.0001$]. There was no significant effect of vowel quality nor any interaction between the factors.

Another two-way RM-ANOVA was performed on the duration of /ne/ in *ne'(ga) nuketa* and *ne'(ga) no'bita* with the accent pattern of the following verb (*nieta* vs. *no'bita*) and the presence/absence of a particle /ga/ as two fixed factors, showing a main effect of the presence/absence of a particle [$F(1, 33)=196.2$; $p<0.0001$]. There was no significant effect of accent pattern nor any significant interaction between the factors.

The next question to be addressed is whether monomoraic nouns as subjects are lengthened even if focus is placed on the predicates. All the utterances of the negative sentences *ne'(ga) nuke na'katta* and *na'(ga) nie na'katta* show the pitch of the syllable /na/ in the auxiliary verb phrase *na'katta* as high as or even higher than that of

the subject /ne/ or /na/. This suggests that the speakers placed focus on the predicate, not on the monomoraic subjects (Kôri, 1997). That is, the monomoraic nouns serving as subjects were not in focus in the negative sentences.

The absence of focus, however, does not affect monomoraic lengthening. As is shown in Table 6-1, the subjects /ne/ and /na/ followed neither by particles nor by pauses in the negative sentences were lengthened to an extent similar to their counterparts in the affirmative sentences. A two-way RM-ANOVA was performed for each pair of affirmative and negative sentences with the placement of focus and the presence/absence of a particle as two fixed factors. The effect of the presence/absence of a particle was observed to be highly significant in both pairs [*ne'(ga) nuketa* & *ne'(ga) nuke na'katta*: $F(1,35)=151.2$, $p<0.0001$; *na'(ga) nieta* & *na'(ga) nie na'katta*: $F(1,34)=152.4$, $p<0.0001$]. No significant effect of the focus nor interaction between the factors was found.

The experimental results indicate that the segmental change of the following syllable from /ga/ to /nu/ has no significant effect on the duration of the target mora /ne/. The mean duration of /ne/ in isolated nonsense utterances of *ne'ga* and that of *ne'nu* averaged across speakers were 163.0ms (SD 26.5) and 163.2ms (SD 22.8), respectively, indicating no significant difference (RM-ANOVA).

Monomoraic lengthening is supposed to take place even when the monomoraic noun is an object of the verb as in *ne'(o) nuita*[3]. In the utterances of *ne' o nuita*, no data for the duration of /ne/ was obtained since the boundary between /ne/ and /o/ is not distinct enough to be measured. Thus, only the duration of /ne/ in *ne' nuita* without a pause is compared with that in *ne' nuketa* without a pause. As shown in Table 6-1, when followed neither by particles nor by pauses, the mean duration of /ne/ functioning as an object in *ne' nuita*

is a little longer than its duration when it functions as a subject in *ne' nuketa*. Also, a pause was inserted into eleven out of thirty utterances of *ne' nuita*. An RM-ANOVA showed no significant difference in duration between the /ne/ in *ne' nuita* and that in *ne' nuketa* with no following pause, suggesting the occurrence of monomoraic lengthening in *ne' nuita* equivalent to that in *ne' nuketa*.

In sum, the above results have made it clear that the lengthening by 40% to 50% of monomoraic nouns without particles was consistently observed despite the difference in the component vowel (/e/ vs. /a/) of the mora and in the accent pattern of the following verb (*nuketa* vs. *no'bita*). The lengthening also occurred in spite of the absence of focus and regardless of the grammatical function of the one-mora nouns.

Domain of lengthening

Table 6-2. Mean duration (ms) and SD of the consonant /n/ of /ne/ or /na/ when followed by particle and followed neither by particle nor by pause for each test sentence.

Test sentence	With particle		Without particle and pause			Repeated measures ANOVA
	/n/ in /ne/ or /na/ (ms)	SD	/n/ in /ne/or /na/ (ms)	SD	Ratio of lengthening	
ne'(ga) nuketa	42.6	13.1	60.9	15.3	1.43:1	F(1,32)=6.19, p=0.018
ne'(ga) no'bita	40.1	14.5	64.0	16.3	1.60:1	F(1,34)=12.58, p=0.0012
ne'(ga) nuke na'katta	43.0	14.3	60.4	10.7	1.40:1	F(1,38)=11.08, p=0.0019
na'(ga) nieta	40.2	12.3	60.1	10.7	1.50 :1	F(1,30)=16.43, p=0.0003
na'(ga) nie nákatta	42.1	17.6	61.1	15.8	1.45 :1	F(1,38)=7.00, p=0.012

Table 6-2 shows the duration of nasal /n/ of the monomoraic nouns /ne/ and /na/ averaged across speakers for each test sentence. "Ratio of lengthening" in the table represents the ratio of the mean duration of /n/ in the target monomoraic nouns followed neither by particles nor by pauses to that suffixed with the particle /ga/. The ratio indicates that the consonant is lengthened by 40% to 60% when the monomoraic noun is not suffixed with /ga/.

As presented in Table 6-2, a one-way RM-ANOVA was conducted for each test sentence, showing a significant main effect of the presence/absence of a particle ($p<0.05$). Thus, the results make it clear that monomoraic lengthening takes place not only in a vowel but also in the preceding consonant, that is, in a unit of a mora.

Length of monomoraic noun + pause

Figure 6-1 illustrates the mean duration of /ne/+/ga/ in *ne' ga nuketa*, and /ne/ without pause and /ne/+pause in *ne' nuketa*. In Figures 6-1 and 6-2, the number in parentheses at the end of each test sentence indicates the number of occurrences for each utterance. We can see from Fig.6-1 that the total length of /nega/ is about 280ms, while /ne/ plus pause amounts to nearly 300ms.

A similar tendency is observed in all the other test sentences with or without particles in Table 6-1. It seems as if the speakers insert pauses so as to compensate for the absence of the particles and to take up about the same amount of time as two moras.

Bimoraic nouns without particles

Table 6-1 shows the mean duration of /ne/ in *ya'ne(ga) nuketa* and *i'ne(ga) no'bita*. A two-way RM-ANOVA was performed on the duration of /ne/ in *ya'ne(ga) nuketa* and *i'ne(ga) no'bita,* with the

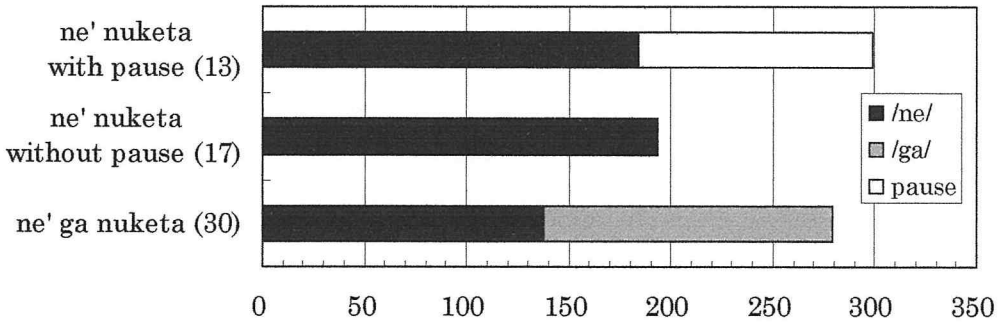


Fig. 6-1. Mean durations (ms) of /ne/+/ga/, /ne/, /ne/+pause in ne' (ga) nuketa

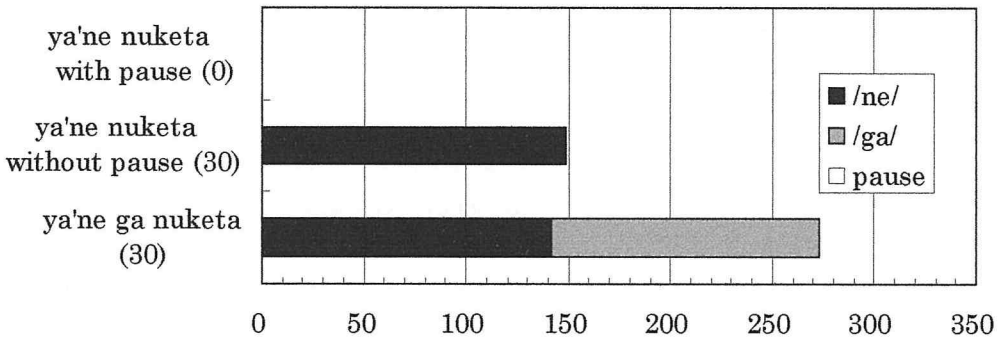


Fig. 6-2. Mean durations (ms) of /ne/+/ga/, /ne/, /ne/+pause in ya'ne (ga) nuketa

bimoraic nouns (*ya'ne* vs. *i'ne*) and the presence/absence of a particle as two fixed factors. Both effects were found to be significant [the bimoraic nouns: $F(1,54)=13.16$, $p=0.0006$; the presence/absence of a particle: $F(1,54)=19.79$, $p<0.0001$].

The significant difference in the duration of /ne/ of *ya'ne* and *i'ne* suggests temporal compensation (cf. 1.1.2) taking place at the word level. It is assumed that /ne/ in *i'ne* is lengthened so as to compensate the shorter preceding mora /i/ composed of only a vowel.

The significant lengthening observed in the second moras of the bimoraic nouns without /ga/ suggests that the second moras are lengthened in the final position when not suffixed with the particle.

Although /ne/ in *ya'ne nuketa* and *i'ne no'bita* is significantly

longer than that in *ya'ne ga nuketa* and *i'ne ga no'bita*, the degree of lengthening is largely different from that of the monomoraic noun /ne/. As can be seen in Table 6-1, the final moras of the bimoraic nouns *ya'ne* and *i'ne* are lengthened by only 6.3 to 7.1ms (4 to 5%) when not suffixed with particles. In addition, pause insertion occurred only a total of four times in the utterances without particles.

Figure 6-2 illustrates the mean durations of /ne/+/ga/, /ne/ without pause, and /ne/+pause in *ya'ne(ga) nuketa*. Compared with *ne' nuketa* without pause in Figure 6-1, it is obvious that the second mora of *ya'ne* is much less lengthened when not suffixed with a particle. In addition, no pause was inserted in any of the tokens.

To sum up, it can be said that the second mora of a bimoraic noun does not show lengthening equivalent to that of a monomoraic noun, when not suffixed with a particle. In other words, the considerable lengthening caused by the absence of a particle is restricted to monomoraic nouns.

6.2.4 *General discussion*

The experimental findings can be summarized as follows:

(1) Sentence-initial one-mora nouns, whether functioning as subjects or objects, are significantly lengthened by about 40 to 50 percent, or accompanied by pauses, when not suffixed with particles. The lengthening is found in both the consonantal and the vocalic elements of the mora, irrespective of variation in vowel quality of monomoraic nouns and accent pattern of the following verb, and despite the absence of focus. In utterances where pauses are inserted, the total length of monomoraic nouns plus pauses is almost the same or a little over the total of one-mora nouns plus monomoraic particles.

(2) The lengthening is specific to one-mora nouns since the second

moras of bimoraic nouns do not undergo lengthening equivalent to that of monomoraic nouns, when not suffixed with particles. Moreover, pause is only rarely inserted after bimoraic nouns.

Monomoraic lengthening & bimoraic minimality

The first question to be addressed herein should be why monomoraic nouns are lengthened by about 40% to 50% or followed by pauses when not suffixed with particles. This phenomenon may be reflection of what Itô (1990) refers to as the "bimoraic minimality" constraint, which is formalized as follows:

> A minimal phonological word of Japanese (and other moraic languages) consists of at least two moras.

In accord with this constraint on a phonological word, it might be hypothesized that the monomoraic noun in question undergoes lengthening when it is isolated in the phrase-final position; that is, when it is not suffixed with a particle and forced to serve as a complete phonological word on its own, the monomoraic noun is lengthened or followed by a pause in an attempt to form a minimal phonological word of two moras.

The present results also suggest that there is a syntactic constraint that a major phrasal boundary between NP and VP cannot be crossed when forming a phonological word of two moras. When there is no mora to combine with it under the syntactic constraint, the isolated monomoraic noun is lengthened and/or followed by a pause in an attempt to form a minimal phonological word on its own as we have observed in /ne/ of *ne' nuketa* or /na/ of *na' nieta* in the current experiment.

Phonological constraint

If the bimoraic minimality constraint is indeed a valid factor in account for the patterns of lengthening observed in the present study, it is also relevant to ask why the speakers did not lengthen all the monomoraic nouns without particles to the length of two moras (recall that, at most, lengthening was in the order of 50%). An important aspect of the phonological system of Japanese appears to be that of providing a constraint on the degree of the lengthening. In Japanese the length of vowels, long vs. short, is considered to be a distinctive feature. It seems possible that monomoraic lengthening of nouns is constrained in such a way as to avoid encroaching upon the length contrast. Previous studies (Beckman, 1982; Hoequist, 1983a) have reported that the ratio of a heavy (two-mora) and light (one-mora) Japanese syllables ranges from 1.66:1 to 1.80:1. The monomoraic lengthening observed in the present study results in smaller mean lengthening ratios, thus differentiating in durational terms the lengthened tokens from those which are phonologically long.

Interestingly, when speakers made use of the alternative strategy of pause insertion, as they actually did in many utterances of the test sentences, it was found that the sequence of monomoraic noun plus pause was able to occupy an amount of time similar to that of two moras, presumably because such a sequence does not have any implications for the length contrast.

Bimoraic nouns without particles

A two-mora noun such as *ya'ne* or *i'ne* as a subject is not lengthened nor a pause is inserted even without a particle. This would be consistent with an analysis whereby these bimoraic nouns form independent bimoraic phonological words on their own, when not

suffixed with particles. Similarly, a monomoraic noun followed by /ga/ could be considered to combine with the particle to form a single phonological word of two moras. In these cases, there is neither substantial lengthening nor pause insertion.

Effect of phrase-final lengthening

Campbell (1992a,b) reports "phrase-related lengthening" in the reading of Japanese sentences, and "the lengthening of particles in Japanese" partly because "they occur, by definition, phrase-finally" (1992b: 414). Thus, it is possible that the particle /ga/ in all the utterances such as *ne' ga nuketa* and *ne' ga no'bita* in the current experiment may undergo lengthening to some extent in the phrase final position.

In addition, as seen in Table 6-1, the second moras of the bimoraic nouns are lengthened by 4 to 5% when not suffixed with particles. Without particles, /ne/ in these nouns or phrases are in the final position of NPs. Thus, they may undergo phrase-final lengthening of 5% or less before a major phrasal boundary between a NP and a VP.

Monomoraic nouns without particles are also in the phrase final position when not suffixed with particles. Thus, there is a possibility that the phrase-final lengthening to a similar extent may partly contribute to the monomoraic lengthening observed in the present study.

6.2.5 *Summany*

Contrary to stress-timed or syllable-timed languages, Japanese moras do not usually undergo a considerable lengthening due to mora-timing constraint. However, Experiment 7 reveals the consistent lengthening of about 40 to 50% in monomoraic nouns

without particles whose duration is too short to constitute minimal phonological words of two moras.

The results also show that Japanese moras are resistant to lengthening by over 50% even under the bimoraic minimality constraint, suggesting the constraint of phonological contrast between long and short vowels in Japanese.

6.3 Lengthening of moras under accent or focus

It has been repeatedly observed in all the experiments of this investigation that accent has no consistent effect on the duration of segments except for prepausal vowels. This finding is consistent with the reports by Kaiki et al. (1992) and Sakaguchi et al. (1999) who observe no correlation between accent and duration through corpus analyses. In addition, this phenomenon has been considered supporting evidence for the weaker version hypothesis of mora timing in Japanese. That is, it has been claimed that there is a strong constraint of mora timing on accent-related durational variability in Japanese moras, by which moras are kept roughly equal in duration.

Even if a significant durational difference is detected due to differing accent in corresponding segments of the test items with the same segmental structure, the degree of difference is within about 10ms and the phenomenon is restricted to that specific segment in a specific test word/phrase/sentence. This suggests that the phenomenon is relevant to a certain segment-specific articulatory constraint.

As to focus realization in Japanese, it has been generally believed that the primary correlate to focus is the higher pitch of a focused word than that of the same word out of focus. Kôri (1989) has showed that "F0 is the most consistent correlate of focus"

whereas "the roles of duration and intensity are inconsistent". Experiment 7 in the present study has provided more supporting evidence for focus realization by pitch, since the monomoraic nouns do not vary in duration due to a focus placement, while apparently the pitch is influenced by the presence or absence of focus.

A question arises why accent and focus have no consistent effect on segmental duration in Japanese. Two accounts for this phenomenon are possible: One is a mora-timing constraint that attempts to maintain roughly equal duration of moras by temporally compensating segmental duration within a unit of a mora, typically a CV syllable. The other is a constraint induced by a phonological contrast between short and long vowels. It should be added here that these two constraints are closely related to each other. In non-prepausal position, a distinction of a monomoraic short vowel from a bimoraic long vowel is possible since there is a tendency to keep the length of one mora roughly equal in duration. On the other hand, each mora should be almost constant in duration since there is a phonological contrast of a short vowel vs. a long one.

It is suggested that these constraints prevent segmental durations from largely varying under accent or focus in non-prepausal position.

6.4 Phrase-final lengthening with no following pauses

Campbell (1992a) reports "phrase-related lengthening" in reading of sentences by four speakers of Tokyo Japanese. He first normalized the raw segmental durations to factor out any phone-specific articulatory differences and visualized the relative lengthening or shortening of segments. What he has found from it is that all speakers have quite similar tendencies in their durational variation,

which is closely related to the phrasing of the utterance. He states that "there is a tendency for both initial shortening and final lengthening throughout the phrases"(p.215). However, he neither defines the domain of the "phrase" in which he observed the phrase-final lengthening nor suggests the degree of lengthening he found. In addition, whether or not the phrase-final position accompanies a pause is not clear in his paper.

Campbell (1992b) has observed lengthening in particles, reporting /a/ in /wa/ is longer than its average duration. Campbell (personal communication) has also found that the vowel /a/ following /g/ is 9.6ms longer than the average duration (83.9ms) of all the 3380 occurrences of /a/ in his data. This suggests a lengthening of about 10% taking place in the vowel /a/ in the particle /ga/ in Japanese. Kaiki et al. (1992) also points out that a slight lengthening is supposed to take place in the final position of a unit that is roughly consistent with an accentual unit.

As has been pointed out in 6.2.3, moras in phrase-final position are longer by up to five percent than the same moras produced in non-phrase-final position. That is, /ne/ in *ya'ne nuketa* or *i'ne no'bita* is 6.3 to 7.1ms (4 to 5%) longer than its counterpart in *ya'ne ga nuketa* or *i'ne ga no'bita*. The former /ne/ locates before a major phrasal boundary between a NP and a VP, the latter does not. Thus, this slight lengthening in the former /ne/ suggests a phrase-final lengthening in the final position of NPs.

From all the above findings, it can be said that there is a slight phrase-final lengthening in Japanese, the degree of which has been reported to be within about 10%.

6.5 Lengthening of penultimate vowels in non-utterance-final position

One of the findings from this study is the consistent lengthening in penultimate vowel, i.e., the second vowel from the end of a pause-bounded unit (4.5.3). The vowels are longer in non-sentence-final position than the same vowels produced in sentence-final position. It is assumed that the increase in vowel duration contributes to the position-specific lengthening.

However, the degree of the lengthening is much smaller than that observed in prepausal vowels. As can be seen in Table4-5, the difference is 12.4ms averaged across sets, ranging from 3.1 to 21.6ms, whereas the prepausal vowels are lengthened by 42.4ms when collapsed across sets, ranging from 29.1ms to 59.7ms.

6.6 Durational variability in prepausal vowels

Throughout the current study, a wide variation in prepausal vowel duration is consistently observed. In isolated words and phrases, the final vowel of an utterance-final word with no accentual fall is lengthened on average by 42% when compared with its counterpart with an accentual fall; the maximum ratio between the former and the latter reaches 1.74:1.

In sentence-final position, the average ratio is about 1.43:1, and the max is 1.70:1. When all the prepausal vowels are compared for each set of long test sentences in Experiment 6, the longest prepausal vowel in NonSF-NoA is on average 74% longer than the shortest one in SF-A, ranging from 66% to 93%. This indicates that, taken together, the effects of accent and position have resulted in over 70% difference in prepausal vowels in sentential utterances.

In sum, the experimental results have clarified that the prepausal vowels show a wide range of durational variability depending on accent (the presence/absence of an accentual fall) and position (sentence-final or non-sentence-final position).

Compared with the other lengthening phenomena observed in Japanese (6.2 to 6.5), it can be said that by far a large amount of durational variation is observed restrictively in the prepausal position. The question we have to ask here is why such a large amount of accent- and position-related variability is allowed exclusively in prepausal position.

Several factors involved in temporal structure of speech should be examined one by one. First, it has been generally believed that segments before pauses consistently gain extra length. It is suggested that this prepausal lengthening partly contributes to the longer duration of prepausal vowels than those not followed by a pause in the present study. However, what we are concerned with here is not a consistent lengthening in prepausal vowels but a wide range of variation in the length of prepausal vowels. Thus, the prepausal lengthening cannot account for the durational variation observed in prepausal vowels.

Second, phonetic environments, i.e., the neighboring segments of prepausal vowels should be taken into account when discussing their temporal variance. However, in this investigation, a comparison has been made between prepausal vowels in words or phrases with exactly the same segmental structure except for the prepausal syllables in sentence-final and non-sentence-final positions in Experiment 6. In order to examine the effect of position in Experiment 6, /a/ in /ga/ was compared with /a/ in /da/ (4.5.2). The results showed that the change of the preceding consonant has caused duration increase within a small range of 1 to 7ms in

prepausal /a/. Thus, the degree of difference is too small an amount to account for the considerable position-related increase in prepausal vowels. Therefore, the effect of phonetic environments can be excluded from the main factors involved in the observed durational variation.

The last and most plausible account may be that there is a physiological or physical constraint in prepausal position that prevents a low-pitch vowel from being lengthened. In other words, when F0 drops abruptly towards the end of an utterance or a PBU, speakers have more difficulty in lengthening a prepausal vowel than they do when F0 does not drop. Thus, they end up with a shorter prepausal vowel as its F0 gets lower.

This interpretation is only a speculation, but had been the most plausible account until Experiment 6 presented the data that showed considerable lengthening of a prepausal vowel in a particle that follows a word with an accentual fall. It was found from the experimental result that the prepausal vowels in non-sentence-final position were lengthened regardless of their relatively low F0s. Recall the correlation coefficient between the prepausal duration and its F0 in non-sentence-final position was as low as 0.19 (4.5.5). Unfortunately, the assumption of the physiological and/or physical factor has turned out to be unable to account for this substantial lengthening in the prepausal vowel in non-sentence-final position. However, it does not deny a possibility that the physiological/physical factor may partly contribute to the accent-related significant difference in prepausal vowels in utterance-final position of isolated words and phrases and in sentence-final position.

In sum, none of these phonetic or physiological/physical factors can successfully account for the observed durational variation in prepausal vowels depending on accent (the presence or absence of an

accentual fall) and position (sentence-final or non-sentence-final position).

6.7 Markedness of prepausal vowels in Japanese rhythmic structure

6.7.1 *Prepausal position outside of mora-timing constraint*

It seems from the above discussion that Japanese speakers lengthen or shorten vowels immediately before pauses as they wish, whereas elsewhere they strictly avoid varying segmental duration to a large extent.

Section 1.1.2, has introduced the weaker version of mora-timing hypothesis in Japanese composed of two main claims. One is that Japanese speakers show a tendency to adjust segmental duration to maintain a roughly constant length of a mora, typically a CV syllable (Sagisaka & Tohkura, 1984; Campbell, 1999). The other is that there is no large increase in duration of Japanese moras under accent or focus.

The experimental results from the present study show that the first claim of the above hypothesis does not hold true of prepausal vowels. As pointed out in 4.5.1 and 5.3, all the prepausal vowels of the seven sets are of almost the same length with no statistic difference, despite differing segments that constitute the prepausal syllables (e.g. *te, to* and *ga*). In other words, the prepausal vowels do not compensate the length of their preceding consonants in an attempt to make the final mora, i.e., the prepausal CV syllable, roughly equal in length with the preceding moras. This anti-rhythmic phenomenon suggests that the temporal compensation in a unit of a mora is not valid for the prepausal mora.

It is also quite clear from the current study that the presence or

absence of an accentual fall results in a consistent and significant difference in the duration of prepausal vowels, while no consistent accent-related effect is observed in the other constituents. This indicates that the second claim of the mora-timing hypothesis that denies any consistent effect of accent on duration is not true of prepausal vowels.

Thus, it seems reasonable to conclude that prepausal vowels are free from the mora-timing constraints asserted by its weaker version hypothesis.

6.7.2 *Phonological constraint on vowel length in prepausal position*

Another aspect suggested from the wide durational variation in prepausal vowels is that the phonological constraint on vowel length weakens in the prepausal position.

In Japanese the length of a vowel is a distinctive feature that is recognized by Japanese speakers as 'making different words'. That is, there is a short vs. long opposition in vowel length. Phonetically, however, the length of a long vowel is not always twice as long as a short vowel (Campbell 1999b). Previous studies have shown the ratio of a heavy syllable (with two moras) to a light syllable (with one mora) is somewhere between 1.66:1 to 1.8:1.

Moreover, as has been observed in 6.2, even in a realization of the bimoraic phonological word, monomoraic nouns without particles are not allowed to be lengthened by over 50%. This suggests a phonological constraint on the lengthening of a mora in Japanese; that is, if a monomoraic word is lengthened by over 50%, it may be perceived as a different word with two moras.

Contrary to these findings, durational variation of over 50% in prepausal vowels is observed in some of the tokens obtained from Experiment 1 to 6 (6.6). It should be noted here that this wide

variation in duration takes place not between long and short vowels but within the same short vowel.

If the phonological constraint of vowel length in prepausal position is as strong as that in non-prepausal position, such a large durational variation will not be possible, or the vowels lengthened by over 50% will be recognized as long vowels in contrast to their non-lengthened short ones. Apparently, however, the durational variation in prepausal vowels did not affect the meaning of the word that speakers produced in the current experiments. Thus, the wide durational variation in prepausal vowels suggests that the phonological constraint on vowel length is weakened in prepausal position and phonetic contrast between short and long vowels are neutralized in that position[4].

Shortening of word-final long vowels has been repeatedly observed on a phonetic level as well as on a phonological level (Sukegawa, Maekawa & Uehara, 1999; Kubozono, 2002). Kubozono (2002) claims that the "position-dependent" shortening of a long vowel indicates the "temporal neutralization" of word-final vowels.

Although most of the previous studies have reported the shortening of a long vowel only, Mori (2002b) reports that, in truncated forms of compound words or phrases, it is sometimes the case that the word- or phrase-final short vowel /i/ or /u/ is lengthened to be a long vowel (p. 126).

hana piasu	→	*hanapii*
heso piasu	→	*hesopii*
banban busaiku	→	*banbuu*
banban mukatuku	→	*banmuu*
banban kimoi	→	*bankii*

What the above data suggest is that in word- or phrase-final position, not only shortening of a long vowel but also lengthening of a short vowel are observed in the formation of clipped compounds. Although this phenomenon allegedly takes place in word final position, there is a possibility that the phonological shortening and lengthening specific to word- or phrase-final position may reflect the observed durational variability in utterance-final or prepausal position since when uttered in isolation word- or phrase-final position is consistent with utterance-final or prepausal position.

6.7.3 *Markedness of prepausal vowels in Japanese rhythmic structure*

The previous sections (6.7.1 and 6.7.2) present two interpretations of results from Experiments 1 to 7: 1) prepausal position is outside of the constraint of mora timing, and 2) phonological constraint on vowel length weakens in prepausal position.

According to these two claims, prepausal position makes a striking contrast with non-prepausal positions in Japanese where the timing of moras (typically CV syllables) are strongly constrained by isochronous tendency of moras and phonological contrast of short and long vowels. Therefore, it seems reasonable to claim that prepausal position is a marked position in Japanese prosody in that it is the only position where a wide range of variation in vowel duration is allowed.

---

[1] The verb phrase *nuketa* is unaccented whereas *no'bita* is initially accented. In utterances without the particle /ga/, /ne/ in *ne' nuketa* is immediately followed by unaccented mora /nu/, whereas /ne/ in *ne' no'bita* is followed by an accented mora /no/ of *no'bita*. The

experiment is designed to examine the effect of pitch accent of the adjacent mora on the monomoraic lengthening.

[2] In these statistic analyses, RM-ANOVAs were not exactly two-way. One factor was done by RM-ANOVA, and the other by ANOVA, since in the case of the former factor (the presence or absence of a particle) there was an exact correspondence of the tokens, whereas it was not the case for the latter factor. It was confirmed, however, whether by ANOVAs or RM-ANOVAs, the statistic results were the same. Only differences were the degree of freedom.

[3] Although /neo/ and /nega/ are phonologically of equal length (two moras), /neo/ is made up of only one consonant and two vowels (CVV), whereas /nega/ consists of two pairs of one consonant and one vowel (CVCV).

[4] It should be noted that phonological contrast between short and long vowels is maintained in prepausal position; for example, there are pairs of lexical items such as *sa'do* vs. *sa'doo* that have phonological contrast between short and long vowels even if the final vowels are prepausal when uttered in isolation.

CHAPTER 7

Theoretical implications and concluding remarks

7.1 Introduction

Chapter 6 has clarified the fact that prepausal position is free from the mora-timing constraint in Japanese. In addition, the markedness of prepausal position has been claimed in the discussion of the position-specific durational variability in prepausal position in contrast to durational invariability in non-prepausal position.

It should be repeated here that the prepausal vowels do not vary irregularly. Instead, they show a similar durational pattern across both sentences and speakers depending on accent (the presence/absence of accentual fall) and position (sentence-final or non-sentence-final position), suggesting some systematic variation.

This chapter explores what this wide durational variability specific to prepausal position suggests in Japanese prosody. In addition, it addresses the theoretical implications of the markedness of the prepausal position in Japanese rhythmic structure.

7.2 *Final shortening in Japanese vs. final lengthening in English*

In discussing what is implied by durational variability in prepausal position in Japanese, it is worth asking what of the durational patterns observed in this study are specific to Japanese and what are universal to languages. However, it is impossible to compare all the languages in the world and draw a line between language-universal and language-specific phenomena. Therefore, in this thesis, the observed variability in Japanese prepausal vowels is compared with the well-known final lengthening in English.

Final lengthening in English has been investigated in detail by many researchers (cf. Klatt, 1975; Cooper & Danly 1981; Edwards & Beckman 1988; Turk & White, 1999). Corpus analyses by Wightman, Shattuck-Hufnagel, Ostendorf & Price (1992) show that final lengthening in English is consistently observed in the rhyme of a syllable immediately preceding prosodic boundaries. That is, final lengthening is realized in the nucleus and coda of the final syllable of phonological units such as a phonological phrase or an intonational phrase.

However, little has been known about how final lengthening is realized in an unstressed English syllable in the final position. Mori (2001b) has carried out several experiments in which test sentences were designed to have stressed vs. unstressed syllables composed of the same segments in sentence-final vs. non-sentence-final positions, thus making it possible to compare durations between stressed vs. unstressed syllables and between sentence-final vs. non-sentence-final syllables. Results from these experiments show that, when the length of a sentence-final segment is compared with that of the same segment produced in non-sentence-final position, the final consonant,

i.e., the coda of an unstressed syllable is lengthened as much as its counterpart of a stressed syllable in sentence-final position. When an unstressed final syllable has no coda, a strong aspiration that follows the sentence-final vowel is consistently observed across speakers.

These results suggest that the lengthened coda and the long aspiration period in sentence-final position may contribute to the realization of final lengthening in English, although apparently speakers have difficulties lengthening the unstressed sentence-final vowel to a great degree when F0 drops abruptly.

Contrary to the above observation in English, no consistent aspiration is observed at the end of Japanese sentential utterances in the present study. Instead, a far shorter vowel is frequently observed when F0 drops sharply by accent nucleus in a sentence-final or utterance-final word/phrase.

Difference in the durational realization of final unaccented (or unstressed) vowels between English and Japanese implies that the final lengthening is obligatory in English, whereas in Japanese it is optional. In other words, it seems that Japanese speakers do not always have to lengthen the final vowel.

Takeda et al. (1989) and Campbell (1992b) have observed through their corpus analyses that the lengthening in Japanese sentence-final position is less marked than the lengthening in the other utterance-final positions. They call the phenomenon "sentence-final shortening", although they do not consider the effect of accentual fall on its occurrence. In comparison with the final lengthening of English, the accent-related durational decrease observed in final vowels of sentence-final words or phrases with an accentual fall in this study is referred to as "sentence-final shortening" in this thesis.

It is suggested from the above difference in the durational realization of sentence-final vowels that syntactic differences between the two languages may influence the phonetic realization of sentence-final vowels. English sentences tend to end with content words, i.e., information carriers. Moreover, there is a well-known "end focus" in English, a tendency to place a focus at the end of an utterance or a sentence if no other element is focused on contextually. It is thereby assumed that English speakers tend to make the final element of an utterance or a sentence phonetically salient.

Japanese sentences are most likely to have a function word in sentence-final position, that is, words that are predictable and in some cases redundant. Thus, it can be assumed that in Japanese sentential utterances, non-lengthening or shortening of the sentence-final vowel is allowed since it is often the case that the final vowel is not crucial as an information carrier. Further studies that explore the realization of final lengthening in agglutinative languages other than Japanese are required in order to provide supporting evidence for this assumption.

7.3 *Durational variability and intonation in prepausal position*

It has been pointed out many times that the aspects of prepausal or sentence-final tones or intonation, i.e., pitch contour, is of great importance in Japanese since the tone in that position is assumed to have a variety of significant functions (cf. Sugitô, 1983; Uemura, 1989; Maekawa, 1999). It is hypothesized that the wide durational variability in prepausal vowels revealed by this study makes it possible to realize a variety of prepausal tones.

Although previous studies have mainly placed focus on pitch

realization in prepausal position, it is presumed that lengthening of prepausal vowels shares an important role in the realization of the wide range of pitch before a pause.

Durational increase in prepausal position is most evident when a speaker attempts to realize a marked rising and/or falling intonation. In Japanese, for instance, questions can be asked with a rising intonation and the lengthening of a sentence-final vowel. Kôri (1997) reports that the F0 starts to rise from the sentence-final mora in this rising intonation and its duration tends to increase accordingly. Anecdotal observation suggests that final /u/ in *Moo ka'eru?* "Will you go home now?" in an interrogative sentence, for instance, is much longer than its counterpart in a declarative sentence *Moo ka'eru* "I'll go home now." It is also observed in our daily conversation that a speaker can drawl a prepausal vowel considerably with a rise-fall intonation, signaling that s/he intends to continue speaking and does not want the listener to interrupt her/his speech; thus it could function as a turn holder.

It is suggested from these findings that a wide range of pitch fall or rise tends to be realized in prepausal position since a great amount of durational increase required to realize the pitch change is allowed in that particular position in Japanese.

It seems that not only a tone of marked rising or falling but also a variety of prosody can be realized by the shortening or lengthening of a prepausal vowel. In daily conversation, it is possible to lengthen the sentence-final vowel even twice or three times as long as its non-lengthened form, if a speaker intends to fill a pause or stall time to think about the next thing to say. Thus, the lengthened vowels could function as a filler or a staller. When the pitch of a sentence-final vowel is abruptly lowered and its length is somewhat

shortened, a speaker's surprise to the information or semantics of a sentence (proposition) could be expressed: *Ka're ga ku'ru* "He is coming!" Or a remarkably short and intensified final vowel with a similarly intensified final syllable or word in sentence-final position could be imperative as in the case of *Koko ma'de ku'ru* "Come here!".

The above findings suggest that elastic final vowels in prepausal position enable speakers to realize a variety of prepausal prosody that has each specific function.

7.4 Phonetic contrast between short and long vowels in prepausal position

Section 6.6.2 has claimed that, in prepausal position, phonological constraint on vowel length weakens and the phonetic contrast of short and long vowels is neutralized.

In daily, informal Japanese conversation, it is frequently observed that long vowels in prepausal position are substantially shortened and that short vowels in prepausal position are lengthened to large degrees.

For instance, *Sa'a iko'o* [sa:iko:] "OK, let's go!" or *So'o desyo'o* [so:deʃo:] "That's it, isn't it?" tends to be pronounced as [sa:iko] and [so:deʃo], respectively, in daily conversation. When this happens, the sentence-final vowel /o/ is considerably shorter, and thus the utterance is often transcribed as *Sa'a iko* or *So'o desyo,* with a short vowel in sentence-final position.

When transcribing speech, in comics or on a bulletin board of an Internet site, for instance, it is frequently observed that prepausal short vowels are transcribed as long vowels as in the case of *Tada'imaa* "Hullo, here I am!", *Konnitiwaa* "How are you?", *Yorosikuu* "I am glad to make your acquaintance." It can be assumed that

these transcriptions reflect lengthened prepausal short vowels in the utterances.

A question may arise as to how long the shortened or lengthened vowels are phonetically or phonologically. Are they monomoraic, bimoraic, or trimoraic? Phonetically, the shortened /o/ in [iko] or [deʃo] may last only half as long as a non-shortened long vowel length. Or the lengthened /a/ and /u/ in the latter two examples may be two, or more than two times as long as non-lengthened forms. However, in daily conversation, we do not have any difficulty in understanding the meaning of the utterances, even if the prepausal vowel is substantially shortened or lengthened. Instead, the shortening or lengthening of the final vowels appears to give us the information about the speaker's speaking style, or attitudes or emotions towards the listener, as in these examples.

The phenomena illustrated above tend to be considered as representing a certain type of sentence-final intonation. Therefore, no previous studies have ever answered the crucial question as to why the marked lengthening of short vowels and shortening of long vowels occur restrictively in prepausal or utterance-final position in Japanese that has previously been claimed to be mora-timing.

The wide range of durational variation exclusively observed in prepausal vowels throughout this study can successfully account for the phenomena. As has been clarified above, a great amount of elasticity in prepausal vowels is available in this specific position, suggesting that the position is not subject to a mora-timing constraint and that phonetic contrast between short and long vowels is neutralized in this position. This markedness of prepausal position makes it possible for prepausal long vowels to shrink into almost one mora, or for short vowels to stretch to the length of two or

more moras. It can also be said that the shortening of a long vowel and the lengthening of a short vowel observed in prepausal position may reflect the wide durational variability specific to the prepausal position that is in turn outside the domain of mora timing.

7.5 Prosody of sentence-final and interjectional particles

Various Japanese particles that are attached to the final position of sentences and phrases are referred to as sentence-final particles and interjectional particles (Yamada 1936)[1], respectively. Yoshida (1984) defines the major function of sentence-final and interjectional particles as that of "exhibiting the relationship with listeners" (p.29). Maynard (1999) refers to sentence-final particles as "the interactional" particles (p.439).

Recently, more attention has been given to the relationship between the functions of various sentence-final particles and their corresponding tones. Although analyses have mostly focused on the pitch of the final syllable, which in most cases consists of one mora (e.g., /yo/ or /ka/), specific pitches tend to combine with varied durations (Kôri 1990; Todoroki 1998; Sugitô 2002). Todoroki (1998) deliniates between short and long durational patterns of the final particles /ne/, /na/, and /ka/, in her analyses of their specific functions.

The sentence-final particle /ka/, for instance, can mark a question with a half-lengthening of the final vowel and a rising intonation as in *Ge'nki ka* "Are you fine?". As has been illustrated in 7.3, this interrogative sentence can be uttered without the particle /ka/ but with the same prosody. A similar rising intonation with a half-lengthening of the vowel in the particle /yo/ could emphasize a warning as in *Abunai yo* "Hey, That's dangerous!"

The interjectional particle /ne/ occurs both in phrase-final and sentence-final positions and may be followed or not followed by a pause. When followed by a pause, however, the particle is most likely to be subject to substantial lengthening. It has been assumed that the particle /ne/ has the general function of softening a statement and inviting confirmation from the listener. A mother could tell her little child a story by frequently inserting /ne/ and a pause: *Oka'asanwa nee kino'o nee outide nee ...* "Yesterday, at home, Mama was, ..." In this case, the slow and gentle final /ne/ with a rising intonation seems to show her affection and intimacy with her child.

Considering all of the above findings, it is suggested that the wide variability in prepausal vowel duration contributes to realizing various kinds of prosodic patterns (specific combinations of pitch and duration) of sentence-final or interjectional particles attached to the prepausal position. It may further be hypothesized that these particles occur, or tend to occur, in prepausal position because that is the position where a large amount of durational variability required to realize their variety of prosody is allowed.

7.6 Functions of prosody in prepausal vowels

In the above examples of 7.3 to 7.5, there seems to be a common point among what is being expressed by the various prosodies of the prepausal vowels. Some of them are concerned with the speaker's emotions or attitudes towards the listener (or to what they say), and/or with the relation to their conversational partners. That is, they have interactional or interpersonal functions that are similar to the functions of discourse markers in English have: e.g. fillers, stallers, turn holders, attention getters, hesitators, intimacy signals,

etc. (cf. Mori, 1996).

In addition, it appears that the prepausal prosody is mainly concerned with the pragmatic aspects[2] of utterances, whereas constituents in non-prepausal position play a crucial role in semantics and syntax of the proposition.

Therefore, it is hypothesized that the prepausal position is a special position in which pragmatic aspects of utterances such as a speaker's attitudes, styles, intentions, comments, or emotions, towards what s/he says or to what the listener says can be displayed with the use of a variety of prosody made available by the marked elasticity in prepausal vowels and each PBS or utterance is connected into a discourse or conversation.

7.7 Prepausal vowels in Japanese rhythmic structure

In the previous chapter (Chapter 6), it is argued that the prepausal position is outside of the mora-timing constraint and that its wide durational variability provides a striking contrast with the durational invariability of the non-prepausal position.

The contrast in elasticity between prepausal position and non-prepausal position is evident in examples of daily conversation. If a speaker lengthens sentence-final /a/ as in *Sono to'oridaa* [sonoto:rida:] "That's right.", the speaker could express that s/he is deeply impressed by what her/his conversational partner has said and that s/he is totally convinced of its veracity. However, if a speaker lengthens non-prepausal /a/ as in *Sono to'oridaa to omo'u* [sonoto:rida:toomo:] "I think that's right." his or her utterance may sound odd since /a/ is too long to be one mora; that is, it goes against normal mora timing and sounds anti-rhythmical. Or the utterance may give his or her listener difficulty in understanding its meaning

since the lengthened [da:] sounds like a different morpheme with a long vowel. Contrary to this, if the speaker inserts a pause between "*Sono to'oridaa*" and "*to omo'u*", the final /a/ can be lengthened to be [sonoto:rida:].

As has been introduced in 1.1.2, Japanese has been long described as mora-timed language in contrast to stress- or syllable-timed languages such as English or Spanish. Although the exact equal length of moras has been denied by acoustical experiments in which actual duration of moras were measured, isochronous tendency to stretch or shrink segments in an attempt to make the duration of each mora constant has been repeatedly reported. Some researchers have proposed phonological units larger than a mora such as bimoraic foot (cf. Poser 1990). Yet, the phonological units they have proposed are assumed to consist of moras that are under the mora-timing constraint.

Contrary to all of the previous studies that support the mora-timing hypothesis in one way or another, the current study has revealed that the prepausal position is outside the domain of mora timing. Therefore, the present study draws a clear line between non-prepausal position and prepausal position in Japanese rhythmic structure. That is, the experimental results have demonstrated that Japanese has two distinct temporal structures: a rigid and invariable structure governed by the mora-timing hypothesis, and a variable and elastic structure free of any mora-timing constraint. Although the latter is restricted to the end of an utterance or PBU, there is a possibility that the varied duration combined with a variety of tones may make the prepausal vowel salient in spoken Japanese and fulfill an important function in Japanese communication.

7.8 *Concluding remarks*

The present study has clarified the argument that the presence or absence of an accentual fall and sentence-final or non-sentence-final positions result in consistent and significant differences in prepausal vowel duration to a substantial degree. This study has also made it clear that, except for prepausal vowels, accent has no consistent effect on duration and that segmental duration stays almost unvaried regardless of the location and presence/absence of an accentual fall. It follows from these findings that prepausal position is a marked position in Japanese that is out of the domain of mora timing.

It is suggested from all of the above findings and discussions that Japanese has two distinct and contrastive temporal structures. In non-prepausal position, the language has a rigid temporal structure in which neither a speaker's intentional shortening nor lengthening is allowed and segments strictly follow the isochronous tendency of moras and phonological contrast between short and long vowels. In prepausal position, by contrast, Japanese has elastic final vowels that make it possible to realize a variety of prosody that enable speakers to display pragmatic aspects of their speech such as their attitudes, styles, intentions, comments, or emotions, to their listeners or to what they say.

---

[1] The classification of particles is in accord with that by Yamada (1936). In his six types of particles, /ne/ is classified into interjectional particles, whereas some phoneticians (Todoroki 1998; Sugitô 2002) consider /ne/ that occurs in sentence-final position sentence-final particles.

[2] The term "pragmatic aspects" is used here to refer to whatever is concerned with the use of language in an act of communication: the

purpose and intention that a speaker has in using the language, whether or not s/he is conscious of it; the way and situation in which the language is used; the effect that use of the language has on a listener and/or on the speaker.

References

Beckman, M. (1982) Segment duration and the 'mora' in Japanese, *Phonetica,* **39,** 113-135.

Beckman (1986) *Stress and non-stress accent.* Dordrecht: Foris.

Beckman, M. & Pierrehumbert, J. (1986) Intonational structure in Japanese and English, *Phonology Yearbook* **3,** 255-309.

Bloch, B. (1949) A set of postulates for phonemic analysis, *Language*, **24,** 3-46.

Bloch, B. (1950) Studies in colloquial Japanese. IV: Phonemics. *Language,* **26**, 86-125.

Bloomfield, L. (1926) A set of postulates for the science of language, *Language*, **2,** 153-164.

Cambier-Langeveld, T. & Turk, A. E. (1999) A cross-linguistic study of accentual lengthening: Dutch vs. English, *Journal of Phonetics*, **27**, 255-280.

Campbell, N. (1992a) Speech timing in English and Japanese. In *Symposium on Research on Japanese and its Pedagogical Applications,* pp. 207-216.

Campbell, N. (1992b) Segmental elasticity and timing in Japanese speech. In *Speech Perception, Production and Linguistic Structure* (Y. Tohkura, E. Vatikiotis-Bateson, Y. Sagisaka, editors), pp. 403-418. Tokyo: Ohmsha.

Campbell, N. (1999) A study of Japanese speech timing from the syllable perspective, *Journal of the Phonetic Society of Japan,*

3, 29-39.

Cooper, W.E. & Danly, M. (1981) Segmental and temporal aspects of utterance-final lengthening, *Phonetica*, **38**, 106-115.

Edwards, J.; Beckman, M. (1988) Articulatory timing and the prosodic interpretation of syllable duraion. *Phonetica* 45, 156-174.

Fletcher, J. (1991) Rhythm and final lengthening in French, *Journal of Phonetics*, **19**, 193-212.

Fries, C. (1949) *The structure of English.* New York: Harcourt, Brace and Company.

Fujisaki, H. & Nagashima S. (1969) A model for the synthesis of pitch contours of connected speech, *Annual Report of the Engineering Research Institute.* **28,** 53-60.

Fujisaki, H. (1992) The role of quantitative modeling in the study of intonation. In *Symposium on Research on Japanese and its Pedagogical Applications*, pp. 163-174, Nara, Japan.

Fujisaki, H. (2002) Modeling in the study of tonal features of speech with application to multi-lingual speech synthesis. In *Proceedings of the Joint International Conferece of SNLP-Oriental COCOSDA 2002*, pp. 1-10, Hua Hin, Thailand.

Han, M. (1962) The feature of duration in Japanese, *The Study of Sounds,* **10,** 65-80.

Han, M. (1994) Acoustic manifestations of mora timing in Japanese, *Journal of the Acoustical Society of America*, **96**, 73-82.

Hattori, S. (1968) *Eigo kiso goino kenkyûu* (A study in the basic vocabulary of English). Tokyo: Sanseidô.

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.

Itô, J. (1990) Prosodic minimality in Japanese. In *Papers from the Parasession on the Syllable in Phonetics and Phonology, CLS* **26-II** (K. Deaton, M. Noske, & M. Ziolkowski, editors), pp. 213-239.

Kaiki, N. & Takeda, K., & Sagisaka, Y. (1992) Vowel duration control using linguistic information, *IEICE Transactions,* **J75-A**, 467-473.

Kindaichi, H. (1967) *Nihongo onin no kenkyû* (Studies in Japanese phonology). Tokyo: TôKyôdô-syuppan.

Klatt, D. H. (1975) Vowel lengthening is syntactically determined in a connected discourse, *Journal of Phonetics*, **3**, 129-140.

Kôri, S. (1989) Acoustic manifestation of focus in Tokyo Japanese: the role of intensity, duration and Fo, *Onsei Gengo,* **3**, 29-38.

Kôri, S. (1990) Tones and communicative functions of sentence-final 'ka', *Onsei Gengo,* **4**, 1-25.

Kôri, S. (1997) Nihongo no intonêsyon (Intonation of Japanese). In *Nihongo-onsei (Japanese Phonetics) 2: Akusento intonêsyon rizumu to pôzu (Accent, intonation, rhythm and pause)* (M. Sugitô, editor), pp.169-202. Tokyo: Sanseido.

Kubozono, H. (1993) *The Organization of Japanese Prosody.* Tokyo: Kurosio Publishers.

Kubozono, H. (2002) Temporal neutralization in Japanese. In *Papers in Laboratory Phonology* VII. (C. Gussenhoven & N. Warner editors.), pp.171-201: Mouton de Gruyter.

Lehiste, I. (1977) Isochrony reconsidered, *Journal of Phonetics*, **5,** 253-263.

Lehiste, I. (1979) The perception of duration within sequences of four intervals, *Journal of Phonetics*, **7,** 313-316.

Liberman, M. & Pierrehumbert, J. (1984) Intonational invariance

under changes in pitch range and length. In *Language Sound Structure*, (M., Aronoff & R. T., Oehrle, editors): MIT Press.

Maekawa, K. (1999) Prosody and communication, *Journal of the Acoustical Society of Japan*, **55-2**, 119-125.

Maynard, S. (1993) *Kaiwa bunseki* (Conversation analysis). Tokyo: Kuosio Pushlishers.

McCawley, J. (1968) *The phonological component of a grammar of Japanese*. The Hague;Paris: Mouton.

Mitsuya, F. & Sugitô, M. (1978) A study of the accentual effect on segmental and moraic duration in Japanese. In *Annual Bulletin*, **12**, 97-112, Research Institute of Logopedics and Phoniatrics, University of Tokyo.

Mori, Y. (1996) Pauses and silence fillers in spoken discourse & time-based perceptual sense unit. Master thesis submitted to Kobe City University of Foreign Studies.

Mori, Y. (2001a) Effect of accentual fall on final lengthening in Japanese. *Journal of Phonetic Society of Japan,* **5-1**, 92-106.

Mori, Y. (2001b) Effect of accent on final lengthening in English. In *Proceedings of the 15th General Meeting of the Phonetic Society of Japan,* pp. 163-168.

Mori, Y. (2002a) Lengthening of Japanese monomoraic nouns. *Journal of Phonetics*, **30-4,** 689-708.

Mori, Y. (2002b) Formation of three-mora compound abbreviations, *Journal of Phonetic Society of Japan,* **6-1**,121-137.

Mori, Y. (2002c) Correlation between pitch and duration in utterance-final position & the prosody of sentential utterances. In *Proceedings of the 16th General Meeting of the Phonetic Society of Japan,* 139-144.

Mori, Y. (2003) Isochrony and variability in Japanese moras. To

appear. *Journal of the Japan Society of Speech Sciences,* **3**.

Mori, Y. Markedness of prepausal vowels in Japanese rhythmic structure. *Journal of Phonetics* (in review).

Nakatani, L., O'Connor, K. & Aston, C. (1981) Prosodic aspects of American English speech rhythm, *Phonetica,* **38**, 84-106.

Oyakawa,T. (1971) How does the Length of Utterance Affect the Segmental Duration? In *Monthly Internal Memorandum* August, 21-35. University of California, Berkeley.

Peterson G. & Lehiste, I. (1960) Duration of syllable nuclei in English, *Journal of the Acoustical Society of America,* **32**, 693-703.

Pierrehumbert, J. (1980) *The phonology and phonetics of English intonation.* PhD thesis, MIT. Published by Indiana University Linguistics Club.

Pierrehumbert, J. & Beckman, M. (1988) *Japanese Tone Structure.* MIT Press: Cambridge, Massachusetts.

Port, R. & Al-Ani, S. & Maeda, S. (1980) Temporal compensation and universal phonetics, *Phonetica,* **37**, 235-252.

Port, R.; Dalby, J.; O'Dell, M. (1987) Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America,* **81**, 1574-1585.

Poser, W. (1990) Evidence for foot structure in Japanese, *Language* **66**, 78-105.

Sagisaka, Y. & Tohkura, Y. (1984) Phoneme duration control for speech synthesis by rule, *IEICE Transactions,* **67-A**, 629-636.

Sakaguchi, T. & Yoshida T. (1999) Correlation between duration and fundamental frequency of syllables. *Techinical report of IEICE,* **SP99-87**, 9-16.

Satô, K. (1977a) Kisoku ni yoru onsei-goosei no kenkyû (Study on rules for synthesizing speech). Ph. D dissertation, Hokkaidô

University.

Satô, K. (1977b) Tango ni okeru onin zizokuzikan to hassei no taimingu (Segmental duration and timing location in speech). *Onseikenkyuukai Siryô,* **S77-31,** 1-8: The Acoustical Society of Japan.

Satô, K. (1994) Bimoraic rhythm and accentuation in Japanese. In *Conference Handbook for the Twelfth National Conference of the English Linguistic Society of Japan,* 85-90: English Linguistic Society of Japan.

Sluijter, A. M. C. & van Heuven, V. J. (1995). Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch, *Phonetica,* **52**, 71-89.

Sugitô, M. (1975) Osaka hôgen ippakugo akusento no pitti-kyokusen to zizokuzikan ni tuite (Pitch pattern and duration of one-mora words of the Osaka dialect), *Syôin-Kokubungaku,* **13**, 80-112.

Sugitô, M. (1983) Nihongono akusentoto intonêsyon (Japanese accent and intonation). *Kotobato Onsei (Language and Sound),* 55-70.

Sugitô, M. (1990) Nihongoto eigono akusentoto intonêsyon (Accent and intonation of Japanese and English). In *Nihongoto Nihongo Kyôiku* 3(2) (M. Sugitô, editor), 349-378.

Sugitô, M. (2002) Syûzyosi 'ne' no imi/kinôto intonêsyon (Meaning, function, and intonation of the sentence-final particle 'ne'). In *Bunpôto Onsei,* **3** (Spoken Language Working Group, editor), pp. 3-16. Tokyo: Kurosio Publishers.

Sukegawa, Y., Maekawa, K., & Uehara, S. (1999) Nihongo tyôboinno tansyukuka gensyôo meguru syoyôinno zikken onseigakuteki kenkyûto nihongo onseikyôikueno sisa (An experimental phonetic study of Japanese vowel shortening and

its implications to pronunciation teaching). In *Gengogakuto Nihongo Kyôiku* (Linguistics and Japanese Language Education) (Y. Sasaki Alam, editor).

Takeda K.; Sagisaka, Y.; Kuwabara, H. (1989) On sentence-level factors governing segmental duration in Japanese. *Journal of the Acoustical Society of America,* **86**, 2081-2087

Todoroki, S. (1998) Tôkyô hôgenno syûjosi ne, na, ka no ontyô ni tuite. (On tones of sentence-final particles /ne/, /na/, and /ka/, in Tokyo Japanese). In *Kotobano Sinrito Gakusyû,* (Kotobano Kagaku Kenkykai, editor), pp.79-92: Kinseidô.

Turk, A. E. & White, L. (1999) Structural influences on accentual lengthening in English, *Journal of Phonetics*, **27**, 171-206.

Uemura, U. (1989) Nihongono intonêsyon. (Japanese intonation) *Kotobano Kagaku* **3**, 193-220.

Wightman, C. & Shattuck-Hufnagel, S & Ostendorf, M. & Price, P. (1992) Segmental durations in the vicinity of prosodic phrase boundaries, *Journal of the Acoustical Society of America,* **91**, 1707-1717.

Yamada, Y. (1936) *Nihon Bunpôgaku Gairon.* (An introduction to Japanese Grammar). Hôbunkan:Tokyo.

Yoshida, K. (1984) Jozitowa Nanika. (What are Japanese particles?). In *Kenkyû Siryô Nihon Bunpô.* (Materials for Japanese Grammar research) (K., Suzuki & Hayashi, editors), pp. 1-32, Tokyo: Meiji Shoin.

Yoshizawa, N. (1960) Intonêsyon. (Intonation). In *Hanasi Kotobano Bunkei* (1) (A research for making sentence patterns in colloquial Japanese), 178-208. The National Language Research Institute.

Appendix 2-1.

Test words, phrases and sentences for Experiment 1 and 2

Carrier sentence: *"Mooitido__to ittekudasai"* (Please say__again.)

[na-set]

na	"name"
nami'	"wave"
na'mida	"tears"
namidame	"eyes with tears"
namidame'ku	"be moved to tears"
namidana'gara	"with tears in one's eyes"
namidana'garani	"with tears in one's eyes"

[ma-set]

ma	"space" or "pause"
ma'do	"window"
madori	"room arrangement"
madori'zu	"floor plan"

[su-set]

su'	"nest"
sumi'	"coal"
sumi'ya	"charcoal dealer"
sumiyaki'	"charcoal making"
sumiyakiba	"coalpit"
sumiyaki'gama	"charcoal kiln"

[si-set]

si'	"city"
si'go	"private talk"
sigoto	"work"
sigotoba	"workplace"
sigotoba'ka	"workaholic"
sigotoba'kari	"only working"

[Test sentences for Experiment 2]

na | ga || sireru	"The name is known."
ma | ga || no'bi | ta	"A pause is/was prolonged."
ma | nobi | si | ta	"Too long of a pause is/was inserted

Appendix 3-1.

Test words for Experiment 3

[Sets of two-mora words]

asa'	"flax"
a'sa	"morning"
hasi	"edge"
hasi'	"bridge"
ha'si	"chopsticks"
sake	"alcohol"
sa'ke	"salmon"
tabi'	"journey"
ta'bi	"(Japanese) socks"

[Sets of three-mora words]

atari	"hit"
a'tari	"neighborhood"
atusa	"thickness"
a'tusa	"hotness"
itigo	"strawberry"
i'tigo	"strawberry"
iti'go	"span of life"
kakeru	"lack"
kake'ru	"hang"
kisetu	"being established"
ki'setu	"season"
kise'tu	"season"

[Sets of four-mora words]

matutake	"*matsudake* mashuroom"
ma'tutake	"pine & bamboo"
neagari	"rise in price"
nea'gari	"roots above ground"
senkyoku	"selection of music"
senkyo'ku	"election district"
senkoo	"selection"
se'nkoo	"incense sticks"
sentoo	"front"
se'ntoo	"public bath"

Appendix 3-2. Test phrases for Experiment 4A

[Sets of three-mora phrases]

kono to	"this door"
ki'no to	"a wooden door"
kono na	"the name of a child"
ki'no na	"the name of a tree"

[Sets of four-mora phrases]

umeno mi	"fruit of a plum tree"
gu'mino mi	"fruit of a *gumi* tree"
zino sake	"the *sake* of the district"
zino sa'ke	"the salmon of the district"
mino gara	"pattern of nuts"
ki'no gara	"pattern of trees"
sakeno na	"the name of *sake*"
sa'keno na	"the name of a salmon"

[Sets of five-mora phrases]

momono gara	"pattern of peaches"
he'bino gara	"pattern of snakes"
uekino na	"the name of a plant"
sa'kusyano na	"the name of an author"
hana'yano na	"the name of a florist"
totino sake	"*sake* of the district"
totino sa'ke	"salmon of the district"

[Sets of six-mora phrases]

kamino atusa	"thickness of sheets of paper"
natuno a'tusa	"hotness of summer"
siborino gara	"white-spot patterns"
o'tibano gara	"pattern of fallen leaves"
higa'sano gara	"pattern of parasols"
sakeno umami	"deliciousness of alcohol"
sa'keno umami	"deliciousness of salmon"

[Sets of seven-mora phrases]

tonarino ueki	"next-door neighbor's plants"
ka'danno ueki	"plants in a flower bed"
hana'yano ueki	"plants in a florist's"
murano matutake	"*matsudake* mashuroom of a village"
murano ma'tutake	"pine & bamboo (decoration) of a village"
kirino getabako	"a shoe cupboard made of paulownia"
ka'sino getabako	"a shoe cupboard made of oak"

[Sets of eight-mora phrases]

yasaino neagari	"rising price of vegetables"
i'hukuno neagari	"roots above ground"
o'okuno senkyoku	"selection of many songs"
ti'isana senkyo'ku	"a small election district"
zyoozaino kikime	"effect of tablets"
syo'oyakuno kikime	"effect of coarse-cut medicines"

Appendix 3-3. Test phrases for Experiment 4B

[siro-set]

sirono kabe	"walls of a castle"
si'rono kabe	"white walls"
mi'gotona siro	"a splendid castle"
mi'gotona si'ro	"beautiful white"

[tabi-set]

tabino na'gasa	"length of a travel"
ta'bino na'gasa	"length of Japanese socks"
naga'i tabi'	"a long journey"
naga'i ta'bi	"long (Japanese) socks"

Appendix 4-1. Test sentences for Experiment 5

A-Group

[hareru-set]

kao-ga hareru	"The/My face is swollen."
ka'ta-ga hareru	"The/My shoulder is swollen."
kao-ga hare'ru	"The/My countenance lightens."
so'ra-ga hare'ru	"The sky is clear."

[hareta-set]

kao-ga hare-ta	"The/My face was swollen."
ka'ta-ga hare-ta	"The/My shoulder was swollen."
kao-ga ha're-ta	"The/My countenance lightened."
so'ra-ga ha're-ta	"The sky was clear."

[tukeru-set]

kyu'uri-o tukeru	"I pickle cucumbers."
kyu'uri-o tuke'ru	"I add cucumbers."

[tuketa-set]

kyu'uri-o tuke-ta	"I pickled cucumbers."
kyu'uri-o tu'ke-ta	"I added cucumbers."

B-Group

[sake-set]

sake-ga totemo oisi-so'oda

"The *sake* looks very delicious."

sa'ke-ga totemo oisi-so'oda

"The salmon looks very delicious."

[siro-set]

yama-no tyoozyo'o-no siro-ga mi'e-ta

"I can see the castle on the top of the mountain."

yama-no tyoozyo'o-no si'ro-ga mi'e-ta

"I can see the white on the top of the mountain."

[atusa-set]

kare-wa kono atusa-ni tomadot'-ta

"He was embarrassed with this thickness."

kare-wa kono a'tusa-ni tomadot'-ta

"He was embarrassed with this hotness."

Appendix 4-2. Test sentences for Experiment 6

[atusa-set]

kono kami-no atusa-ga ka're-ni-wa odoroki-da't-ta

"He was surprised at this thickness of paper."

kono natu-no a'tusa-ga ka're-ni-wa odoroki-da't-ta

"He was surprised at this hotness of summer."

ka're-ga odoro'i-ta-no-wa kono kami-no atusa-da

"What he was surprised at was this thickness of paper."

ka're-ga odoro'i-ta-no-wa kono natu-no a'tusa-da

"What he was surprised at was this hotness of summer."

[hasi-set]

hosonaga'i nuno-no hasi-ga ka'nozyo-ni-wa totemo atukai-ni^{0}ku^{0}-kat-ta

"The narrow edge of the cloth she found very difficult to deal with."

hosonaga'i nuri-no ha'si-ga ka'nozyo-ni-wa totemo atukai-ni^{0}ku^{0}-kat-ta

"The long chopsticks she found very difficult to use."

hosonaga'i isi-no hasi'-ga ka'nozyo-ni-wa totemo watari-ni^{0}ku^{0}-kat-ta

"That long and narrow bridge she found very difficult to cross."

ka'nozyo-ni-to't-te atukai-ni^{0}ku^{0}i-no-wa hosonaga'i nuno-no hasi-da

"What she found difficult to deal with is that narrow edge of cloth."

ka'nozyoni-to't-te atukai-ni^{0}ku^{0}i-no-wa hosonaga'i nuri-no ha'si-da

"What she found difficult to use is those long chopsticks."

ka'nozyoni-to't-te watari-ni^{0}ku^{0}i-no-wa hosonaga'i isi-no hasi'-da

"What she found difficult to cross is that long and narrow bridge."

[siro-set]

yama-no tyoozyo'o-no siro-ga yuuhi-ni ha'e-te kagaya'i-te-iru

"The castle on the top of the mountain is glowing with the splendor of the setting sun."

yama-no tyoozyo'o-no si'ro-ga yuuhi-ni ha'e-te kagaya'i-te-iru

"The white on the top of the mountain is glowing with the splendor of the setting sun."

yuuhi-ni ha'e-te kagaya'i-te-iru-no-wa yama-no tyoozyo'o-no siro-da

"What is glowing with the splendor of the setting sun is the castle on the top of the mountain."

yuuhi-ni ha'e-te kagaya'i-te-iru-no-wa yama-no tyoozyo'o-no si'ro-da

"What is glowing with the splendor of the setting sun is the white on the top of the mountain."

[sake-set]

sono tugitate-no sake-ga totemo oisi-so'o-ni mi'e-ta

"The *sake* that has just been poured looked very delicious."

sono toretate-no sa'ke-ga totemo oisi-so'o-ni mi'e-ta

"The salmon that has just been caught looked very delicious."

ka'rega i'ma itiban hosi'i-no-wa sono tugitate-no sake-da

"What he wants most now is the *sake* that has just been poured."

ka're-ga i'ma itiban hosi'i-no-wa sono toretate-no sa'ke-da

"What he wants most now is the salmon that has just been caught."

[sake-set2]

me-no ma'e-ni okare-ta sake-ga totemo oisi-so'o-ni mi'e-ta

"The *sake* before his eyes looked very delicious."

me-no ma'e-ni okare-ta sa'ke-ga totemo oisi-so'o-ni mi'e-ta

"The salmon before his eyes looked very delicious."

ka're-ga i'ma itiban hosi'i-no-wa me-no ma'e-ni okare-ta sake-da

"What he wants most now is the *sake* before his eyes."

ka're-ga i'ma itiban hosi'i-no-wa me-no mae-ni okare-ta sa'ke-da

"What he wants most now is the salmon before his eyes."

[harete-set]

kao-ga hare-te ka're-wa totemo kurusi-so'o-dat-ta

"His face lightened and he looked very happy."

kao-ga ha're-te ka're-wa totemo uresi-so'o-dat-ta

"His face is swollen and he looked very painful."

[kamuto-set]

sono ko-wa hana-o kamu-to kanasi-so'o-ni kotira-o mi'-ta

"The boy gnawed his nails and looked sadly at me."

sono ko-wa tume-o ka'mu-to kanasi-so'o-ni kotira-o mi'-ta

"The boy blown his nose and looked sadly at me."

Appenix 4-3. Ratio of accent- and/or position-related lengthening

Target items	Position	Accent type	Vowel duration (ms)	Accent effect 1/2, 3/4	Position effect 1/3, 2/4	Effects of accent & position 1/4
1. atusaga	NonSF	a.	168.9	1.11	1.21	1.66
2. a'tusaga	NonSF	b.	152.3		1.50	
3. atusada	SF	a.	139.8	1.38		
4. a'tusada	SF	b.	101.6			
1. hasiga	NonSF	a'.	167.0	1.24 (1/3)	1.23 (1/4)	1.66 (1/5)
2. ha'siga	NonSF	b'.	147.5		1.47 (2/5)	1.93 (1/6)
3. hasi'ga	NonSF	b'	134.3		1.55 (3/6)	
4. hasida	SF	a'.	135.7	1.57 (4/6)		
5. ha'sida	SF	b'.	100.5			
6. hasi'da	SF	b'	86.5			
1. sakega	NonSF	a.	162.9	1.14	1.24	1.69
2. sa'kega	NonSF	b.	143.1		1.49	
3. sakeda	SF	a.	130.9	1.36		
4. sa'keda	SF	b.	96.3			
1. sakega2	NonSF	a'.	161.7	1.10	1.26	1.80
2. sa'kega2	NonSF	b'.	147.0		1.63	
3. sakeda2	SF	a'.	128.3	1.43		
4. sa'keda2	SF	b'.	90.0			
1. siroga	NonSF	a'.	169.4	1.06	1.23	1.69
2. si'roga	NonSF	b'.	159.8		1.60	
3. siroda	SF	a'.	137.5	1.37		
4. si'roda	SF	b'.	100.1			
1. kamuto	NonSF	a.	162.1	1.13		
2. ka'muto	NonSF	b.	143.4			
1. harete	NonSF	a.	166.1	1.11		
2. ha'rete	NonSF	b.	149.3			

Appendix 6-1. Test sentences and phrases for Experiment 7

(1) Basic test sentences with one-mora nouns as subjects or objects

ne' ga nuketa	ne' nuketa	"The root is/was pulled out."
ne' ga no'bita	ne' no'bita	"The root has grown."
na' ga nieta	na' nieta	"The vegetables are/were boiled."
ne' o nuita	ne' nuita	"I have pulled out the root."

(2) Negative sentences with focus on the predicate

ne' ga nuke na'katta	ne' nuke na'katta	"The root was not pulled out."
na' ga nie na'katta	na' nie na'katta	"The vegetables were not boiled."

(3) Test sentences with two-mora nouns as subjects

ya'ne ga nuketa	ya'ne nuketa	"The roof was blown away."
i'ne ga no'bita	i'ne no'bita	"The rice plants have grown."