Lexical Effects in Japanese Vowel Reduction

Setsuko Shirai

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2005

Program Authorized to Offer Degree: Linguistics

UMI Number: 3163406

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Setsuko Shirai

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

Richard A. Wright

Reading Committee:

Richard A. Wright

Sharon L. Hargus

Alicia B. Wassink

Date: 17 December 2004

University of Washington

Abstract

Lexical Effects in Japanese Vowel Reduction

Setsuko Shirai

Chair of the Supervisory Committee:

Associate Professor Richard Wright

Department of Linguistics

This dissertation reports the results of a study of a vowel reduction in Japanese. Vowel reduction (durational and spectral) has been observed in many languages; however, only a few studies have been conducted research on Japanese phonetic vowel reduction. Furthermore, in these studies, factors influencing duration and formants were not controlled. I investigated vowel reductions in Japanese, while controlling factors influencing duration and formants. In my study, vowels in the functional particles /ga/, /de/, and /to/ were compared with word-final vowels in lexical words.

The overall results show that there was a significant effect for LEXICAL on DURATION and F1 values but not on F2 values. When I looked at the lexical effect for individual vowels, vowels [a] and [o] showed lexical effect on DURATION but [e] did not. I sought the reasons for no lexical effect on DURATION of [e] and concluded that lexical effects on [e] were not observed because of phrase final lengthening.

For formants, only [a] showed a significant lexical effect on F1 and F2 values but [e] and [o] did not. The lack of observation of a lexical effect was probably due to the short locus distance between alveolar consonants, [d] and [t], and vowels, [e] and [o],

respectively. Furthermore, I investigated the possible coarticulation effects of consonants on vowels and V-to-V coarticulation effects.

Durational and spectral vowel reductions for [a] were observed, which raised a question: whether duration is the only factor that causes spectral vowel reduction. In the undershoot hypothesis, the short duration of reduced vowels leads to the spectral vowel reduction; in other words, duration is the only factor that causes spectral vowel reduction. I investigated whether or not there was a lexical effect in addition to a durational effect. The results indicate that there was a significant effect for LEXICAL on F1 values. Function particles have a low information value; consequently speakers do not articulate these function particles clearly. In my hypothesis, there are two targets (content and function) instead of one target.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

This thesis reports the results of a study of vowel reduction in Japanese. There are two kinds of vowel reduction: phonetic and phonological. An example of phonological vowel reduction from English is the first vowel in "major" [mejdʒər], which becomes a schwa in "majority" [mədʒɔ́rəɾi]. In phonological vowel reduction, there is a categorical change from one vowel quality, [ej] in this case, to a neutral vowel, [ə]. Thus there is a neutralization of vowel contrast. On the other hand, phonetic vowel reduction is characterized by continuous and incremental changes in vowel quality and duration. The phonetic changes range from subtle effects that are barely distinguishable to dramatic changes that may neutralize contrasts. For example, in colloquial speech, the article "a" in "I bought a book" is pronounced [ə] in a declarative sentence, while it is pronounced [ej] in a focus position as in the question "Is it a book or many books". In my dissertation, I focus on phonetic vowel reduction[1] in Japanese.

Phonetic vowel reduction has been studied in many languages such as English, Swedish and Dutch, which have a stress contrast. The phonetic characteristics of unstressed vowels are decreased amplitude, shortened duration, and shallow pitch excursion. In addition, unstressed vowels are located close to the center of the vowel space. Vowel reduction is typically observed in unstressed vowels, and has two characteristics: durational and spectral reduction. Japanese is a pitch accent language and does not have a stress contrast so researchers may not expect to observe vowel reduction in Japanese. Furthermore, the Japanese language is mora-timed, which means that each mora is

[1] The Japanese language does not have known phonological vowel reduction except perhaps devoiced vowels.

supposed to have a fixed duration. This is probably the reason that there are few studies about Japanese phonetic vowel reduction. However, Ueyama (1997) observed that function /a/ was shorter than content /a/ so it is possible that phonetic durational reduction occurs in Japanese. In addition, the research conducted by Wright (2003) observed degrees of lexically based phonetic vowel reductions in English stressed syllables. While there are many factors that contribute to vowel reduction, one of the most frequently cited factors is the lexical status of the word in which a phone occurs. Lexical factors, such as word frequency and word class (content, function) may increase or decrease the predictability of a word. As a word's predictability rises, speakers tend to allow a greater degree of phonetic reduction in speech (ex Lindblom 1990, Hunnicut 1985, and Liberman 1963). Word frequency and lexical word class (content, function) have been shown to affect the predictability of a word thereby affecting the degree of phonetic vowel reduction (ex: Balota 1989, Bybee 1994 and Zipf 1935). Therefore, it is possible that spectral vowel reduction occurs in Japanese.

In this dissertation, I examine the degree to which word class (content and function) affects phonetic durational reduction and spectral reduction in Japanese. I then investigate the degree to which the observed spectral reduction can be attributed to target reduction rather than vowel undershoot (Lindblom 1963) (the tendency for vowels to fail to achieve their articulatory targets in shorter syllables). Thus, I have three research questions about Japanese: whether or not durational reduction occurs, whether or not spectral reduction occurs, and whether or not observed spectral vowel reduction is solely the result of undershoot.

2 Vowel reduction I: Duration

One way in which vowel reduction is manifested is in overall duration: the shorter the vowel relative to its length in other full stressed positions the greater the degree of reduction. For example, in English, durations of stressed syllables are longer than those of unstressed syllables (ex: Fourakis 1991). Thus, the degree of reduction can be measured by comparing the duration of reduced vowels with the duration of full vowels. As was noted in the introduction, word class has been shown to affect vowel duration in languages such as Dutch (ex: Van Bergem 1993). Therefore, I expect to see similar lexical effects in Japanese. That is, I expect that vowels from function particles will have shorter durations than equivalent vowels in content words.

2.1 Background

Segmental duration is influenced by several factors. Known factors influencing segmental duration are the position of a phoneme in a word (ex: English, Beckman & Edwards 1990), the presence or absence of stress (ex: Dutch, Sluijter and van Heuven 1995), adjacent phonemes (ex: English, Port 1981, Japanese, Han 1962), the number of syllables in a word (ex: English, Lehiste 1972, Port 1981), the rate of speech (ex: English, Port 1981), and syntactic structure (ex: English, Klatt 1976).

Lexical status also influences the duration of a phoneme, which is the topic of this study. In English, function words are reduced (ex: Bolinger and Sears 1981). However, there has been scant research cross-linguistically on lexically motivated vowel reduction;

therefore, it is not yet established whether function vowels in Japanese are shorter than content vowels or not. The existing research is summarized in the following sections.

2.1.1 Campbell (1992)

A study by Campbell (1992) indicated that Japanese particles, (i.e. function words) were longer on the average than their content word counterparts. In the study, a single male Japanese speaker read aloud sentences taken from newspapers and magazines. Campbell's aim was to determine whether each segment would have a different pattern for lengthening. He also compared the duration of the topic marker (the function word), [wa], with the duration of [wa] in content words. The result indicated that lexical status did not influence the duration of the phones. However, Campbell did not control other factors affecting the duration of content vowels such as the number of moras, the level of phrase boundary, the preceding consonants, etc. Since other factors affecting duration were not controlled, it is possible that the position of the target segment and the number of moras or other factors influenced Campbell's duration results. In addition, the number of observations was not large (content [wa] 82 and function [wa] 139). Therefore, the lack of observed effect might be due to low statistical power.

In addition, Campbell offered an argument for the short duration of function [a] (Kaiki *et al* 1990) in the sentence final position. His explanation of this phenomenon was that the preceding voiceless consonant [t], which predominated in the corpus, led to the short durations of vowel [a]. He also mentioned that "the past-tense marker may be informationally redundant (p415)" which might cause durational reduction. Thus while his findings indicate no lexical effect, or even an effect that was the reverse of previous findings in other languages, the lack of control makes his findings questionable.

2.1.2 Ueyama (1997)

Ueyama (1997) conducted research on vowel duration in Japanese, in which two questions were asked: whether or not lexical status (content vs. function) would lead to a difference in vowel duration and whether or not differences in prosodic boundary levels would lead to duration effects of phrase final vowels. Her results show that function [a] in [obaba-ga] 'grandma_SUB[1]' was shorter than content [a] in [obaba] 'grandma' and that there were differences of vowel durations between the end of major phrases (INTONATIONAL PHRASE and SENTENCE) and the end of lower phrase boundaries (ACCENTUAL PHRASE and INTERMEDIATE PHRASE). She measured the duration in of the vowel in sentence final position, using a sentence which ended with the noun phrases, [obaba] and [obaba-ga]. The sentence final vowel was not significantly shorter than the vowel at the end of the INTONATIONAL PHRASE. Thus, the short duration of sentence final [a] in Kaiki's study might be due to the frequent occurrence of [ta], rather than to its final position. Both studies indicate that part of speech affects the duration of segments. However, because the numbers of syllables and adjacent consonants were not controlled in Ueyama's study, there remains the possibility that the different numbers of moras in a word (3 vs. 4) led to the shorter duration of function vowels.

[1] SUB: the subject marker

2.2 Experiment 1: Duration

There is an apparent conflict in the results of the two experiments. One appears to indicate that vowels in function words (hereafter function vowels) are longer than vowels in content words (hereafter content vowels), while the other seems to indicate that vowels in function words (hereafter function vowels) are shorter. To resolve this discrepancy, I tested the effect of lexical status (content, function) on vowel duration in Japanese controlling for other factors which can possibly influence duration such as the number of moras in the word.

2.2.1 Purpose

I designed the current experiment to investigate the difference between vowels in content words and vowels in function particles in Japanese.

2.2.2 Japanese function particles

A typical Japanese function word is a particle such as the subject marker [ga]. Although there are other function words such as [made] 'until,' I focused on particles because they are more plentiful and easier to control. For example, the function word, [made], consists of two moras so it is hard to find a similarly shaped content word. In addition, it is not clear whether or not both moras are involved in reduction. The function particles are compared with an equivalent syllable from content words in this experiment.

These particles are not affixes, as there are notable distinctions between particles and affixes. Particles do not undergo phonological processes that affixes undergo such as vowel deletion. I will use Japanese diminutives as an example. To make a diminutive, use the first two moras of the first name, and add [tʃan]. The diminutive of [satʃiko] would be [satʃitʃan], but it is [sattʃan] due to vowel deletion (the deleted vowel is from the root, not the suffix in this example). On the other hand, this phonological process does not occur with particles. In addition, sentences without particles are grammatical[2] while sentences without suffixes are not. Examples of particles are shown in sentences (2-1) and (2-2). Sentence (2-1) is often heard in conversation, while sentence (2-2) is written. Both sentences are grammatical.

(2-1) hon		**katta?**	**"Did (you) buy a book?"**
book		**bought**	
(2-2) hon	o	**katta?**	**"Did (you) buy a book?"**
book	OBJ[3]	**bought**	

The examples of affixes are sentences (2-3) and (2-4). The word [hajai][4] is an adjective, while [hajakɯ] is an adverb. So [i] is an adjectival suffix while [kɯ] is an adverbial suffix. The sentence in (2-3) is ungrammatical while the sentence in (2-4) is grammatical.

(2-3)	*** haja**	**arɯkɯ**	**"(I) walk fast."**
	fast	**walk**	
(2-4)	**hajakɯ**	**arɯkɯ**	**"(I) walk fast."**
	fast	**walk**	

---

[2] The subject marker, [ga], the object marker [o], and topic marker [wa] can be dropped while other particles are not dropped (Martin 1988). The postposition [to] 'with' can be dropped in idiomatic phrases.

[3] OBJ: the object marker

[4] There is no word without a suffix. The corresponding noun is [hajasa] so [sa] is a nominal suffix. The word, [haja] is not related to these words.

Therefore, affixes cannot be deleted, while particles can be. Both facts indicate that Japanese particles should not be analyzed as suffixes. I will refer to these particles as function words.

2.2.3 Hypothesis 1

Hypothesis 1: In English, function words such as the definite article, *the*, or the preposition, *in*, are reduced in normal speech. Likewise, I predict that the duration of the vowels in Japanese function particles will be shorter than the duration of the matched word-final vowels in content words.

2.2.4 Methodology

Researchers generally use two methods to control factors influencing phone duration. The first is to use a carrier phrase such as "Say __ again" and measure only the target words. The second is to measure all phones in a corpus. However, these methods are problematic when observing the reduction of function particles. The first method may lead speakers to place focus on the target word, consequently rendering the reduction of the target word unobservable. The second method is characterized by lack of control over other factors which could influence the duration of vowels. We cannot observe the effect of the factor of interest unless the corpus is large enough to ensure that the data are not biased or to ensure that there are not uncontrolled factors that may have a larger effect than the lexical factor of interest to this study.

Van Bergem (1993) used a different method, in which a set of sentences contained the target syllables whose lexical status and other factors were varied. My method was similar to Van Bergem's; and I used pairs of sentences that contain target syllables of two lexical types: content and function. For example, one function sentence had the function particle [-ga] preceded by the word [dʒinbɯtsɯ] 'person' while the counter-part sentence had the matched word, [dʒinbɯtsɯga] 'portrait'. In this way, I control for prosodic factors such as the number of syllable, and position in the word.

Speaking rate is also an influencing factor in phone duration. Two factors at work in increasing speaking rate are sentence length and phrase position. If the sentence is long, speakers tend to speak fast; consequently the length of each phone is shortened. As for phrase position, speakers tend to speak slowly at the beginning of the sentence and then speed up in the middle, only slowing on the final syllable. Therefore, one must control for sentence length and phrase position in the experimental design. To control for these factors, the function and content token sentences were almost identical in length and prosodic structure.

If a pair of sentences is read together, subjects may notice the aim of the experiment, and alter their speed, hyperarticulate or vary their pronunciation in a variety of ways. In addition, the speaking rate may change from the beginning of the list to the end of the list. To avoid the influence of sentence order, the sentences were pseudo-randomly[5] presented.

In English, presence or absence of stress affects duration of phones. However, previous research revealed that Japanese pitch accent does not influence the duration of

[5] I used the randomizing function in Excel to determine the order of sentences. When paired sentences were next to each other, I changed the order so that subjects would not change their speaking rates.

syllables (Sugito and Mitsuya 1977). In addition, Cutler and Otake (1999) briefly state that the pitch of the target syllable does not affect the duration of the syllable in Japanese. Therefore, I decided it was not necessary to control for pitch accent in this study.

2.2.4.1 Tokens

I created two pairs of tokens for three target syllables [ga], [de] and [to], so I had 12 sentences. These sentences contain the same phoneme sequences with the target syllable located word-finally for both content and function word tokens. Each of the token pairs is as follows[6]:

(2-5) Tokens

[ga]: two-syllable word

a. Content:
bokushi, **ki<u>ga</u>** kikitsukeru
pastor starvation heard
A pastor heard about the starvation

b. Function:
katoo_san_wa **ki_<u>ga</u>** kiku.
Mr. Kato__TOP[7] spirit_SUB be effective
Mr. Kato is tactful.

[ga]: four-syllable word

c. Content:
nakamura_san_wa **jinbutsu<u>ga</u>** toku_ni osukide irasshaimashita
Mr. Nakamura_TOP portraits especially like (honorific and past form)
Mr. Nakamura especially liked the portraits.

---

[6] I used Romanization to transcribe tokens for this manuscript, but presented the tokens to readers in Japanese.

[7] TOP: the topic marker, SUB: the subject marker, CONJ: conjunction, NEG: the negative marker, GEN: the genitive marker, DAT; the dative marker, LOC: the locative particle

d. Function:
ano eiga_wa **jinbutsu_ga** toku_ni yoku byousha sareteimasu
that movie_TOP people_SUB especially well describe
In that movie, people are described especially well.

[to]: two-syllable word

e. Content:
hato kubi_o mawashita
pigeon neck_OBJ turned
The pigeon turned his head.

f. Function:
ha_to kuki dake shika nai
leaf_CONJ stem only NEG
There is only a leaf and a stem

[to]: three-syllable word

g. Content:
kookoosei **kogoto** karuku kikinagasu
high school students sermon lightly listen without attention
High school students do not listen to the sermon seriously

h. Function:
kookoo_no **kogo_to** kanbun_wa muzukashii
high school_GEN old Japanese_CONJ classic Chinese_TOP difficult
Studying old Japanese and Classic Chinese in high school is difficult

[de]: three-syllable word

i. Content:
kotowaza_ni, **tade** kuu mushi mo suki dzuki to
proverb_DAT glass eat bug_OBJ (also) like_COMP
A proverb says that every one has his own taste.

j. Function:
taue_no ato **ta_de** kuu meshi_wa umakatta
rice planting_GEN after paddy_LOC eat meal_TOP delicious
After rice planting, meals eaten at the rice paddy were delicious.

[de]: three-syllable word[8]

k. Content:
suizoku_kan_ni iku_made, **hitode**, kon'nani kireidanante, shiranakatta
aquarium DAT go_until starfish like this beautiful did'nt know
Until coming to the aquarium, I did not know that starfish were so beautiful.

l. Function:
kato_san_wa, Tokyo no **hito_de**, korokoro_to
Mr. Kato_TOP Tokyo_GEN person_COP pleasently
yoku waratte, aisoo_ga ii.
well laugh friendly
Mr. Kato of Tokyo frequently laughs and is friendly.

2.2.4.2 Subjects

Six male and six female native Japanese speakers participated in the experiment. One speaker was born in Yokohama, and the others were born in Tokyo. Yokohama is close to Tokyo, and the Yokohama dialect is considered to be the same as the Tokyo dialect. Thus, all subjects are considered to be Tokyo dialect speakers. Most speakers were in their 20's, except for one female speaker in her 40's, and three speakers in their 30's. Another female speaker participated in this study; however, she produced nasalized /a/ so I did not use her speech in this study. When I recorded the first subject, "kn", I mistakenly presented the word /hitode/ with the subject marker /ga/ (i.e. /hitode_ga/) so her data for /hitode/ are missing. Since I recorded /tade/, I used her data except for token effect. Each subject read randomized lists of approximately 100 sentences, 12 of which were of interest in this study.

---

[8] Because of palatalization, Japanese [hi] is [çi]. For the convenience I use [hi]. Japanese /u/ is unrounded [ɯ].

2.2.4.3 Recordings

Recordings took place in the sound attenuated recording booth in the phonetics lab at the University of Washington. Recordings were made using an Electro-Voice RE20 microphone with a flat response to 20 kHz, an amplifier (Shure model FP32A) and a professional analog cassette tape recorder (TASCAM 122MK III). The speech was digitized at 11 kHz, 16 bit, using Sound Edit 16 version 2 on a computer equipped with an Audiomedia III card.

2.2.4.4 Measurements

I measured the vowel duration from onset to offset. In measuring vowel duration, the following criteria were used. For vowel onset, I measured from the second clear vowel cycle where formant structure was clearly visible. For vowel offset, I measured to the last peak where formants were clearly visible, excluding creaky voicing. The waveform and spectrogram in Figure 2.1 illustrate the measuring points in a typical case.

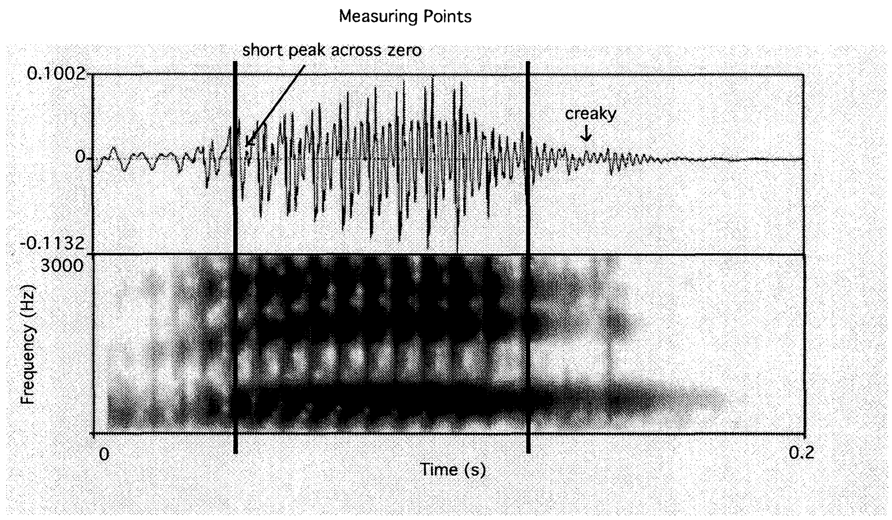

Figure 2.1 Wave and spectrograph (horizontal axis: TIME, vertical axis: AMPLITUDE for waveform, FORMANTS for spectrogram) illustrating measuring points. The end of the speech is creaky, and prolonged. I chose the point before the creaky voice.

Figure 2.1 shows that the end of the vowel is creaky and prolonged. Since creaky voicing has unpredictable effects on duration, the creaky part has been excluded from duration measures. In addition, the subjects produced fricative [ɣ] and nasal [ŋ] as the allophones of voiced velar stop [g]. In order to measure the duration of a vowel after a stop in the same way as the duration after a fricative, I needed criteria for measuring them consistently. Figure 2.1 provides an example of these criteria. At the beginning of the waveform in Figure 2.1, cycles consist of two peaks, but later a cycle consists of three peaks. I chose as the starting point of a vowel where the middle peak or short peak crosses zero. So the short peak crossing zero is the criterion for this measurement.

2.3 Results of Experiment 1: Duration

Overall results indicate that function vowels were reduced. That is, the average durations of function vowels were shorter than their content counterparts. In analyzing individual vowels, the averages confirmed this tendency. However, not all data showed significant differences.

2.3.1 Overall results

The overall results indicate that vowels in function words (i.e. function vowels) were shorter than content vowels. The averages and standard deviations are summarized in Table 2.1. Averages and standard deviations in the tables are calculated based on all measurements.

Table 2.1 Means and standard deviations of DURATION (ms)

	all	content	function
DURATION (ms)	77	81	72
(SD)	(29.0)	(30.4)	(26.8)

Table 2.1 shows that content vowels were longer function vowels, and the standard deviation of content vowels was slightly greater than that of function vowels. A boxplot (horizontal axis: LEXICAL, vertical axis: DURATION) in Figure 2.2 illustrates the median values and ranges of durations of the content and function vowels. The boxes in the figures were drawn with horizontal lines indicating the 25th, 50th (median), and 75th percentiles. The whiskers above and below the boxes extend to one-and-a-half times the upper and lower interquartile ranges, respectively.

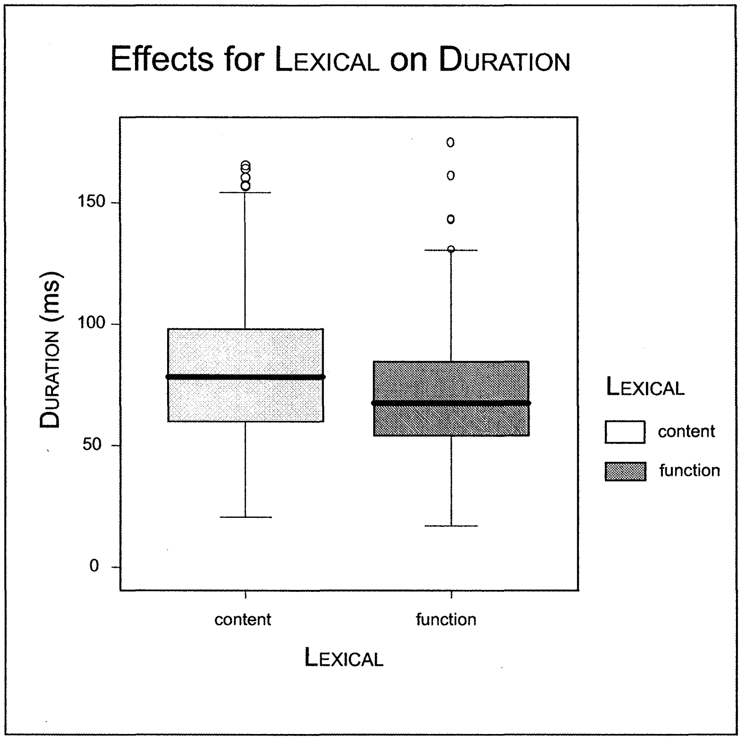

Figure 2.2 Box plot illustrating that the durations of content vowels and function vowels. The middle lines in the boxes indicate the median (the 50th percentile).

The boxplot in Figure 2.2 clearly shows that the durations of function vowels were shorter than those of content vowels. Next I will use a repeated measures ANOVA. Since there is a strong correlation among vowels produced by each speaker, the average durations of content and function for each speaker and each vowel are calculated individually and used to describe the relationships among factors. However, averages may be affected by outliers, so I also used all measurements to see which vowels show effects for LEXICAL. An alpha level of 0.05 was used for all statistical tests. Results were submitted to a repeated-measures ANOVA with AVERAGE DURATION as the dependent variable and LEXICAL (content, function) as the independent variable. There was a significant effect for LEXICAL on DURATION, $F(1, 10) = 5.145, p < 0.047$. However, this overall result might have been due to the disproportional effect in one vowel overwhelming the marginal effects in the other vowels. Therefore, I next considered individual vowel effects.

2.3.2 Individual vowels

The measurements are summarized in Table 2.2.

Table 2.2 Means and standard deviations of DURATION (ms) split by VOWEL

VOWEL		all	content	function
a	DURATION	81	92	70
	(SD)	(27.3)	(28.7)	(20.7)
e	DURATION	84	81	87
	(SD)	(27.0)	(24.9)	(28.8)
o	DURATION	66	71	60
	(SD)	(29.5)	(33.4)	(24.0)

To investigate individual vowel effects, a two-way repeated-measures ANOVA was performed with AVERAGE DURATION as dependent variable and VOWEL ([a], [e] and [o]) and LEXICAL (content and function) as independent variables. The results indicate that there was a significant effect for VOWEL and for the VOWEL by LEXICAL interaction, and a near-significant effect for LEXICAL. These results were based on averages so the magnitude of lexical effect appears to be weak. The results are summarized in Table 2.3.

Table 2.3 Effects for VOWEL and LEXICAL on AVERAGE DURATION

	VOWEL	LEXICAL	VOWEL x LEXICAL
df	[2, 20]	[1, 10]	[2, 20]
F	12.27	4.723	7.933
p	< 0.001	0.055	0.003

Next I used all measurements for duration to test which vowels show lexical difference. Since I focus on lexical difference for individual vowels and it is possible that differences by subject (i.e. within-subject differences) may exceed differences among subjects (i.e. between subject differences) to a great extent, I used all measurements. The durations split by VOWEL show that, with the exception of [e], function vowels were shorter than content vowels.

Table 2.4 Effects for LEXICAL on DURATION split by VOWELS

	a	e	o
df	[1, 65]	[1, 62]	[1, 65]
F	26.33	1.849	5.377
p	<0.001	0.179	0.024

There were statistically significant lexical effects for [a] and [o] but not for [e]. I also used Wilcoxon Signed-Rank test for paired data for the lexical effect. The Wilcoxon Signed-Rank test is distribution free; in other words, even if data are not normally distributed, we can use this test (Kirk 1999). When I looked at individual subjects, the results of most subjects favored the hypothesis but a couple of subjects showed the opposite direction. I used this test to see whether or not these subjects would be problematic. In this test, the data are sorted based on the absolute differences of paired data, and assigned ranks (the pair with smallest difference is rank 1). Then the sums of ranks for positive and for negative differences are separately calculated. If there is no difference between two levels, the sums of ranks for positives should be about the same as the sum for negatives. I used all measurements of durations for this test. The results are summarized in Table 2.5. There were no ties in the results so the table does not have a column for ties.

Table 2.5 Effects for LEXICAL on DURATION (Wilcoxon signed-rank tests)

VOWEL	*Z*	*p*	Σ negative ranks	Σ positive ranks
all	-3.632	< 0.001	6689[9]	12421
a	-4.264	< 0.001	1773	438
e	-0.815	0.415	889	1127
o	-2.578	0.010	1509	702

The results of Wilcoxon signed rank tests are similar to the results of the repeated-measures ANOVA. There is a significant effect for LEXICAL on DURATION for /a/ and /o/ but not for /e/. With this test, I was able to compare the frequency of cases where function vowels were shorter than content vowels with that of the other cases. The results of the Wilcoxon tests indicate that function vowels were consistently shorter than content vowels.

All measurements are plotted in a boxplot (horizontal axis: VOWEL and LEXICAL, vertical axis: DURATION) in Figure 2.3.

[9] This number is not the sum of ranks (a, e and o) because the rank starts from 1 for each vowel.

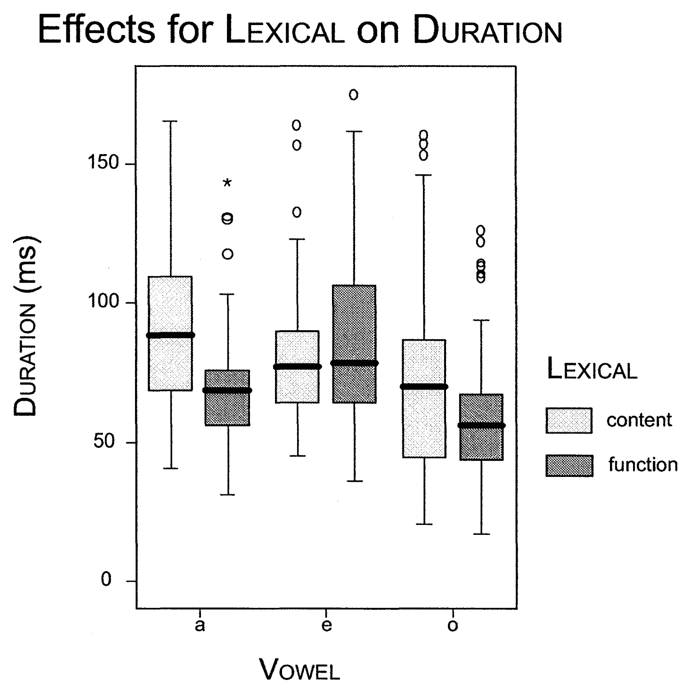

Figure 2.3 Boxplot illustrating the lexical difference of vowel duration. Horizontal axis: vowels, Vertical axis: duration (ms).

The boxplot in Figure 2.3 reveals that function [a] and [o] were clearly shorter than content ones. However, there was very little difference in the median values for [e], and the variance for function [e] was greater than for content [e].

Below in Figures 2.4 – 2.8, I have provided pairs of waveforms (horizontal axis: TIME, vertical axis: AMPLITUDE in all waveforms) and pairs of spectrograms (horizontal axis: TIME, vertical axis FORMANTS in all spectrograms) produced by the subject, "ss". These waveforms and spectrograms illustrate the differences between content vowels and function vowels. The length of each window is 300 ms. In extreme cases, function [a] and [o] consist of only three glottal cycles, but the waveforms shown here are more representative of the majority of cases.

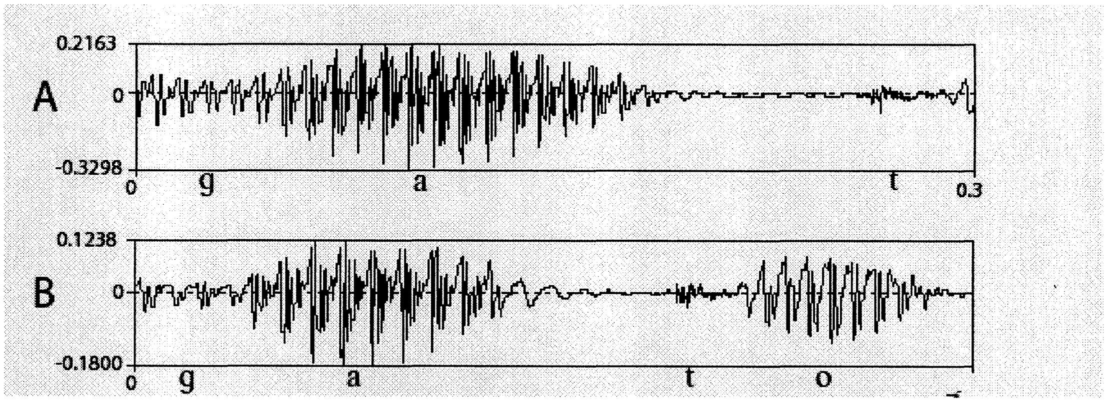

Figure 2.4 Waveforms illustrating content [a] (A) and function [a] (B) in [dʒinbɯtsɯga] produced by the subject, "ss."

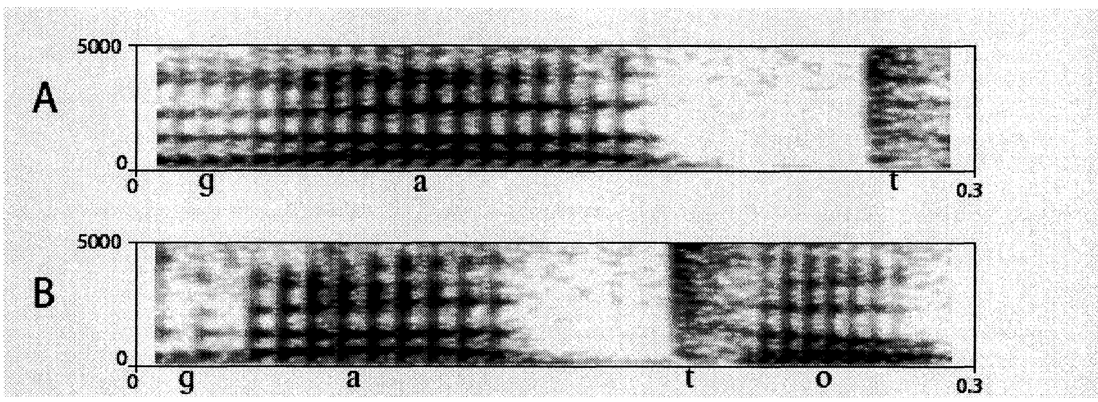

Figure 2.5 Spectrograms illustrating content [a] (A) and function [a] (B) in [dʒinbɯtsɯga] produced by "ss."

The waveforms in Figure 2.4 and spectrograms in Figure 2.5 illustrate the noticeable difference between content [a] (A) in [dʒinbɯtsɯga] and the equivalent function [a] (B).

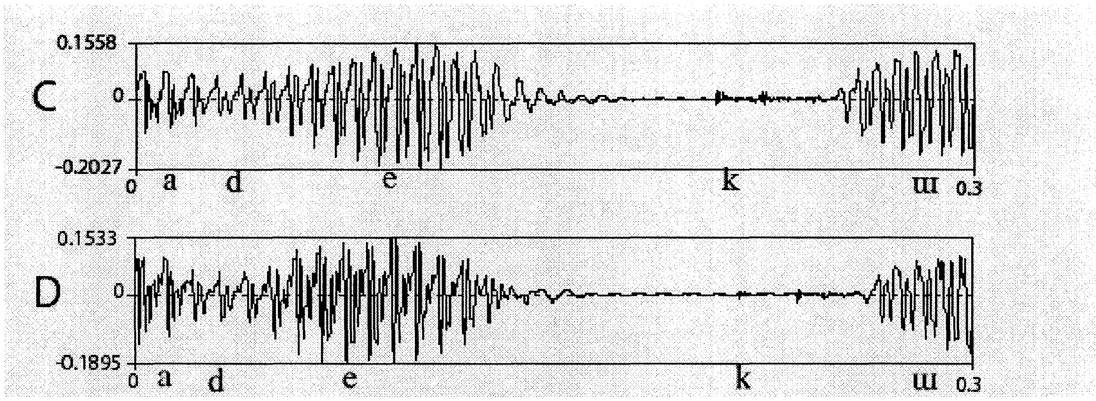

Figure 2.6 Waveforms illustrating content [e] (C) and function [e] (D) in [tad<u>e</u>] produced by the subject, "ss."

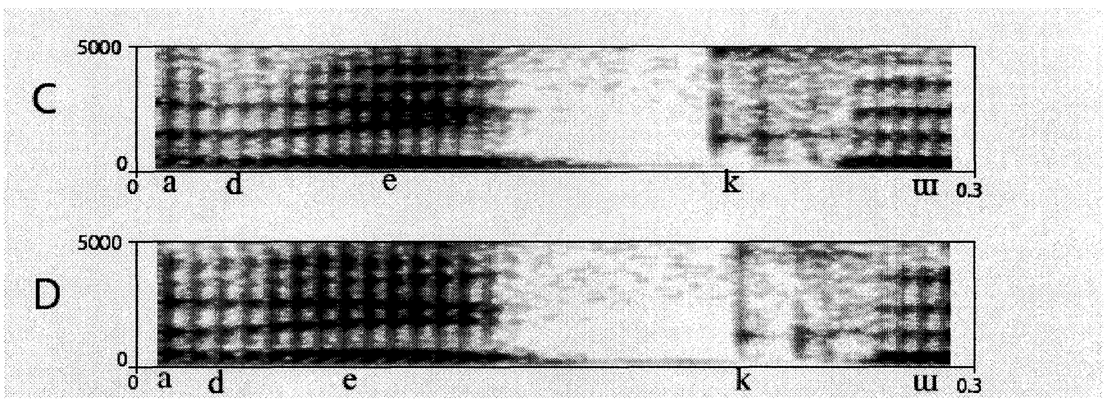

Figure 2.7 Spectrograms illustrating content [e] (C) and function [e] (D) in [tad<u>e</u>] produced by the subject, "ss."

The waveforms in Figure 2.6 and spectrograms in Figure 2.7 illustrate that content [e] (C) is slightly longer than function [e] (D) in [tade]. Waveforms in Figure 2.8 and spectrograms in Figure 2.9 illustrate [o] in [kogoto].

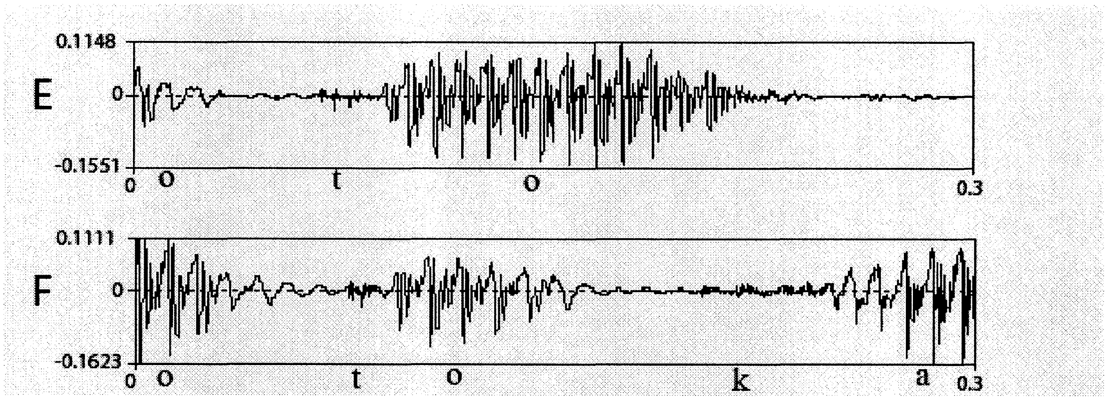

Figure 2.8 Waveforms illustrating content [o] (E) and function [o] (F) in [kogoto] produced by the subject, "ss."

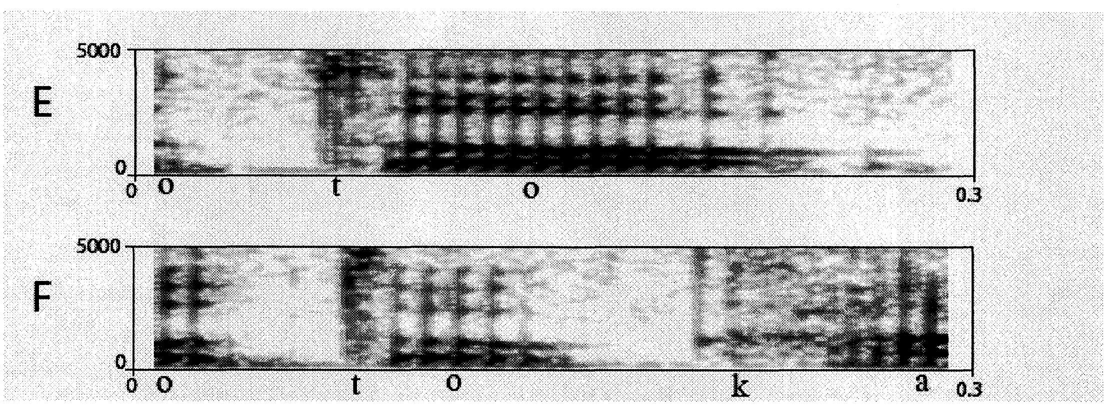

Figure 2.9 Spectrograms illustrating content [o] (E) and function [o] (F) in [kogot<u>o</u>] produced by the subject, "ss."

The waveforms in Figure 2.8 and spectrograms in Figure 2.9 illustrate a noticeable difference between content [o] (E) in [kogoto] and function [o] (F).

In summary, these pairs of waveforms and spectrograms illustrate the differences between function and content vowels. Overall function vowels tended to be shorter than content vowels, except for [e]. I also investigated possible gender differences; however, there was no significant effect for GENDER.

2.3.3 Token differences

To see whether or not the effect of LEXICAL status was uniform, individual words were examined. The duration measures are summarized in Table 2.6.

Table 2.6 Mean DURATION (ms) split by TOKEN

VOWEL	TOKEN		content	function
a	kiga	DURATION (SD)	85 (26.9)	71 (17.8)
	jinbutsuga	DURATION (SD)	100 (28.8)	69 (23.5)
e	hitode	DURATION (SD)	88 (30.2)	102 (31.9)
	tade	DURATION (SD)	74 (16.8)	73 (16.4)
o	hato	DURATION (SD)	68 (28.0)	64 (24.0)
	kogoto	DURATION (SD)	74 (38.3)	57 (23.8)

The table of the mean DURATIONS shows that, with the exception of the token in [hitode], the function vowels were shorter than those of content vowels in all tokens. The durations of [dʒinbɯtsɯga] show clearly that the duration of function [a] in [dʒinbɯtsɯga] was 65 % of the duration of content [a]. In addition, the duration of function [o] in [kogoto] was 70 % of the duration of content [o]. The duration measures are plotted in Figure 2.10, a boxplot showing TOKEN and LEXICAL (horizontal axis) by DURATION (vertical axis:).

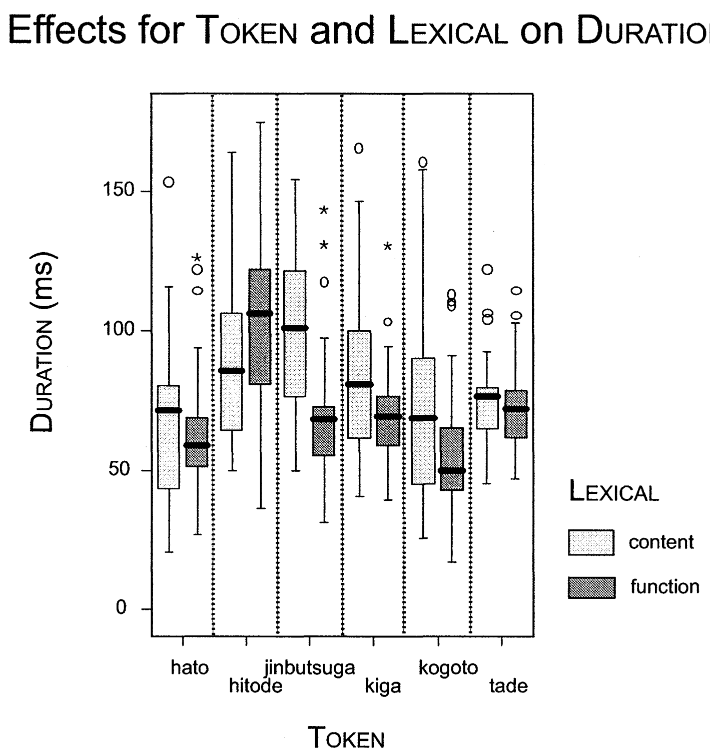

Figure 2.10 Boxplot illustrating the difference among tokens.

The boxplot in Figure 2.10 shows that with the exception of [e] in [hitode], content vowels were longer than function vowels, and content vowels in [dʒinbɯtsɯga] and [kogoto] were noticeably longer than function vowels.

To test for effects for TOKEN, the duration measurements are submitted to a two-way repeated-measures ANOVA. Since no token appears in two vowel categories, I separately ran factorial ANOVA for individual vowels. The results of a repeated measures ANOVA with AVERAGE DURATION as the dependent variable, and LEXICAL and TOKEN as independent variables are summarized in Table 2.7. For the vowel /e/, the degree of the freedom is [1, 9] because I did not record the vowel /e/ in content /hitode/[10].

[10] Since I presented a different word, "hitode_ga" to one subject, the data is missing.

Table 2.7 Effects for TOKEN and LEXICAL on AVERAGE DURATION split by VOWELS

VOWEL		TOKEN	LEXICAL	LEXICAL x TOKEN
a	*df*	[1, 10]	[1, 10]	[1, 10]
	F	1.334	13.794	2.704
	p	0.275	0.004	0.131
e	*df*	[1, 9]	[1, 9]	[1, 9]
	F	9.392	0.600	2.045
	p	0.013	0.459	0.187
o	*df*	[1, 10]	[1, 10]	[1, 10]
	F	0.022	3.537	2.978
	p	0.885	0.089	0.115

The results of the two-way repeated-measures ANOVA indicate that for [a] and [o], there was no significant effect for TOKEN on AVERAGE DURATION, but there was a significant effect for LEXICAL. In addition, there was a near significant interaction for [o]. Conversely, for [e], there was a significant effect for TOKEN on AVERAGE DURATION but no effect for LEXICAL nor for the interaction of the two. Even though the [e] in [hitode] showed the opposite tendency to that of [tade], there was no significant interaction.

Next I will show error bar graphs (horizontal axis: LEXICAL, vertical axis: DURATION) in Figures 2.13 – 2.15. These error bar graphs show the interaction between LEXICAL and TOKEN. Although Figure 2.10 shows the differences between tokens, it is hard to see how differently function vowels were reduced between tokens. The error bar graphs show the different reduction between tokens. In addition, these error bar graphs show the interaction between TOKEN and LEXICAL. The first error bar graph illustrates the interaction between TOKEN and LEXICAL for [a] in Figure 2.11 (horizontal axis: LEXICAL, vertical axis: DURATION).

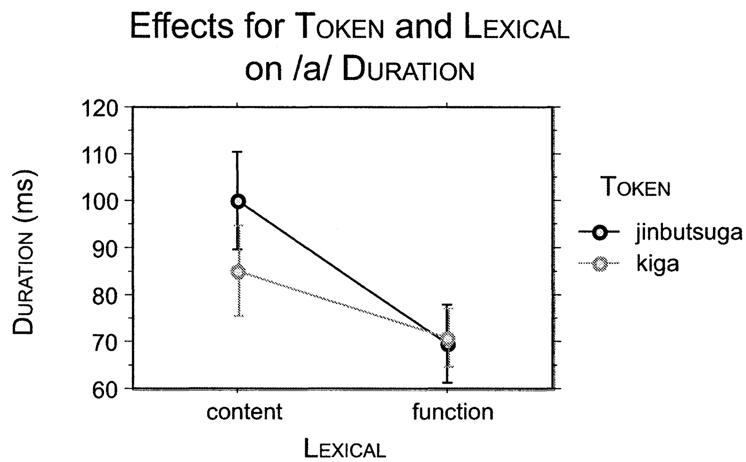

Figure 2.11 Error bars illustrating the interaction between LEXICAL by TOKEN for [a]

Figure 2.11 shows that the different reduction rate between [dʒinbutsɯga] and [kiga] led to the nearly significant interaction between of the two. The next error bar graph illustrates the interaction between LEXICAL and TOKEN for [e].

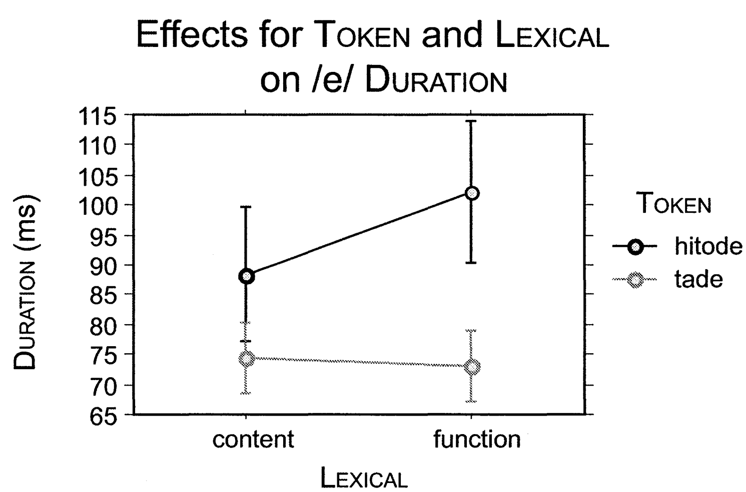

Figure 2.12 Error bars illustrating the interaction between LEXICAL by TOKEN for [e]. The error bars indicate the 95% Confidence interval.

In Figure 2.12, durations of both content and function for [hitode] were longer than durations for [tade]. Furthermore, content [e] in [tade] was slightly longer than function [e] but [e] in [hitode] showed the opposite tendency. In the case of [o], there was clear interaction between [hato] and [kogoto], which is illustrated in Figure 2.13.

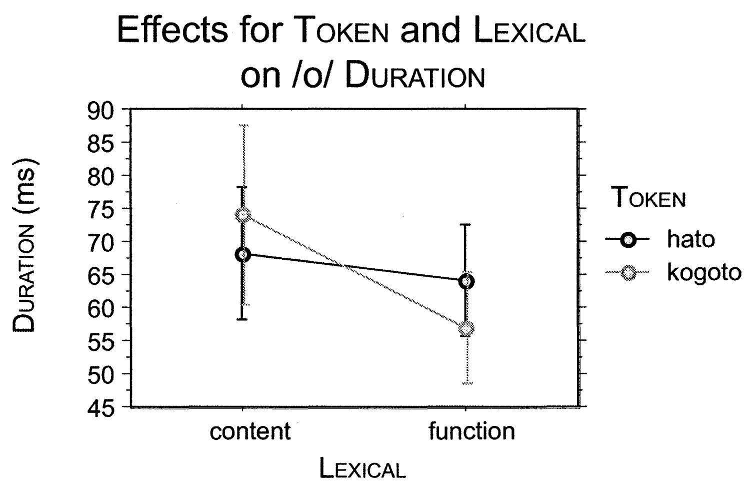

Figure 2.13 Error bars illustrating the interaction between LEXICAL by TOKEN for [o]

The token difference between [hitode] and [tade] was large. The following waveforms of [e] in [hitode] in Figure 2.14 and those of [e] in [tade] in Figure 2.15 illustrate the token difference between [e] in [hitode] and [e] in [tade]. In addition, spectrograms in Figure 2.16 are provided.

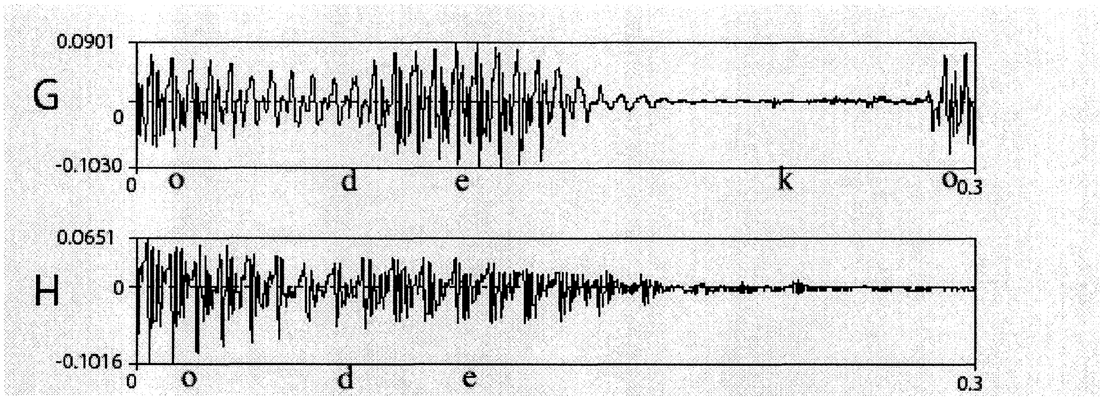

Figure 2.14 Waveforms illustrating content [e] (G) and function [e] (H) in [hitode] produced by the subject, "m2h."

The waveforms in Figure 2.14 illustrate that content [e] (G) was shorter than function [e] (H) in [hitode]. The next pair of waveforms illustrates [e] in [tade].

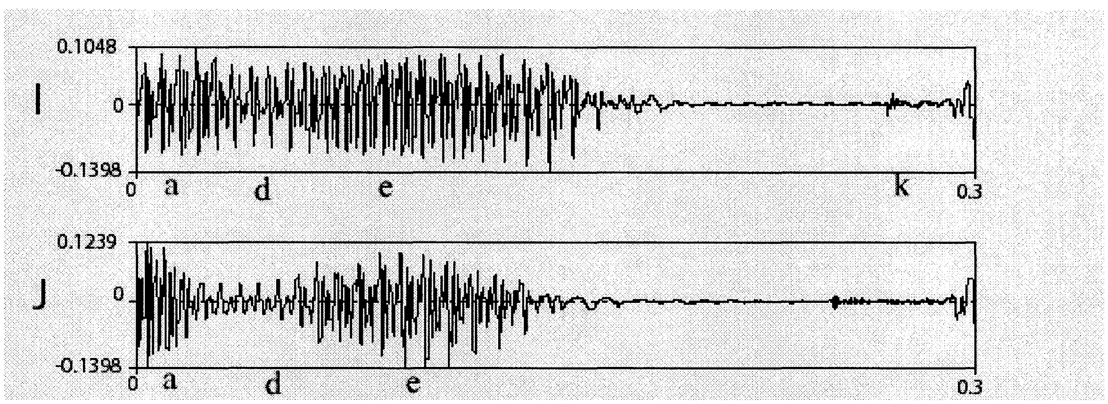

Figure 2.15 Waveforms illustrating content [e] (I) and function [e] (J) in [tade] produced by the subject, "m2h."

The waveforms in Figure 2.15 illustrate that content [e] (I) in [tade] was longer than function [e] (J). So [e] in [tade] accorded with the hypothesis although there is no significant effect for LEXICAL on DURATION with the token [tade], $F\,(1, 70) = 0.146$, $p = 0.7032$.

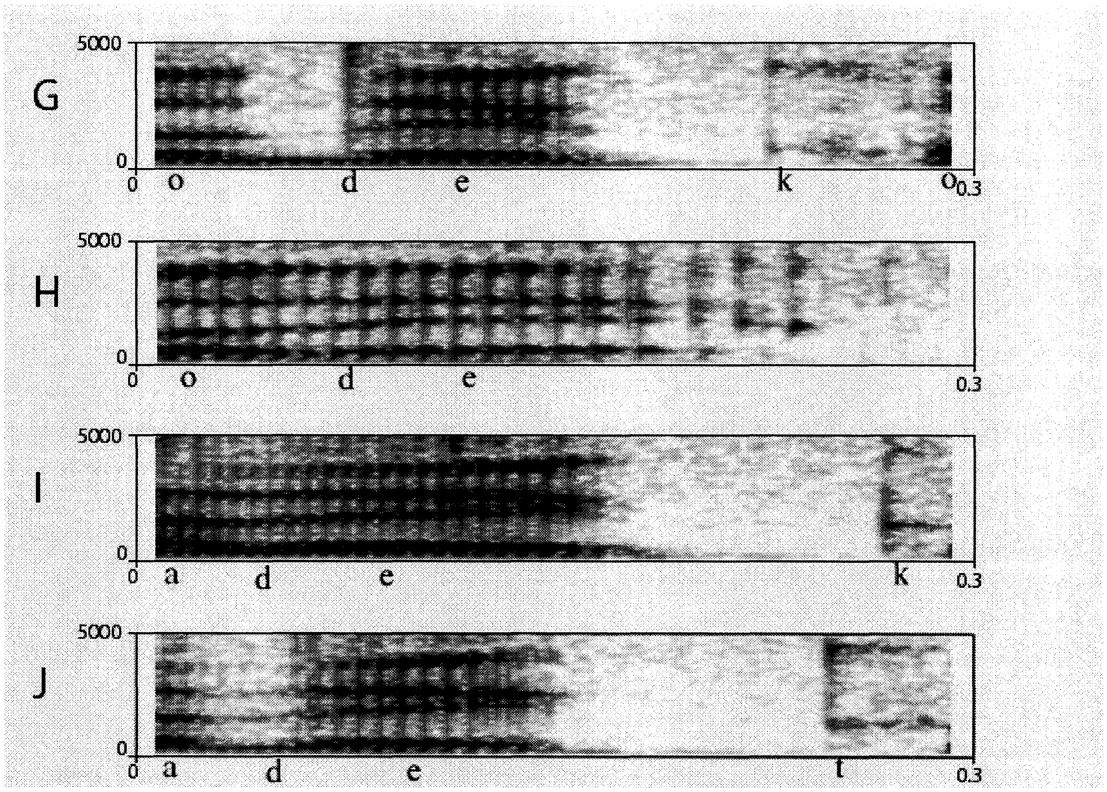

Figure 2.16 Spectrograms illustrating the variation in duration for [e]. G: Content [e] in [hitod<u>e</u>], H: function [e] in [hitod<u>e</u>], I: content [e] in [tad<u>e</u>], J: function [e] (J) in [tad<u>e</u>]. These speeches were produced by "m2h."

For [hitode], function [e] (H) was longer than content [e] (G). On the other hand, function [e] (J) was shorter than content [e] (I) for [tade]. The waveform and spectrogram (H) illustrate that function [e] in [hitode] was unexpectedly long.

2.4 Discussion 1

The overall results indicate that Japanese function vowels behave differently than content vowels. Durational reduction occurs in Japanese function vowels. In a previous study by Campbell (1992), the reduction was not observed probably because the investigator did not control for various factors influencing duration, which include the number of moras in a word, the preceding and following phones, and speaking rate.

The results split by vowel indicate that the durational differences of vowels [a] and [o] between content vs. function were statistically significant. On the other hand, the duration of content [e] (91ms) was shorter than the duration of function [e] (100ms), although the difference was not significant. The average duration for each token shows that unexpected results come from the token [hitode], while the content [e] in [tade] (76ms) was longer than function [e].

This unexpected result of [e] in [hitode] can be explained as follows. Vowel duration was influenced by syntactic structure. Ueyama (1999) found that prefinal lengthening occurred after major prosodic boundaries, the characteristics of which were long silent durations after major phrase boundaries and pitch resets.

I checked pitch and silent durations after the target vowels, and found that pitch resets occurred after function [e] in most utterances while it did not occur after content [e]. Figure 2.17 (horizontal axis: TIME, vertical axis: AMPLITUDE for waveform, FORMANT for spectrogram, PITCH for pitch track) illustrates that pitch movement of content [e], and Figure 2.18 illustrates that that of function [e].

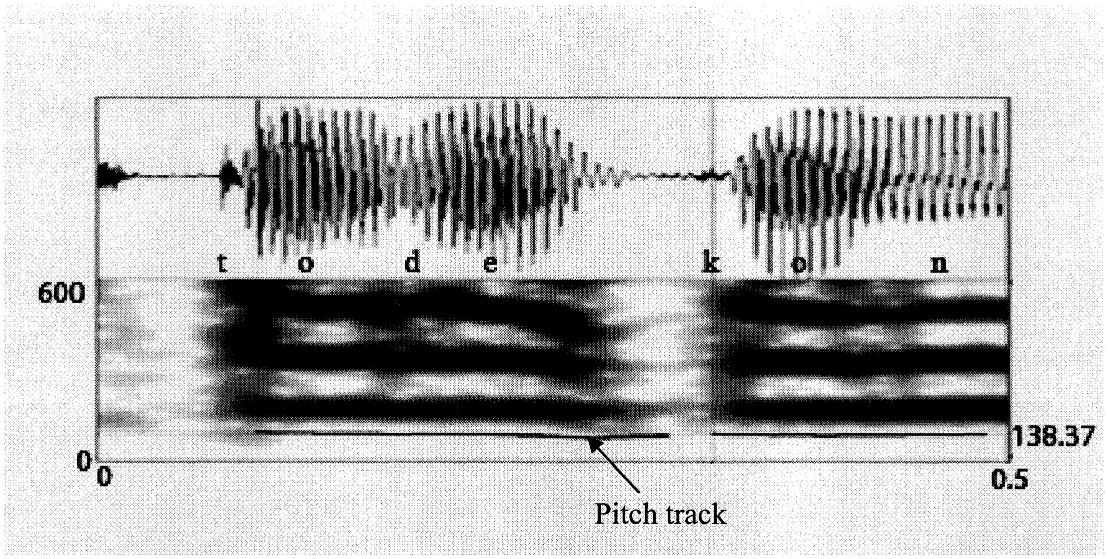

Figure 2.17 Narrow-band spectrogram and pitch track illustrating the pitch of the phones, [e] and [k] after content [e]. The range of the narrow spectrogram is from 0 to 600 Hz while the range of pitch track is 75 to 500 Hz. The number on the right, 138.37, is the pitch at [k] shown with the vertical line.

The pitch at the end of content [e] was almost the same as the pitch of the following syllable [ko], and creaky voicing was absent in Figure 2.17. Both indicate that no pitch reset occurred after content [e].

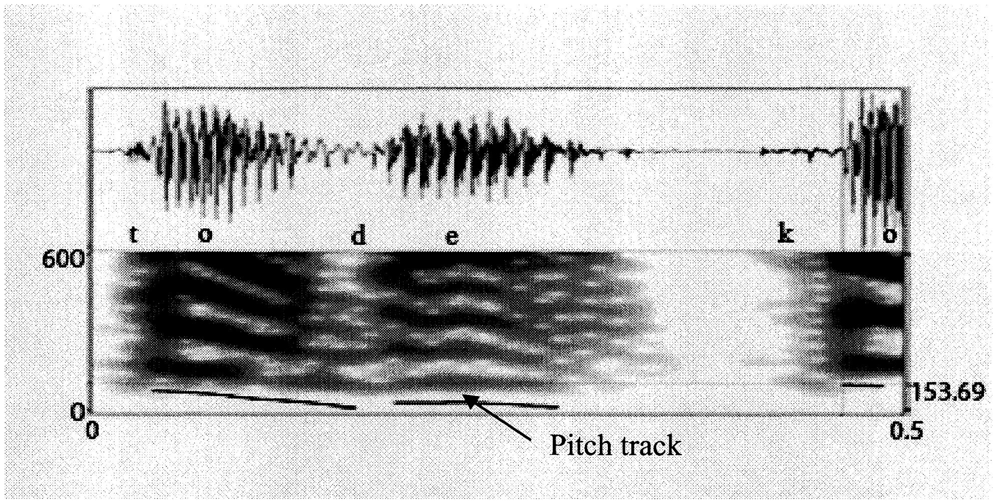

Figure 2.18 Narrow-band spectrogram and pitch track illustrating pitch reset after the function [e]. The number on the right, 153.69, is the pitch at the beginning of [o] shown with the vertical line. The pitch at [k] cannot be measured.

The dynamic fall in pitch on the syllable of function [e] with creaky voicing and the relatively large difference between the onset of [e] and the onset of the following [ko] in Figure 2.25, indicate pitch reset after the function [e].

I measured the pitch at the end of the target vowel and at the beginning of the following syllable [ko], and calculated the ratio of pitch of the following syllable, [ko], to the pitch of the target vowel, [e]. The results are plotted in Figure 2.19 (horizontal axis: LEXICAL, vertical axis: RATIO of pitch). In some cases, I could not measure pitch for function [e] because of creaky voicing; however, the harmonics showed that pitch resets occurred after function [e] but not after content [e].

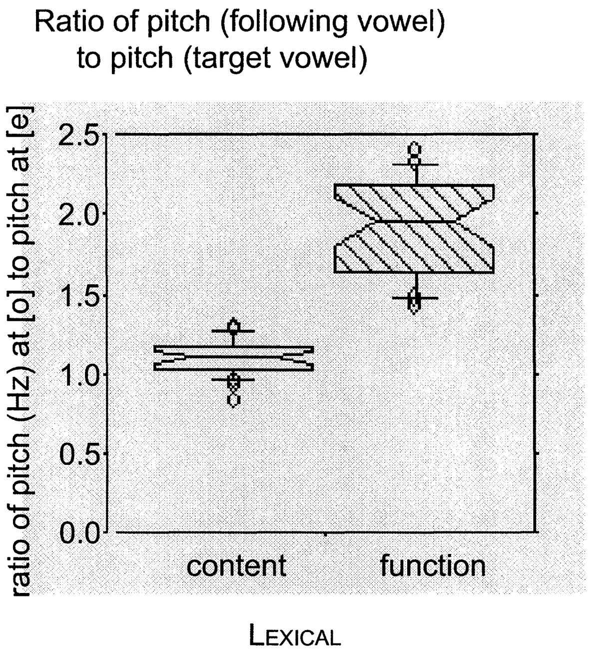

Figure 2.19 Boxplot illustrating the ratio of pitch (Hz) at the end of the target vowel [o] to pitch (Hz) at the beginning of the following syllable [e].

For comparison, I provide Table 2.8, which shows the pitch ratio of the target vowel and the following syllable for each token. The difference between the ratio for content vowel and that for function in [hitode] is noteworthy. This indicates that pitch resetting did take place; consequently, the duration of function [e] in [hitode] was lengthened.

Table 2.8 Pitch ratio: pitch (Hz) at the beginning of the following vowel to the pitch at the end of the target vowel

TOKEN	content	function
kiga	1.13	0.97
jinbutsuga	1.27	1.48
tade	1.16	1.19
hitode	1.92	1.11
hato	1.09	0.99
kogoto	1.07	1.29

Since silent duration is also a main indicator of a phrase boundary, I measured the duration of the silent period between the offset of voicing in the target vowel and the onset of the release burst in the following stop. The silent duration after [e] in [tade] and [hitode] are illustrated in the Figure 2.20.

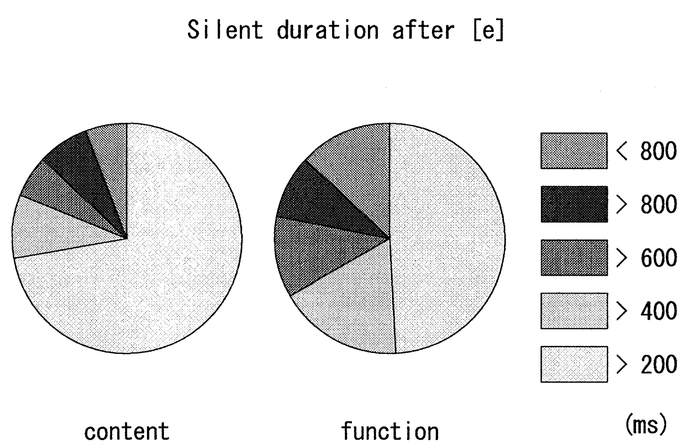

Figure 2.20 Pie graphs illustrating the difference between silent duration after content [e] (left) and silent duration after function [e] (right).

More than a half of the silent durations after function [e] were longer than 200 ms while only 27 % of those after content [e] were longer than 200 ms. The average silent duration after content [e] was 277 ms while those after function [e] was 459ms. This difference indicates that there was a difference of phrase levels between content vowel and function vowel. Probably this difference of phrase level has led to the anomalous durations of many of the function [e] cases.

2.4.1 Token differences

The number of moras and the length of sentences are two factors known to influence vowel duration. In general, the duration of a vowel becomes shorter as the number of moras increase. I had expected durations of vowels in long words to be short. However, the results showed the opposite tendency. For example, [e] in [hitode] was longer than the counterpart in [tade] although [hitode] consists of three moras and [tade] consists of two moras. In the case of [e] in [hitode], the different levels of phrase boundary influenced the durational difference.

3 Vowel reduction II: Spectral dimension

A second way in which vowel reduction can be measured is in the spectral domain. That is, reduced vowels tend towards centralization in the speaker's vowel space. Again, word class has been shown to be a factor in spectral reduction in other languages (ex. Van Bergem, 1993), and therefore, I expect to see similar effects in Japanese. Specifically, I expect vowels from function words will show more spectral reduction than equivalent vowels from content words.

3.1 Background

Research on spectral vowel reduction has been conducted for decades. I introduce some examples from previous research on spectral vowel reduction.

3.1.1 Lindblom (1963)

Lindblom (1963) conducted research on Swedish vowels in three consonantal frames. In Swedish, stress was accompanied by long duration as in American English. In the experimental design, six vowels were accompanied with stress and without stress in three consonantal environments: [b_b], [d_d] and [g_g]. According to Lindblom, the presence or absence of stress and the phrase positions caused variation in the duration of the target vowels, although he did not discuss it at length. The results of his experiments have been copied in Figure 3.1 (horizontal axis: TIME, vertical axis: FORMANTS), in which the relation between formants and duration was plotted, although the contrast between stressed and unstressed vowels was not.

Lindblom's conclusion that vowel centralization was due to short duration of reduced vowels referred to vowel undershoot hypothesis. I explain vowel undershoot model proposed by Stevens and House (1963). In vowel undershoot model, there are hypothetical F2 values for vowels and consonants. When a consonant and a following vowel are pronounced together, there is a transition of formant values from the F2 values of the consonant to the F2 values of the vowel. Formant displacement, which is the difference between hypothetical F2 values and observed F2 values, depends on two factors: the duration of a vowel and the distance between the F2 value of consonant and the F2 value of the vowel (hereafter, locus target distance). Since Lindblom concluded that the short duration led to the centralization of reduced vowels, his hypothesis was called "vowel undershoot hypothesis."

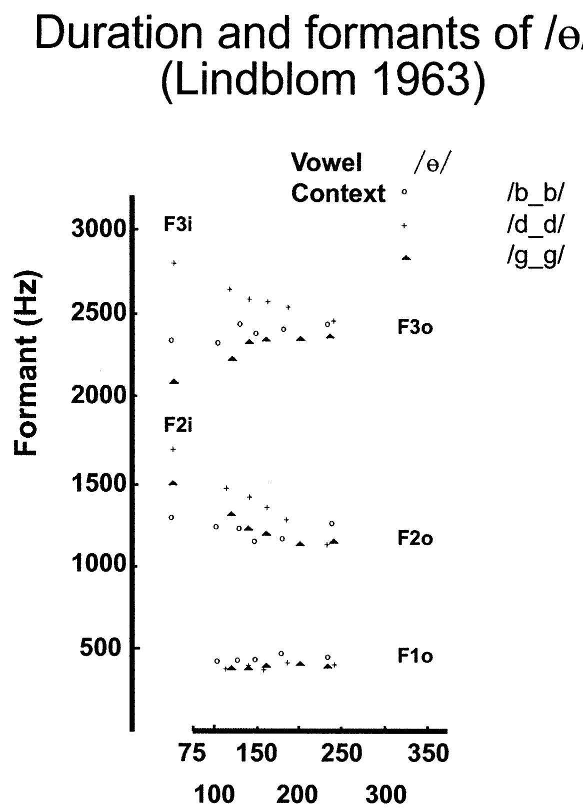

Figure 3.1 Scatter graph from Lindblom (1963) illustrating the relations between duration and formants of the vowel, [ɵ][1], in three contexts [b_b], [d_d] and [g_g]. F_{1o}, F_{2o} and F_{3o} indicate the series of F1, F2 and F3 averages in three consonantal contexts. Each point represents the averages of 4 individual measurements. The points referred F_{2i} and F_{3i} represent the averages of 20 measurements. F2s and F3s are "shown as a function of vowel-segment duration and consonantal context (1963 p1775)".

Figure 3.1 shows that there were three series of points for F2 and F3 and one series of points for F1, and that F2-values and F3-values changed as DURATION increased. In addition, F2-values gradually decreased as DURATION increased, but F1-values did not change. The change of F3-values depended on the context; F3 values in [b_b] decreased but those in [b_b] increased as DURATION increased. The results indicated that the F2 and

---

[1] The vowel [ɵ] is a rounded close-mid central vowel (http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/course/chapter1/vowels.html).

F3 values of vowels were dependent on DURATION and the context. Since the presence or absence of stress changed the duration of vowels, which was the main factor for the formant displacements, Lindblom concluded that the centralization of reduced vowels was due to undershoot. However, Lindblom did not test how the presence or absence of stress might affect formants.

3.1.2 Nord (1986)

The research conducted by Nord (1986) on Swedish vowels shows that the presence or absence of stress influences the formants; in other words, duration is not the sole factor affecting formant displacements. The results of formant measurement have been copied in Figure 3.2.

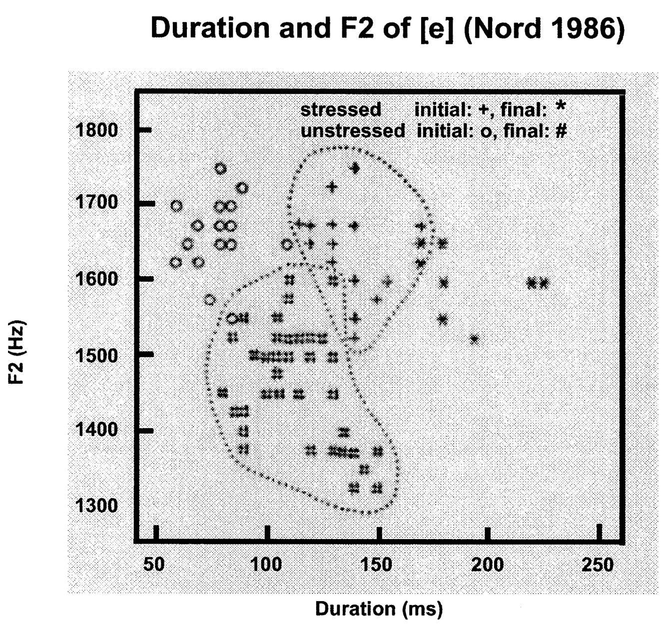

Figure 3.2 Scatter graph from Nord (1986) mainly illustrating the relation between F2s and durations. Plus marks: stressed initial [e]s, asterisks: stressed final [e]s, open circles: unstressed initial [e]s, and pound marks: unstressed final [e]s.

In Figure 3.2, although unstressed final [e]s were as long as stressed initial [e]s, F2s of unstressed [e] were lower than those of stressed [e]. He concluded that DURATION was not the only determinant of formant displacements. However, F2 values of unstressed initial [e]s were almost the same as those of stressed initial [e]s so the word-position might have stronger effects on F2 than the presence or absence of stress.

3.1.3 Van Bergem (1993)

Van Bergem conducted research on Dutch. He investigated the various effects of word class (function words vs. content words), stress and sentential accent. In addition, isolated syllables were also recorded as the ideal vowel position or the target formants. The subjects were 15 male speakers. His results were that vowels in function words, which were monosyllabic, and unstressed syllables in unaccented content words were strongly reduced. These reduced vowels seemed to be centralized; however, reduced vowels in various contexts show that it was not centralization but rather the results of consonant coarticulation effect. Van Bergem examined the shifts of formant frequencies for the vowel [ɛ] in different contexts. His measurements are reproduced in Figure 3.3.

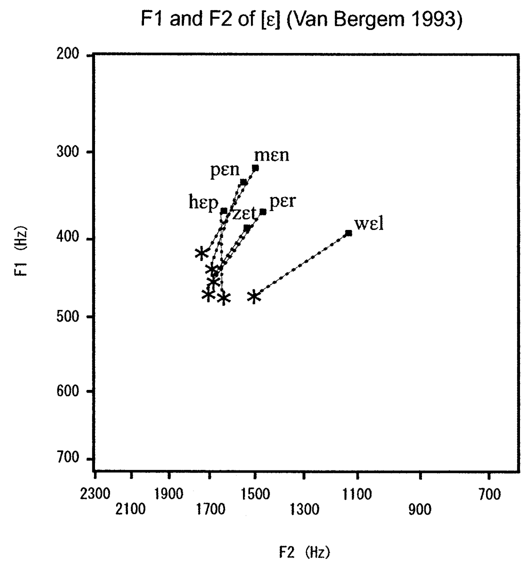

Figure 3.3 Scatter graph from Van Bergem (1993) illustrating the comparison of reduced vowels with vowels in isolated words in six different contexts. Dots: reduced vowels, asterisks: isolated syllables.

Figure 3.3 shows that F1 and F2 values of reduced vowel [ɛ] in [w_l] context were pronounced higher and further back in the mouth than those in isolated syllables, respectively; in other words, the F2 values of reduced [ɛ] in [w_l] context were not centralized. Furthermore, the F2 values of isolated [ɛ] in the [w_l] frame were less than F2 values in other contexts, so even in the isolated syllable, the vowel [ɛ] in [w_l] was affected by the surrounding consonants. Van Bergem concluded that spectral vowel reduction is caused by an undershoot effect, and that it is due to increased contextual assimilation.

3.1.4 Moon and Lindblom (1994)

Moon and Lindblom (1994) conducted research on American English. The purpose of their research was to prove that spectral vowel reduction was the result of coarticulation (i.e. vowel undershoot). The tokens were front vowels embedded in [w_l] frames and in [h_d] frames. F2 values in [h_d] context served as the targets, since this environment causes the least coarticulation effect. They predicted that the coarticulation of the consonants [w], [r] and [l] would cause greater decrease of the F2 of front vowels than the same vowels in [h_d] environments. The results indicate that the F2 values in [w_l] context were lower than the F2 values in [h_d] context and that F2 values in citation form were lower than F2 values in clear speech. Thus, they interpreted their results, the lack of centralization, was due to coarticulation. They confirmed that DURATION was the determinant of formant displacement. Furthermore, the results indicate that the displacements of formants in citation form depend on the duration more than those in clear speech[2].

In their experiments, Moon and Lindblom compared vowel formants in citation form with those in clear speech so they did not compare stressed vowels with unstressed vowels or content words with function words. Thus, the magnitude of difference between formants in citation forms and formants in clear speech was small.

Although they did not compare stressed vowels with unstressed vowels, one speaker differentiated clear speech from citation form for all vowels, shown in Figure 3.4, although other four other speakers showed only minor differences.

---

[2] For “clear speech”, subjects read the tokens as clearly as they could, while for “citation forms” the subjects read tokens at a comfortable rate and loudness.

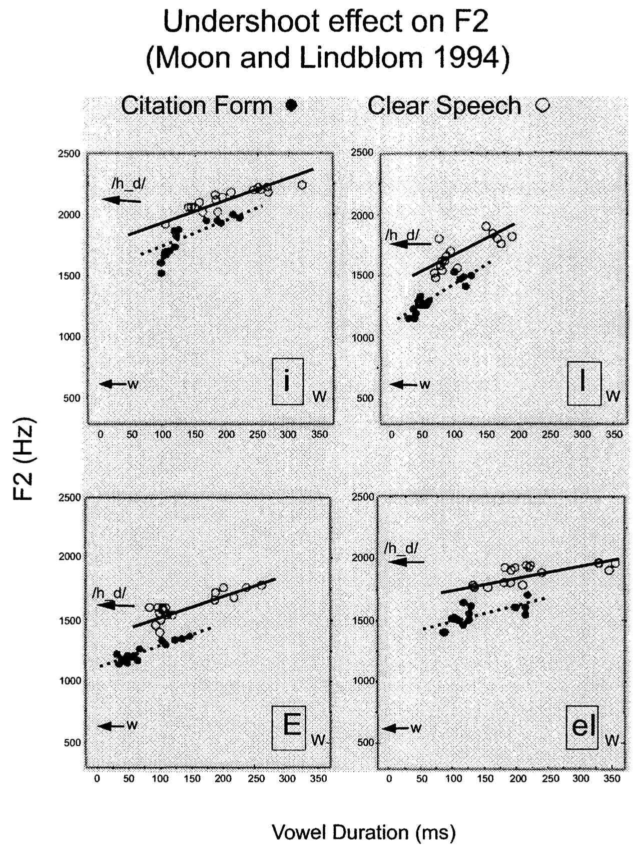

Figure 3.4 Scatter graphs from Moon and Lindblom (1994) illustrating F2-values of front vowels and durations. This subject clearly differentiated clear speech from citation form although other four subjects showed little difference. Closed circles: citation form, open circles: clear speech. I have added solid lines and dashed lines for clear speech and citation form, respectively.

In Figure 3.4, there seem to be two trend lines of F2 values by DURATION: one for clear speech and one for citation form. The F2 values in clear speech were closer to the F2 values in [h_d] context than those in citation form. Additionally, the F2s of [ɛ] for citation form were clearly less than those of [ɛ] for citation speech when the duration was in the vicinity of 100 ms as seen in Figure 3.4. These graphs indicate that formant displacements for clear speech and citation form were dependent on DURATION, but the trend line for clear speech did not agree with the trend line for citation form.

3.1.5 Keating and Huffman (1984)

Keating and Huffman (1984) conducted research on Japanese vowels. The tokens were words in a word-list and prose, uttered by seven male speakers. The researchers compared formants of vowels in words spoken in the word-list task with words spoken in the prose task. The results show that vowels spoken in the prose spread toward the center of the vowel space, although this study did not provide statistical analysis. Keating and Huffman considered this phenomenon to be centralization. However, the data in the prose task were possibly taken from non-equivalent consonant environments. According to the researchers, the onsets of [o] and [e] in the prose words were mainly apical (i.e. [t], [d], [s], [z] and [ɾ]); on the other hand, the onsets of all vowels in the word-list task were [h_b], and [b_#]. Furthermore, they wrote that most [a]s in the prose words were taken from a small number of morphemes, which suggests that onsets of [a]s were probably [t] (past marker [ta]), [s] (causative marker [sas]), [ɾ] (passive marker [ɾaɾ]), and [m] (polite [mas]). Thus, it is not clear whether spreading was due to centralization or coarticulation. Thus the current project includes investigation of whether or not Japanese function vowels are reduced in the spectral dimension.

3.2 Experiment 2: vowel centralization

Experiment 2 is designed to measure the degree of spectral reduction based on a lexical difference. I will propose the hypothesis in Section 3.2.1, and then describe the methodology of the experiment in Section 3.2.2, including the normalization. The results will be presented in Section 3.3.

3.2.1 Hypotheses

Hypothesis2: In Dutch, formants of vowels in function words such as "heb" [hɛp] are more central than those in full vowels (Van Bergem 1993). Likewise, I predict that the formants of vowels in Japanese function words will be more central than the formants of matched word-final vowels in content words.

3.2.2 Methodology

3.2.2.1 Materials

Recordings were the same as used in Experiment1.

3.2.2.2 Japanese vowels

The Japanese language has five vowels: /a/, /i/, /u/, /e/ and /o/. However, Japanese /u/ is unrounded so it is [ɯ]. Japanese vowels in vowel space are plotted in Figure 3.5.

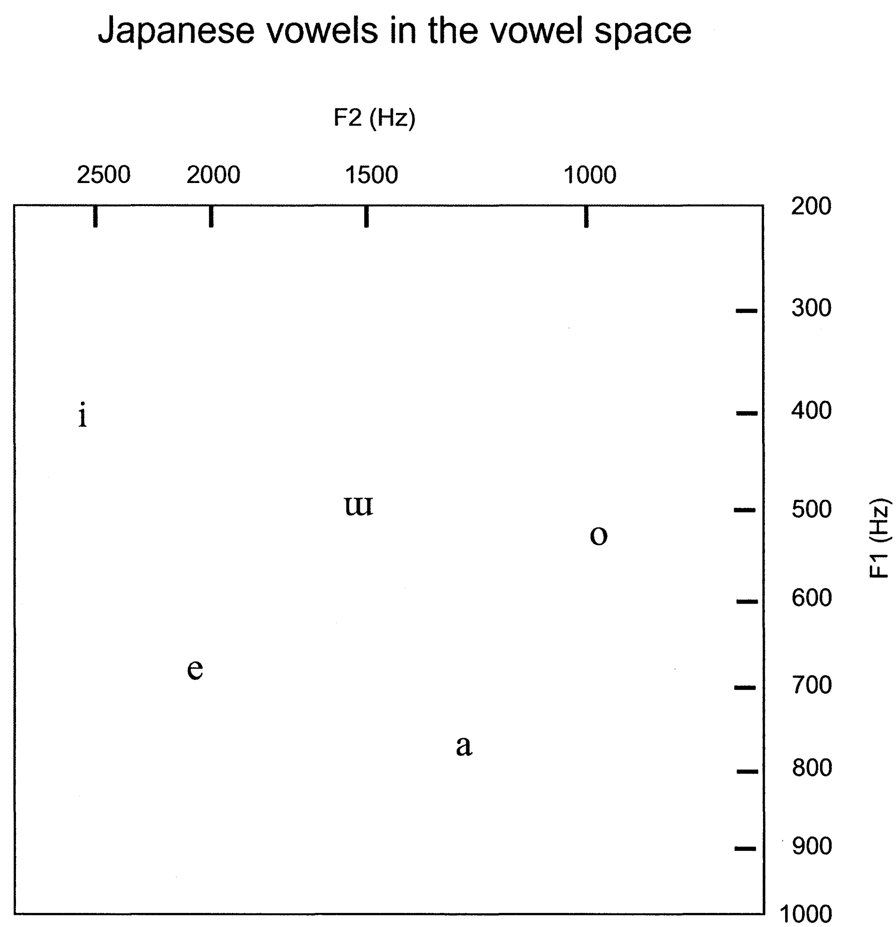

Figure 3.5 Formant plot illustrating an example of vowel space. These vowels are produced by a female speaker as isolated syllable in the context of /tVː/. These formants were measured at the mid-point of long vowels[3].

Since [ɯ] is unrounded, the F2 value of [ɯ] is lower than that of [u]. The Japanese vowels, which illustrated in Figure 3.5, show that [ɯ] seems to be a central vowel[4], although acoustically the symbol, [ɯ], represents a back vowel. In addition, Figure 3.5 shows that Japanese vowels are unevenly distributed. Next I will discuss the normalization, in which formants are normalized relative to the center of vowel space.

---

[3] The duration of each vowel was longer than 200 ms.

[4] A couple of Korean speakers pointed out that Japanese [ɯ] sounds like [ɨ] for them. Although Japanese /u/ is probably a central vowel (i.e. [ɨ]), I follow the convention to use /u/ and [ɯ] in writing.

3.2.2.3 Normalization

To discuss the centralization of reduced vowels, formants should be normalized. Formants are determined by the size of vocal tract; the vocal tract serves as a resonator, and the formants are determined to some degree by vocal tract length. Since men typically have larger vocal tracts, than women overall the formants of vowels produced by men are usually lower than those produced by women. For example, the mean F1 of [a] produced by the male subjects in this study was 597 Hz while the mean F1 produced by the female subjects was 720 Hz. Thus, formant normalization is necessary.

Nearey (1978) proposed Constant Log Interval Hypothesis (CLIH). In his hypothesis, log-transformed formants (hereafter, logF1 and logF2) were used, because the human auditory system distinguishes sounds based on Mel scaling, which is similar to log scaling. His hypothesis is that the ratio of logF1 of a speaker for a vowel to logF1 of another speaker for the same vowel is constant across vowel qualities. In other words, the ratio of a logF1 of Speaker S1 for vowel [a] to a logF1 of Speaker S2 for vowel [a] is the same as the ratio of a logF1 of Speaker S2 for vowel [i] to a logF1 of Speaker S2 for [i]. Based on his hypothesis, Nearey proposed a vowel normalization method (1989). This procedure is applied to F1 and F2 separately for each speaker. Below is an example of the process, using F1.

1. Calculate the mean F1 and F2 values for each vowel quality.
2. Convert all F1 values to log (base 10). I call the results logF1.
3. Sum up the means of log F1 and divide it by the number of categories. The result is a single grand mean of logF1. This grand mean serves as the center of the vertical vowel space.
4. Each logF1 is subtracted from F1 grand mean. The results are normalized F1.

Since Japanese vowels are distributed unevenly[5], the mean of F2 values for all vowels would be greater than the expected F2 values of the center of the vowel space. Thus, modification is necessary. In Japanese, F1 vowel category /i/ demonstrates the lowest F1 vowel while F1 vowel category /a/ demonstrates the highest F1 vowel. I used the mean of F1_i and F1_a for the F1 grand mean. Since the category /i/ demonstrates the highest F2 and /o/ the lowest, I attempted to use the mean of both F2s for the F2 grand mean. However, after plotting the F1 grand mean and the F2 grand mean, I realized that the F2 grand mean based on /i/ and /o/ was not located at the center of the vowel space[6]. So I calculated the mean of F2_e and F2_o and used this mean for the grand mean of F2. Then I subtracted the F1 and F2 grand means from the individual logF1 and logF2 values, respectively. I introduce the calculation procedure with an example (the example was written in `Courier fonts`):

1. The means of the formants for individual vowels are calculated.
```
For example, means are as follows:
   F1-a: 800, F2-a: 1827,
   F1-e: 600, F2-e: 2302,
   F1-i: 388, F2-i: 2688,
   F1-o: 743, F2-o: 1066,
   F1-u: 482, F2-u: 1696.
```

2. The results are log transformed.
```
log10 (F1-a): 2.903, log10 (F2-a): 3.262,
log10 (F1-e): 2.778, log10 (F2-e): 3.362,
log10 (F1-i): 2.589, log10 (F2-i): 3.429,
log10 (F1-o): 2.871, log10 (F2-o): 3.028,
log10 (F1-u): 2.684, log10 (F2-u): 3.229.
```

---

[5] There is only one back vowel, [o], while there are two front vowels, [i] and [e]. The calculated F1 and F2 values based on Nearey's method, using the same data, are 582 Hz and 1828 Hz respectively. It is located between the central and front vowels.

[6] I attempted to investigate the gender difference, which I did not pursue because there was no interaction between LEXICAL and GENDER. When the means of vowels were plotted, the pattern of vowels produced by female speakers was parallel to that by male speakers.

3. For F1, the mean of log F1_i and log F1_a is used as the center of vowel space.
```
F1: (2.589 +2.903) /2 = 2.746
```
For F2, the mean of F2_e and F2_o is used as the center of vowel space.
```
F2: (3.362 + 3.028) /2 = 3.195
```
The log10 of raw data (each measurement) are subtracted from the center of vowel space.
```
A measurement: F1 = 881.79 Hz,
               log10 (881.79) = 2.9454,
               Center of F1: 2.7704,
               G1 = 2.9454- 2.7704 = 0.1750
```

Based on this modified method, I calculate the normalized F1 and F2, and call these normalized F1 and F2 as G1 and G2, respectively.

In Figure 3.6 which is a vowel plot showing F2 (horizontal axis) by F1 (vertical axis) values with Bark scaling, I plotted Japanese five vowels, whose F1 and F2 values were used in the calculation above. The calculated center of vowel space seems indeed to be located in the center of vowel space.

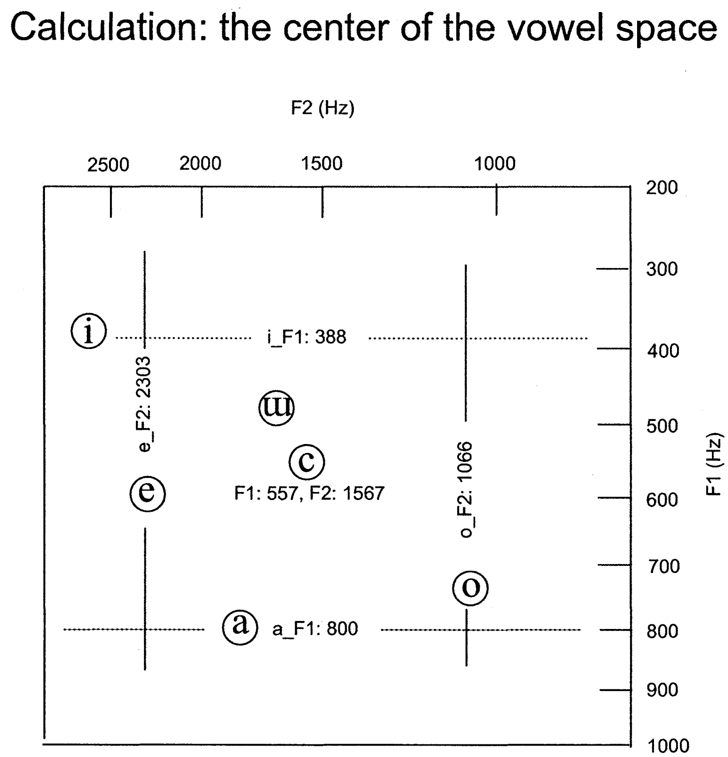

Figure 3.6 "Plot Formant" illustrates the calculation of the center[7] of the vowel space. The vertical lines show the F2 values of [e] and [o] while the horizontal dashed lines show that the F1 values of [i] and [a]. Based on these values, the F1 values and F2 values for the center of the vowel space were calculated. The F1 value at the center of the vowel space was at the middle point of F1-a and F1-i. The F2 value at the center of the vowel space is at the middle point of F2-e and F2-o. The circled c represents the calculated center.

3.2.2.4 Distance from the center of vowel space

In my hypothesis, function vowels are more centralized than content vowels. In order to test if function vowels are more central than content vowels, I calculate the distance from the center of vowel space. The distance from the center of the vowel space

[7] The center of vowel space in Figure 3.6 was calculated as follows:
F1 of center of vowel space in logarithmic scaling: 2.746
F2 of center of vowel space in logarithmic scaling: 3.195
MidF1 = power (10, 2.746) = 557
MidF2 = power (10, 3.195) = 1567

for each person was calculated based on normalized F1 and F2 (i.e. G1 and G2). Since G1 and G2 were relative locations to the center of the vowel space, the Euclidean distance was calculated as follows:

$$distance = \sqrt{G1^2 + G2^2} \qquad \textbf{(3-1)}$$

Hereafter, distance refers to the distance from the center of the vowel space.

3.3 Results of Experiment 2: Spectral dimension

I will test the effects for Lexical in three measurements: F1 & F2, G1 & G2 and DISTANCE. The analyses of formants for individual vowels showed mixed results: for [a], there was a significant effect for LEXICAL on F1, but for [e] and [o], there was no significant effect for LEXICAL on F2. Overall results for distance indicate that function vowels were more central than content vowels. That is, the mean distance of function vowels from the center of the vowel space were significantly shorter than their content vowel counterparts. However, in analyzing individual vowels, only vowel [a] showed significant differences. I present the results of F1 & F2, the results of G1 & G2 and the results of distances. Inferential statistical results will be presented for F1 & F2 and G1 & G2, i.e. for both the raw and normalized data, for comparison. The limitations of comparing non-normalized data were discussed in Section 3.2.2.3. Discussions of overall patterns in the data will be based primarily upon the results provided for normalized data.

3.3.1 Formants

The means and standard deviations are provided in Table 3.1.

Table 3.1 Means and standard deviations of F1 and F2 (Hz)

VOWEL		all	content	function
a	F1	653	696	610
	(SD)	(105.9)	(98.1)	(95.9)
	F2	1502	1484	1519
	(SD)	(187.1)	(202.5)	(170.0)
e	F1	514	513	516
	(SD)	(79.9)	(84.0)	(76.2)
	F2	1987	1998	1977
	(SD)	(245.1)	(243.4)	(248.4)
o	F1	501	505	496
	(SD)	(71.4)	(69.6)	(73.3)
	F2	1097	1086	1107
	(SD)	(169.8)	(170.5)	(169.8)

The means of measurements are plotted in Figure 3.7, a vowel plot showing F2 (horizontal axis) by F1 (vertical axis) values with Bark scaling.

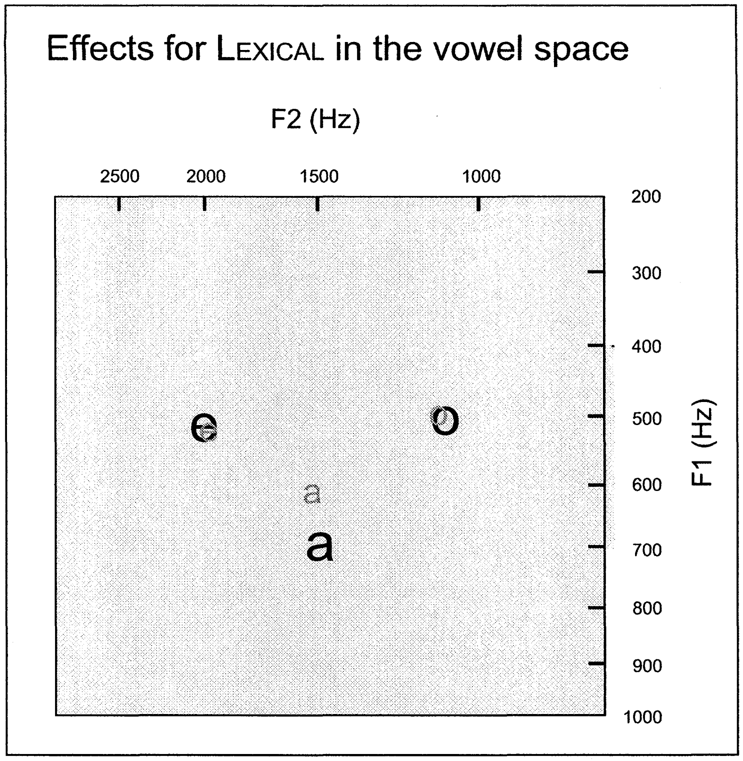

Figure 3.7 "Plot Formants" illustrating the means of formant values with large fonts for content vowels and small gray fonts for function vowels. This graph was plotted based on the means of raw data. The graph shows that F1 of function [a] was clearly lower than content [a]. In addition, function [e] and [o] were slightly closer to the center of vowel space than the content vowel counter-parts.

The formant plot in Figure 3.7 illustrates that F1 values of [a] in function words were noticeably less than content [a] values. In addition, [e] and [o] in function words were slightly closer to the center of the vowel space than content-vowel counterparts. This graph illustrates that function [a] shows robust lexical effect.

The results of a one-way repeated-measures ANOVA with AVERAGE F1 as dependent variable and LEXICAL (content and function) as independent variables indicate that there is a significant effect for LEXICAL $F(1, 10) = 15.858, p = 0.003$. The results of a one-way repeated-measures ANOVA with AVERAGE F2 as dependent variable and LEXICAL (content and function) as independent variables indicate that there is no significant effect for LEXICAL $F(1, 10) = 1.757, p = 0.215$. However, the direction of F2 values for reduced [e] is the opposite for that for [o]. To confirm it, I calculated the differences between content and function for AVERAGE F2 [e] and F2 [o] and call them DIFF_E and DIFF_O. The result shows that there is a negative correlation between DIFF_E and DIFF_O Pearson's $r(9) = -0.587, p = 0.058$. Although it is not significant, it indicates that it is inappropriate to compare F2-values of all vowels in content words with function counterparts. Next, I will discuss individual vowels.

3.3.1.1 Individual vowels

For [a], change in F1 is an indicator of centralization, while for [e] and [o], change in F2 is an indicator of centralization. Thus F1 of [a] and F2 of [e] and [o] are discussed here. A boxplot (horizontal axis: individual vowels, vertical axis: formants) in Figure 3.8 illustrates medians and ranges of F1 and F2 values.

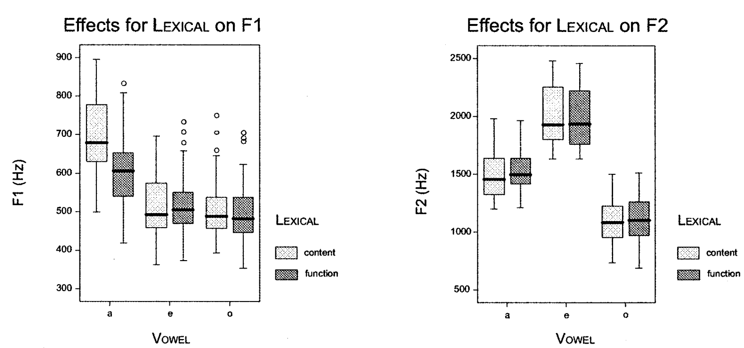

Figure 3.8 Boxplot illustrating F1 and F2

The boxplot in Figure 3.8 illustrates that F1-values of function [a] were noticeably less than those of content [a]; however, there was no differences between content and function for F2-values of [e] and [o] in Figure 3.8.

The results of a two-way repeated-measures ANOVA with AVERAGE F1 as dependent variables, and LEXICAL (content, function) and VOWEL as the independent variables are summarized in Table 3.2.

Table 3.2 Effects for LEXICAL and VOWEL on AVERAGE F1

	VOWEL	LEXICAL	VOWEL x LEXICAL
df	[2, 20]	[1, 10]	[2, 20]
F	120.968	15.966	24.095
p	< 0.001	0.003	< 0.001

The results show that there is a significant effect for LEXICAL, for VOWEL and the interaction of two. Results of a two-way ANOVA with AVERAGE F2 as dependent variables, and LEXICAL (content, function) and VOWEL as the independent variables are summarized in Table 3.2.

Table 3.3 Effects for LEXICAL on AVERAGE F2

	Vowel	Lexical	Vowel x Lexical
df	[2, 20]	[1, 10]	[2, 20]
F	252.463	1.542	3.065
p	< 0.001	0.243	0.076

There was a significant effect for VOWEL on F2 but no significant effects for LEXICAL on F2. The lack of significance is expected since the direction of F2 reduced [e] is the opposite to the direction of F2 reduced [o].

Since I am interested in the effect for LEXICAL for individual vowels, all measurements of F1 and F2 values for individual vowels are submitted to a repeated-measures ANOVA with F1 or F2 as dependent variable and LEXICAL as independent variable.

Table 3.4 Effects for LEXICAL on F1 and F2 split by VOWELS

		a	e	o
	df	[1, 65]	[1, 62]	[1, 65]
F1	*F*	73.401	0.262	2.064
	p	< 0.001	0.611	0.156
F2	*F*	5.498	2.863	1.685
	p	0.022	0.096	0.199

The results show that there is a significant effect for LEXICAL on F1 and F2 values for [a]. However, there is no significant effect for LEXICAL on F2 values for [e] and [o] as expected as the evidence of centralization.

I used Wilcoxon signed-rank tests to see the tendency. The results of the Wilcoxon signed-rank tests are summarized in Table 3.5.

Table 3.5 Effects for LEXICAL on F1 (Wilcoxon signed-rank test)

VOWEL	*Z*	*p*	Σ negative ranks	Σ positive ranks
all	-5.068	< 0.001	13554	5556
a	-6.442	< 0.001	2114	97
e	-0.828	0.407	887	1129
o	-1.549	0.121	1348	863

The results indicate that there is a significant effect for LEXICAL on F1 only for [a] but not for [o] or [e]. Next table shows the lexical effect for F2 values by the Wilcoxon signed-rank test.

Table 3.6 Effects for LEXICAL on F2 (Wilcoxon signed-rank tests)

VOWEL	*Z*	*p*	Σ negative ranks	Σ positive ranks
all	-2.037	0.042	7948	11162
a	-2.654	0.008	690	1521
e	-1.677	0.093	1253	763
o	-2.073	0.038	781	1430

The results of the Wilcoxon signed-rank tests show that there is a significant effect for LEXICAL on overall F2 and on F2 for [a] and [o].

3.3.1.2 Token differences

To obtain possible statistical differences for tokens, individual words were examined. The means and the standard deviations are summarized in Table 3.7.

Table 3.7 Means and standard deviations of F1 and F2 split by TOKEN

VOWEL	TOKEN		content	function
a	kiga	F1 (Hz)	673	602
		(SD)	(95.7)	(101.0)
	jinbutsuga	F1 (Hz)	720	618
		(SD)	(96.3)	(91.3)
e	hitode	F1 (Hz)	532	540
		(SD)	(76.0)	(56.7)
	tade	F1 (Hz)	495	494
		(SD)	(88.1)	(85.4)
o	hato	F1 (Hz)	500	492
		(SD)	(56.5)	(70.2)
	kogoto	F1 (Hz)	511	500
		(SD)	(81.1)	(77.3)
a	kiga	F2 (Hz)	1549	1573
		(SD)	(220.7)	(182.6)
	jinbutsuga	F2 (Hz)	1420	1466
		(SD)	(161.4)	(139.4)
e	hitode	F2 (Hz)	1936	1917
		(SD)	(244.6)	(239.8)
	tade	F2 (Hz)	2054	2031
		(SD)	(231.8)	(246.9)
o	hato	F2 (Hz)	1128	1127
		(SD)	(182.0)	(166.1)
	kogoto	F2 (Hz)	1045	1088
		(SD)	(149.5)	(173.7)

Since there were no tokens used for two individual vowels, and the differences among individual vowels were large, the data for individual vowels were separately submitted to a two-way repeated-measures ANOVA with AVERAGE F1 or F2 as a dependent variable and LEXICAL and TOKEN as independent variables. The results of the ANOVA are summarized in Table 3.8.

Table 3.8 Effects for TOKEN and LEXICAL on AVERAGE F1 and F2

	VOWEL		TOKEN	LEXICAL	LEXICAL x TOKEN
F1	a	*df*	[1, 10]	[1, 10]	[1, 10]
		F	5.213	31.387	1.253
		p	0.046	< 0.001	0.289
	e	*df*	[1, 9]	[1, 9]	[1, 9]
		F	8.090	0.112	0.175
		p	0.019	0.745	0.685
	o	*df*	[1, 10]	[1, 10]	[1, 10]
		F	0.577	1.554	0.031
		p	0.465	0.241	0.864
	a	*df*	[1, 10]	[1, 10]	[1, 10]
F2		*F*	23.928	2.729	0.210
		p	0.001	0.130	0.656
	e	*df*	[1, 9]	[1, 9]	[1, 9]
		F	168.694	1.394	0.000
		p	< 0.001	0.268	0.990
	o	*df*	[1, 10]	[1, 10]	[1, 10]
		F	27.504	2.959	0.802
		p	< 0.001	0.116	0.392

The results indicate that there was a significant effect for LEXICAL on AVERAGE F1 for only vowel [a], and on AVERAGE F2 for no vowels. There was a significant effect for TOKEN on AVERAGE F1 for [a] and [e], and on AVERAGE F2 for all three vowels. To illustrate token differences, means[8] of F1 and F2 values are plotted in Figure 3.9, a scatter graph (horizontal axis: F2, vertical axis: F1).

[8] Means are calculated based on all measurements of F1 and F2 values.

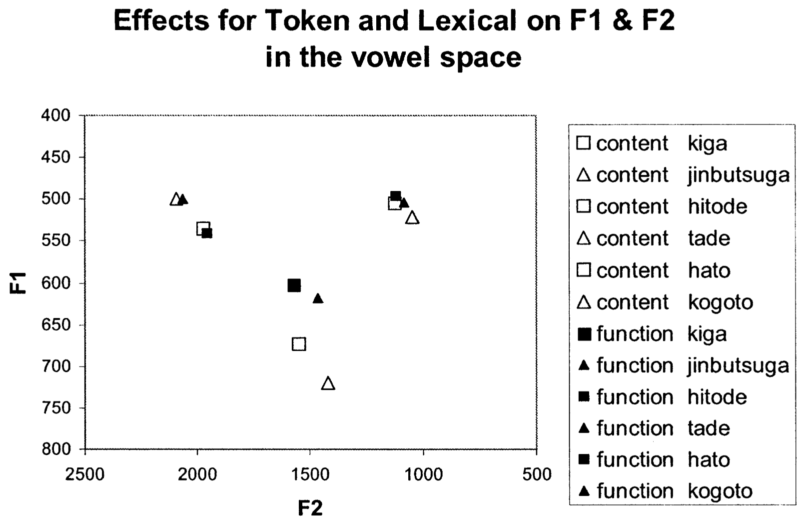

Figure 3.9 Scatter graph illustrating token differences. Content: unfilled, Function: filled. For [a], lexical difference was greater than token difference although token difference was noticeable. For [e], token difference was greater than lexical difference. Lexical difference of [o] in [kogoto] was greater than that in [hato].

The scatter graph in Figure 3.9 illustrates that for [a], lexical difference was greater than token difference although token difference was noticeable. For [e], token difference was greater than lexical difference, and for [o], the differences for TOKEN and LEXICAL were not clear. For [o], the mean F2 of content [o] in [hato] was greater than function counterparts.

The graph shows that function vowels were consistently more central than content vowels, and that lexical difference was greater than token difference for [a], but not for [e] and [o]. The means of content vowels and function vowels indicate that with exception of [hato], function vowels were more central than content vowels. Since the difference was subtle, it was not clearly apparent in Figure 3.9. Thus, F2 values of content [o] in [hato] could be considered almost the same as those of function [o]. To illustrate the interactions between LEXICAL and TOKEN, error-bar graphs are provided. The error-bar graph in Figure 3.10 illustrates the interaction for F1 [a] between TOKEN and LEXICAL.

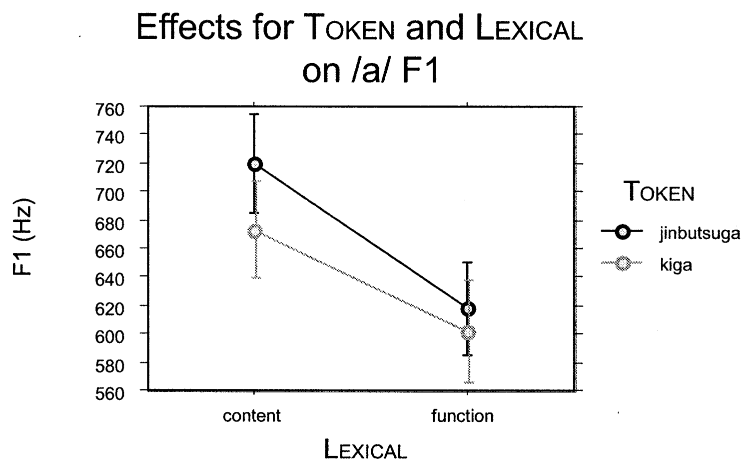

Figure 3.10 Error-bar graph illustrating token difference for F1 of vowel [a]. The F1 means of [a] in [dʒinbɯtsɯga] were greater than those in [kiga]. The token difference was less than the lexical difference.

The error-bar graph in Figure 3.10 shows that the F1 means of [a] in [kiga] were less than those in [dʒinbɯtsɯga], and that the lexical difference was greater than the token difference. In addition, the error-bar for [kiga] was not parallel to that for [dʒinbɯtsɯga] so there was interaction between TOKEN and LEXICAL. The next error-bar graph illustrates the interaction between TOKEN and LEXICAl for F2 [a].

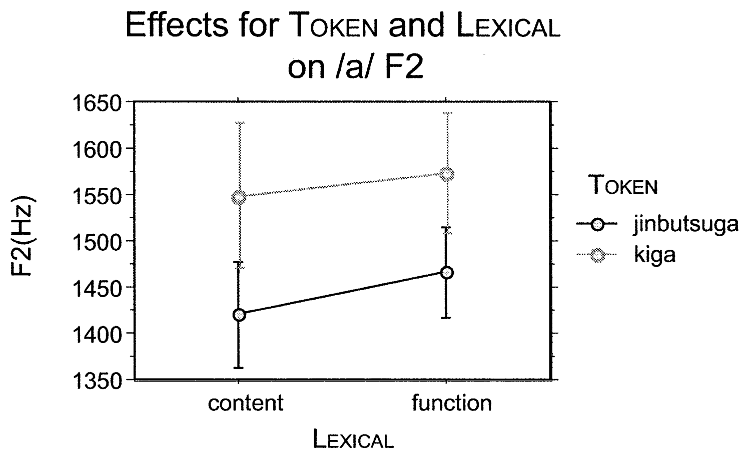

Figure 3.11 Error-bar graph illustrating the interaction between LEXICAL and TOKEN for F2 [a]. The F1 means of [a] in [kiga] were greater than those in [dʒinbɯtsɯga]. The error-bar for [dʒinbɯtsɯga] was parallel to that for [kiga] so there was no interaction between LEXICAL and TOKEN. The lexical difference was less than the token difference.

The error-bar graph in Figure 3.11 shows that the F2 means of [a] in [dʒinbɯtsɯga] were greater than those in [kiga], that the lexical difference was less than the token difference, and that the error-bar for [dʒinbɯtsɯga] was parallel to that for [kiga] so there was no interaction between LEXICAL and TOKEN. The next error-bar graph in Figure 3.13 illustrates the token difference of F1 [e] in [hitode] and [tade].

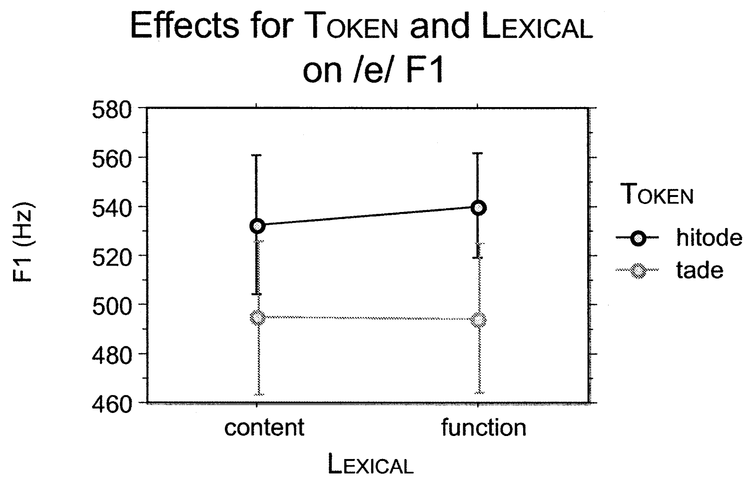

Figure 3.12 Error-bar graph illustrating the interaction between LEXICAL and TOKEN for F1 [e]. The mean F1 of content [e] in [tade] was almost the same as that of function [e]. The F1 means of [e] in [hitode] were greater than those in [tade]. The token difference was greater than the lexical difference.

The error-bar graph in Figure 3.12 shows that the F1 means of [e] in [hitode] were greater than those in [tade], that the F1 mean of content [e] in [tade] was almost the same as that of function [e], and that the error-bar for [tade] was almost parallel to that of [hitode] so there was no interaction between LEXICAL and TOKEN. The next error-bar graph illustrates the interaction between LEXICAL and TOKEN for F2 [e].

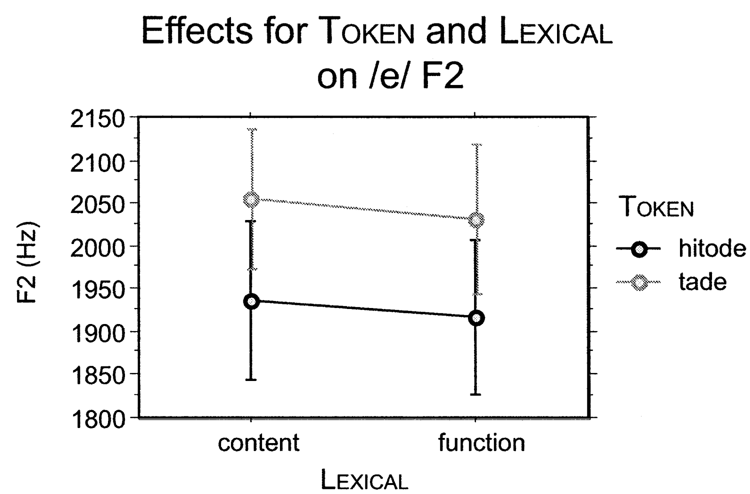

Figure 3.13 Error-bar graph illustrating token difference and lexical difference for F2 of [e]. The F2 means of [e] in [hitode] were less than those in [tade]. The token difference was greater than the lexical difference.

The error-bar graph in Figure 3.13 shows that the F2 means of [e] in [hitode] were less than those in [tade], and that the token difference was greater than the lexical difference. In addition, the error-bar for [tade] was almost parallel to that for [hitode] so there was no interaction between LEXICAL and TOKEN. The next error-bar graph illustrates the interaction of F1 [o] between TOKEN and LEXICAL.

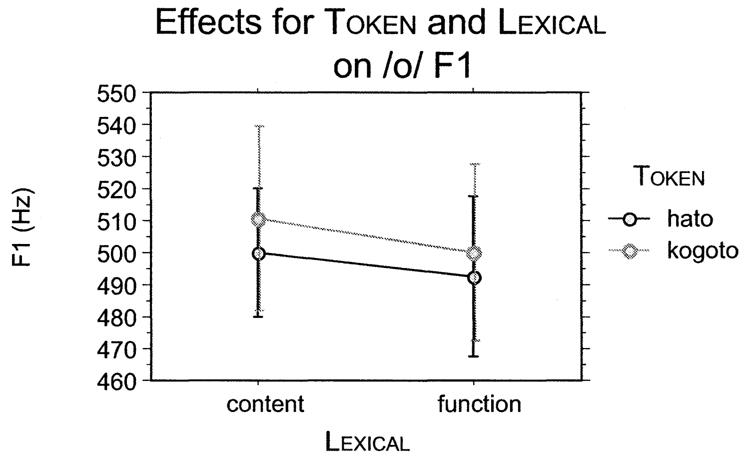

Figure 3.14 Error-bar graph illustrating the interaction between Lexical and Token for F1 [o]. The F1 means of [o] in [hato] were less than those in [kogoto]. There was almost no lexical difference for [hato] while there was clear lexical difference for [kogoto]. The error-bar for [hato] was not parallel to that for [kogoto].

The error-bar graph in Figure 3.14 illustrates that the F1 means of [o] in [hato] were less than those in [kogoto], and that F1 mean of content [o] in [kogoto] was clearly greater than that of function [o]. The error-bar of [o] in [hato] was not parallel to that of [kogoto] so there was interaction between the LEXICAL and TOKEN as the results of the factorial ANOVA indicate. The next error-bar graph illustrates the interaction between LEXICAL and TOKEN for F2 [o].

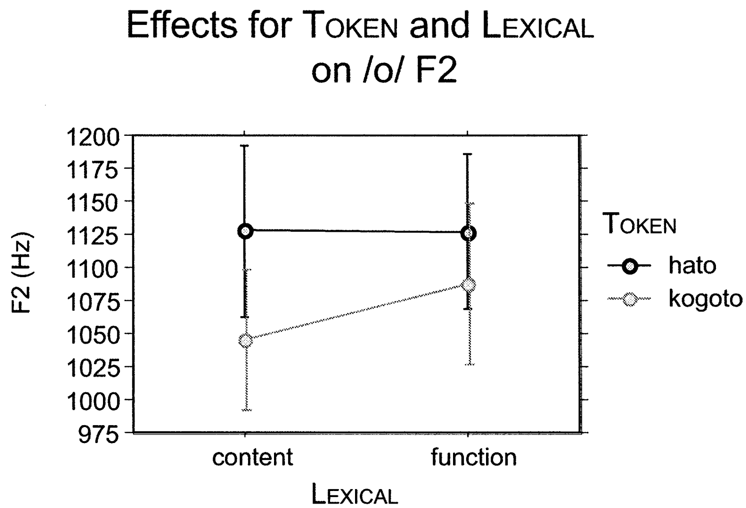

Figure 3.15 Error-bar graph illustrating the interaction between TOKEN and LEXICAL for F2 of [o]. The F2 means of [o] in [hato] were greater than those in [kogoto]. The F2 mean of content [o] in [hato] was slightly greater than that of function [o], while that of content [o] in [kogoto] was clearly less than that of function [o].

The error-bar graph in Figure 3.15 shows that the F2 means of [o] in [hato] were greater than those in [kogoto] and that the F2 mean of content [o] in [hato] was slightly greater than that of function [o] while that of content [o] in [kogoto] was clearly less than that of function [o]. Again, token [o] in [hato] shows unexpected results. However, the results of the ANOVA show that there was no significant effect for LEXICAL on F2 of [o] in [kogoto]. Since the lexical effect on F2 of [o] in [hato] moved the opposite direction of that in [kogoto], there was not significant interaction between TOKEN and LEXICAL. The syllable following [to] in [kogoto] was [ka], while that in [hato] was supposed to be [kɯ]. However, [ɯ] in the following syllables of function [to] in [hato] tended to be devoiced. The following waveforms in Figure 3.16 and spectrograms in Figure 3.17 illustrate it.

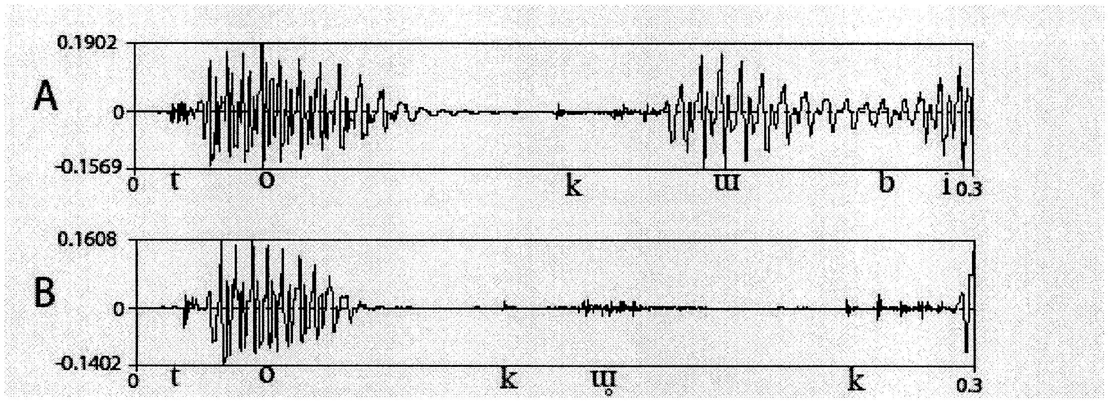

Figure 3.16 Waveforms illustrating regular vowel, [ɯ], after content [to] (A), but devoiced vowel, [ɯ̥], after function [to] (B).

In Figure 3.16, there were glottal pulses with very tiny amplitude in the waveform after function [to] (B), while there were clear peaks after content [to] (A).

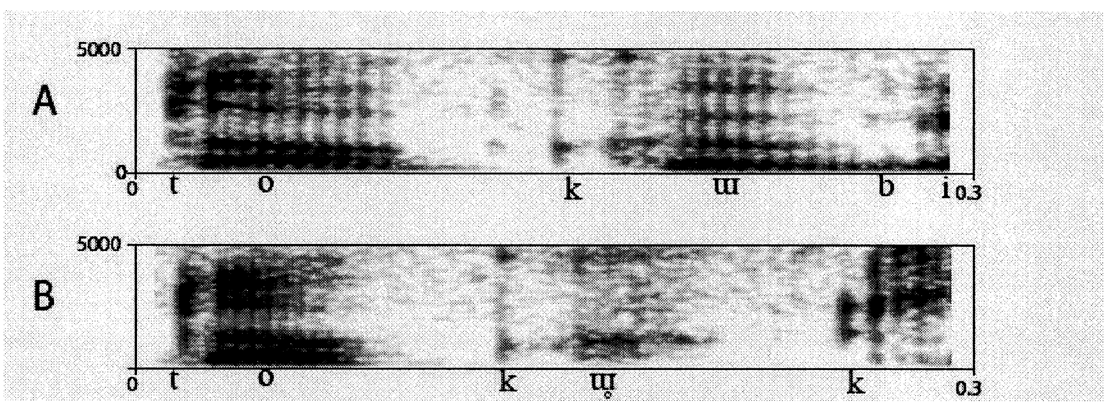

Figure 3.17 Spectrograms illustrating regular vowel [ɯ] after content [to] (A), but devoiced vowel [ɯ̥] after function [to] (B).

In Figure 3.17, [ɯ̥] looks like fricative noise with a light line after function [to] (B), while [ɯ] with four clear dark lines appear after content [to] (A). The waveforms and spectrograms illustrate devoiced [ɯ̥] after function [to] and regular [ɯ] after content [to]. This devoiced vowel probably contributes to the unexpected results to some extent.

3.3.2 Results of normalized formants

I provide the results of G1 and G2, which were calculated based on the modified Nearey's method. The means and standard deviations of G1 and G2 are provided in Table 3.9.

Table 3.9 Means and standard deviations of G1 and G2

VOWEL		all	content	function
a	G1	0.1563	0.1878	0.1247
	(SD)	(0.06542)	(0.05347)	(0.06121)
	G2	0.0581	0.5078	0.0655
	(SD)	(0.04502)	(0.04549)	(0.04367)
e	G1	0.0516	0.0504	0.0529
	(SD)	(0.05842)	(0.05785)	(0.05942)
	G2	0.1842	0.1870	0.1814
	(SD)	(0.02591)	(0.02556)	(0.02618)
o	G1	0.0449	0.04932	0.0405
	(SD)	(0.04957)	(0.04636)	(0.05257)
	G2	-0.0814	-0.0864	-0.0765
	(SD)	(0.05081)	(0.05216)	(0.04932)

In Table 3.9, the range of standard deviations for G1 is similar to that for G2. On the other hand, in Table 3.1, the range of standard deviations for F1 is almost half of F2 one. This indicates that I compare the differences between content and function G1 with counterpart of G2. G1 and G2 were plotted in Figure 3.18.

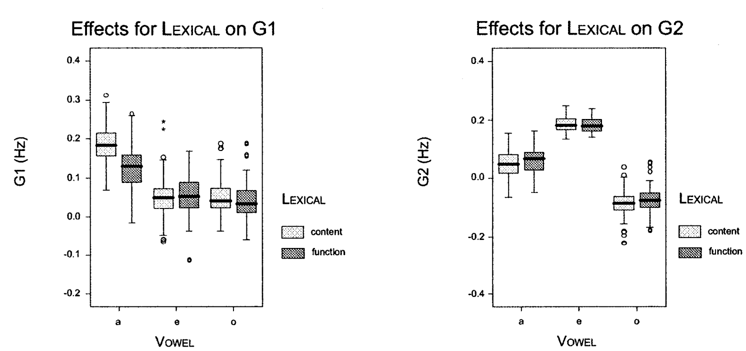

Figure 3.18 Boxplot illustrating G1 and G2 values

In Figure 3.18, the G2 median for [e] in function words was slightly less (i.e. more back) than that for [e] in content words, while the G2 median for [o] in function words was slightly greater (i.e. more front) than that for [o] in content words. G2 values in Figure 3.18 show clearer centralization than F2 values in Figure 3.8, although the magnitudes of centralization for [e] and [o] were small. The results of a one-way repeated measures ANOVA with AVERAGE G1 as dependent variable and Lexical (content and function) as independent variables indicate that there is a significant effect for Lexical $F(1,10) = 12.197$, $p = 0.006$. The results of a one way repeated measures ANOVA with AVERAGE G2 as dependent variable and Lexical (content and function) as independent variable indicate that there is no significant effect for Lexical $F(1, 10) = 3.232$, $p=0.102$. Next I will discuss individual vowels.

3.3.2.1 Individual vowels

The results of a two-way repeated-measures ANOVA with AVERAGE G1 as dependent variables, and LEXICAL (content, function) and VOWEL as the independent variables are summarized in Table 3.10.

Table 3.10 Effects for LEXICAL and VOWEL on AVERAGE G1

	VOWEL	LEXICAL	VOWEL x LEXICAL
df	[2, 20]	[1, 10]	[2, 20]
F	109.110	11.885	18.698
p	< 0.001	0.006	< 0.001

The results show that there is a significant effect for LEXICAL, for VOWEL and the interaction of two. Results of a two-way ANOVA with AVERAGE G2 as dependent variables, and LEXICAL (content, function) and VOWEL as the independent variables are summarized in Table 3.11.

Table 3.11 Effects for LEXICAL on AVERAGE G2

	Vowel	Lexical	Vowel x Lexical
df	[2, 20]	[1, 10]	[2, 20]
F	248.562	3.062	3.136
p	< 0.001	0.111	0.065

The results of the two way repeated measures ANOVA for G1 and G2 were the similar to the results for F1 and F2. There was a significant effect for VOWEL on G1 and G2, and there was a significant effect for LEXICAL on G1 but not on G2. The results of a repeated measures ANOVA with G1 or G2 as dependent variable and LEXICAL as independent variable is summarized in Table 3.12.

Table 3.12 Effects for LEXICAL on G1 and G2 split by VOWELS

		a	e	o
	df	[1, 65]	[1, 62]	[1, 65]
G1	*F*	62,559	0.169	2.285
	p	< 0.001	0.683	0.135
G2	*F*	8.703	4.131	2.503
	p	0.004	0.046	0.118

The results in Table 3.12 are similar to the results in Table 3.4. However, there was a significant effect for LEXICAL on /e/ G2 but not on /e/ F2. The next table shows the result of the Wilcoxon signed rank test for G1.

Table 3.13 Effects for LEXICAL on G1 (Wilcoxon signed-rank test)

VOWEL	Z	p	Σ negative ranks	Σ positive ranks
all	-4.842	< 0.001	13376	5734
a	-6.398	< 0.001	2107	104
e	-0.657	0.511	912	1104
o	-1.492	0.136	1339	872

The results of Table 3.13 are similar to the results of Table 3.5. There was a significant effect for Lexical on /a/ G1, but not on other vowels. Next table shows the lexical effect for G2 values by the Wilcoxon signed-rank test.

Table 3.14 Effects for LEXICAL on G2 (Wilcoxon signed-rank tests)

VOWEL	Z	p	Σ negative ranks	Σ positive ranks
all	-2.529	0.011	7559	11551
a	-2.986	0.003	638	1573
e	-2.102	0.036	1315	701
o	-2.341	0.019	739	1472

The Wilcoxon results show that there was a significant effect for LEXICAL on G2 for all vowels although there was not a significant effect for LEXICAL on F2 for /e/. The results also show that the direction of reduction for [e] was the opposite of that for [o].

3.3.2.2 Token difference

To illustrate token differences, means of G1 and G2 values are plotted in Figure 3.19, a scatter graph (horizontal axis: G2, vertical axis: G1).

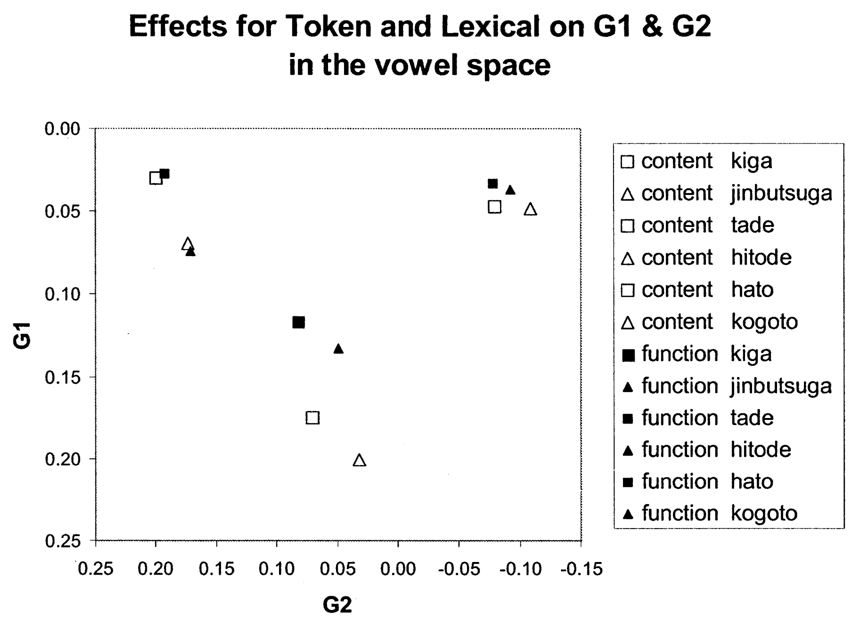

Figure 3.19 The graph illustrating difference lexical by token. Content: unfilled, Function: filled. The lexical effect on [a] was noticeable; the lexical difference was grater than token difference. However, token differences were greater than lexical differences for [e] and [o].

The G1 and G2 means in Figure 3.19 look similar to F1 and F2 means in Figure 3.9. The G1 and G2 means for [o] show clearer lexical difference.

3.3.3 Results of distance

The distances were calculated based on G1 and G2. As explained, Euclidean distances were used. The overall results indicate that the distances for function vowels were shorter than content vowels.

3.3.3.1 Overall results

The means and standard deviations of DISTANCE are provided in Table 3.15.

Table 3.15 Means and standard deviations of DISTANCE

VOWEL		all	content	function
all	DISTANCE (SD)	0.1621 (0.05892)	0.1715 (0.06026)	0.1527 (0.05613)
a	DISTANCE (SD)	0.1772 (0.05184)	0.2022 (0.04289)	0.1521 (0.04800)
e	DISTANCE (SD)	0.1991 (0.03188)	0.2001 (0.03334)	0.1974 (0.03052)
o	DISTANCE (SD)	0.1063 (0.04861)	0.1093 (0.05272)	0.1033 (0.04432)

The AVERAGE DISTANCES were submitted to a one-way repeated-measures ANOVA, whose results show that there was a significant effect for LEXICAL on AVERAGE DISTANCE, $F(1,34) = 5.953$, $p < 0.020$. The distances for function vowels (mean 0.1527) were shorter than those for the content vowels (mean 0.17115).

Next I discuss individual vowels.

3.3.3.2 Individual vowels

The results of a two-way repeated-measures ANOVA with AVERAGE DISTANCE as dependent variable and VOWEL and LEXICAL as independent variable are summarized in Table 3.16.

Table 3.16 Effects for VOWEL and LEXICAL on AVERAGE DISTANCE

	VOWEL	LEXICAL	VOWEL x LEXICAL
df	[2, 20]	[1, 10]	[2, 20]
F	53.920	10.189	24.472
p	< 0.001	0.010	< 0.001

The results indicate that there were significant effects for LEXICAL, VOWEL and LEXICAL x VOWEL. I investigated the effect for LEXICAL for individual vowels. The results of a one-way repeated-measure ANOVA with DISTANCE as a dependent variable and LEXICAL as an independent variable are summarized in Table 3.17.

Table 3.17 Effects for LEXICAL on DISTANCE split by VOWELS

	a	e	o
df	[1, 65]	[1, 62]	[1, 65]
F	59.379	0.885	0.898
p	< 0.001	0.350	0.347

Table 3.17 shows that vowel [a] contributes to the lexical difference of DISTANCE. The next table shows the results of the Wilcoxon signed-rank tests.

Table 3.18 Effects for LEXICAL on DISTANCE (Wilcoxon signed-rank test)

VOWEL	*Z*	*p*	Σ negative ranks	Σ positive ranks
all	-5.230	< 0.001	13682	5428
a	-6.244	< 0.001	2083	128
e	-0.472	0.637	1077	939
o	-1.326	0.185	1313	898

Table 3.18 shows that the results of the Wilcoxon signed-rank tests were similar to the results of the ANOVA. Only [a] contributes to the lexical difference on DISTANCE. The boxplot in Figure 3.20 illustrates the lexical differences for individual vowels.

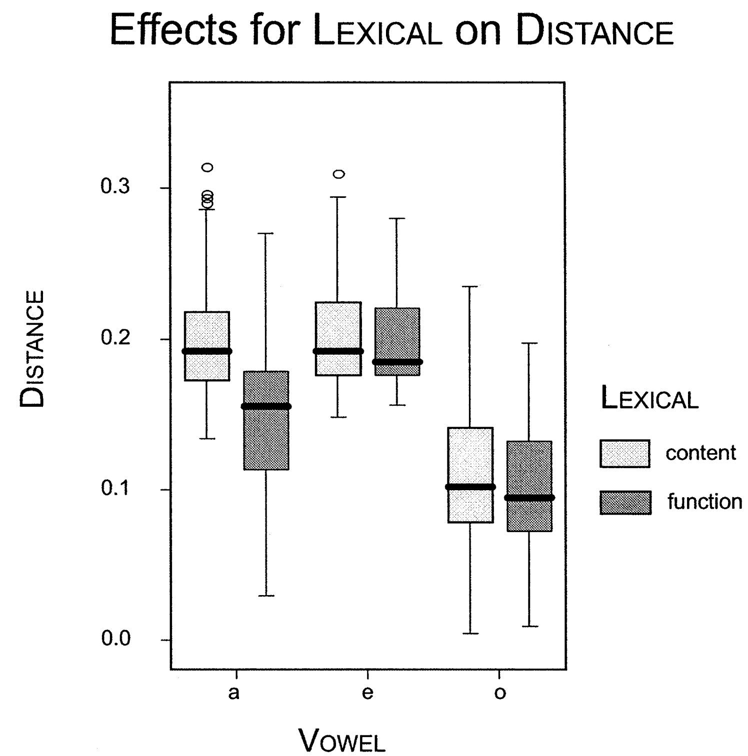

Figure 3.20 Boxplot illustrating the lexical differences of DISTANCES for individual vowels.

The boxplot in Figure 3.20 illustrates that, for all individual vowels, the median distances for function vowels were shorter than for content vowels although the lexical differences for [e] and [o] were subtle.

3.3.3.3 Token difference

To illustrate possible token difference, for each vowel a two-way repeated-measures ANOVA with AVERAGE DISTANCES as a dependent variable and TOKEN and LEXICAL as independent variables was performed. The results of the ANOVA are summarized in Table 3.19.

Table 3.19 Effects for TOKEN and LEXICAL on AVERAGE DISTANCE split by VOWELS

VOWEL		TOKEN	LEXICAL	LEXICAL x TOKEN
a	*df*	[1, 10]	[1, 10]	[1, 10]
	F	0.016	23.850	1.162
	p	0.902	0.001	0.306
e	*df*	[1, 9]	[1, 9]	[1, 9]
	F	3.899	0.476	0.044
	p	0.080	0.508	0.838
o	*df*	[1, 10]	[1, 10]	[1, 10]
	F	6.734	0.709	0.473
	p	0.027	0.419	0.507

Since onsets of target vowels were the same in both tokens, we expect similar AVERAGE DISTANCES for individual vowels in both tokens. However, overall results indicate that there was a significant effect for TOKEN on AVERAGE DISTANCE for [o] and a near significant effect for [e] but not for [a], and there was a significant effect for LEXICAL on AVERAGE DISTANCE for [a] but not for [e] and [o]. In addition, there was no significant interaction LEXICAL x TOKEN. The DISTANCES split by TOKEN and LEXICAL are plotted in Figure 3.21.

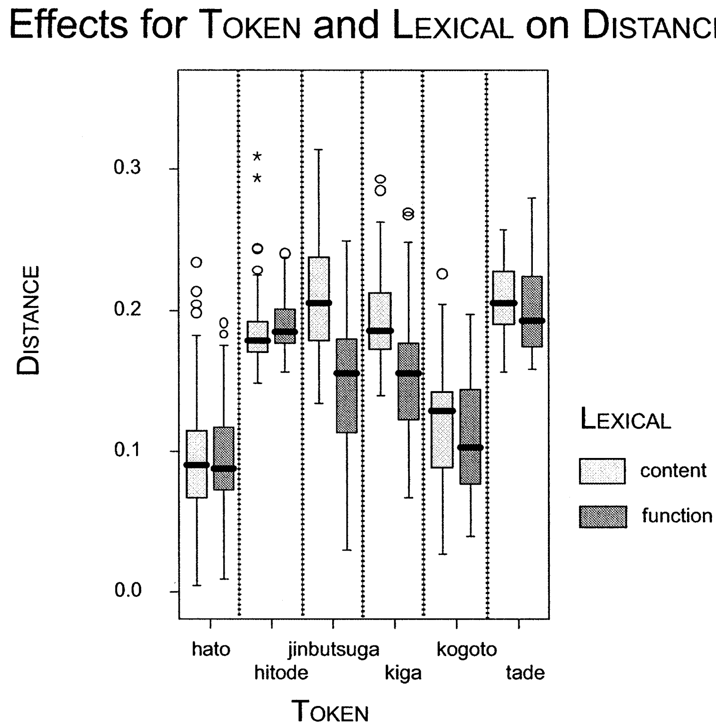

Figure 3.21 Box plot illustrating that distances from the center of the vowel space for function vowels and content vowels.

The boxplot of DISTANCE split by TOKEN shows that, with some exceptions, the median DISTANCE for function vowels was shorter than for content vowels. With the exception of the token, [hato], this boxplot is similar to the boxplot for DURATION. In Figure 3.21, the median distances for the function vowels of [hato] and [hitode] were slightly longer than those for content vowels although the mean distances for function in both tokens were shorter than for content.

3.4 Discussion

3.4.1 Coarticulation effects

Here I will summarize the results which showed the significant formant differences between content and function. First, F1 values of function [a] were lower than while F2 values were higher in function counterparts. Second, G2 values of function /e/ were significantly lower than content /e/. In addition, F2 values of function [e] showed a similar trend towards lower values. When I designed this experiment, I focused on the centralization. However, the centralization appears to be due to robust coarticulation effects.

Function [a] was significantly shorter than content [a], so the coarticulation effect more intensely influenced function [a] than content [a]. The onset of [a] in [dʒinbɯtsɯga̲ to] and [kiga̲ ki] was velar, the velar consonant, [g], lowers F1 values but raises F2 values. Thus, it is reasonable to interpret the lowered F1 and the raised F2 as resulting coarticulation effects.

Unexpected result was that there was not a significant effect for LEXICAL on F2 values of [o], although the duration of function was significantly shorter than content counterparts. A possible reason for this unexpected result is as follows. The onset of [o] was alveolar, which has been known as little coarticulation effects on F2 values; in other words, locus target distance between alveolar consonants and vowels are short. The research conducted by Van Son and Pols (1992) had the similar results. In their research, duration and formants of vowels in alveolar–vowel–alveolar with two reading speed were measured. The results showed that there was a significant effect for READING SPEED on F1 values but no significant effect on F2 values. The results of their research also showed a little coarticulation effect of alveolar on F2 values. In addition to the coarticulation effects by the preceding consonants (i.e. onsets), following consonants, preceding and following vowels also affect the formants of the target vowels. Thus, the unexpected results were probably due to the combination of these coarticulation effects.

The results of the Wilcoxon signed-rank test showed that the F2 values of function [o] were frequently greater than content counterparts. On one hand, ANOVA tests compares the variation of F2 values within groups with that between groups so the direction (less or greater) is not important. On the other hand, Wilcoxon signed-rank tests compare mainly frequencies of directions (content is less or function is less). The results of the ANOVA test for F2 [o] indicate that the F2 difference between groups (content vs. function) was not significantly larger than the differences within groups (difference in content group and that in function group). The results of the Wilcoxon test for F2 [o] indicate that content G2 values for [o] were commonly greater than function ones, and content F2 for [o] were commonly less than function ones. So the results of the Wilcoxon test for [o] indicate that it is possible that the effect for LEXICAL would be observed if the number of token were large enough.

In addition, Japanese is a mora-timed language, so the duration of each mora is supposed to the same. Consequently, the magnitude of vowel reduction in Japanese is probably weaker than those in other languages.

The other possible reason comes from the different level of redundancy in Japanese function particles. Although some Japanese function particles can be omitted, not all function particles can be omitted. The subject marker [ga], the object marker [o] and the topic marker [wa] can be omitted but [de] and [to] in the contexts of my study cannot. So [de] and [to] should be pronounced more clearly than the subject marker [ga], which possibly led to no lexical effect of [e] and [o]. There may be other reasons for this unexpected result of F2 [o].

3.4.2 Using three methods to test lexical effects

In this chapter, I used three methods to test the centralization. I discuss merits and demerits of each method. First, I used F1 and F2 values. F1 and F2 directly show the effects for LEXICAL. However, the range of F2 is almost twice as that of F1 so it is not appropriate to compare F1 with F2. Second, I used G1 and G2. Since each person has different range of inherent formants, normalization is necessary to compare formants produced by one subject to those by another. In addition, using G1 and G2, I could compare F1 and F2. Since people have non-linear auditory system, it is dangerous to discuss the difference of vowels based on raw formants. Differences in F1 seem to smaller than those in F2 based on raw formants. Or difference in /e/ F2 seems larger than difference in /o/ F2. So using G1 and G2, we can compare the differences fairly. The results of a repeated measures ANOVA with G1 and G2 are similar to the results with F1 and F2. It indicates that it is not the case that some subjects with high voices (outliers) led to the results. Finally I used Distance. The range of differences in G1 is similar to the range of in G2. If I use F1 and F2 to calculate the distance, then F2 difference affects more than F1 difference. The calculated DISTANCES based on G1 and G2 reflect the human auditory system. This distance is useful for comparing two vowels whose F1 and F2 are totally different. For example, I can compare [ɛ] distance with [ɔ] distance, etc.

3.4.3 Token differences: Coarticulation effects

I summarize the effects for TOKEN on F1 and F2. There were significant effects for TOKEN on F1 for [a] and [e] and on F2 for all vowels. In token pairs, the preceding consonants (i.e. onsets) were the same in both tokens. However, the following consonant, the preceding and following vowels were different. These differences caused the significant token differences. Without controlling each phoneme, I could not discuss the individual coarticulation effects.

3.4.4 Lexical effects?

The fact that Japanese function [a] was shorter than content vowels, and that function [a] was centralized raises another question: does spectral vowel reduction depend on the duration? In other words, is there lexical effect in the absence of durational difference? I will test whether or not spectral vowel reduction is the result of durational vowel reduction.

4 Vowel reduction and undershoot in Japanese

In Chapters 2 and 3, I showed that function [a] was reduced in durational and spectral dimensions. Both facts raise the following question: is spectral vowel reduction due to durational reduction? In this chapter, I will discuss the relation between durational reduction and spectral reduction (in terms of both [a] F1 and G1). First, I will briefly reintroduce previous studies.

4.1 Background

Lindblom (1963) hypothesized that formant displacement was a function of duration; in other words, duration was the sole determinant of formant displacements based on the results of F2 values in stressed vowels and unstressed vowels. I provide a graph in Figure 4.1 to show Lindblom's hypothesis.

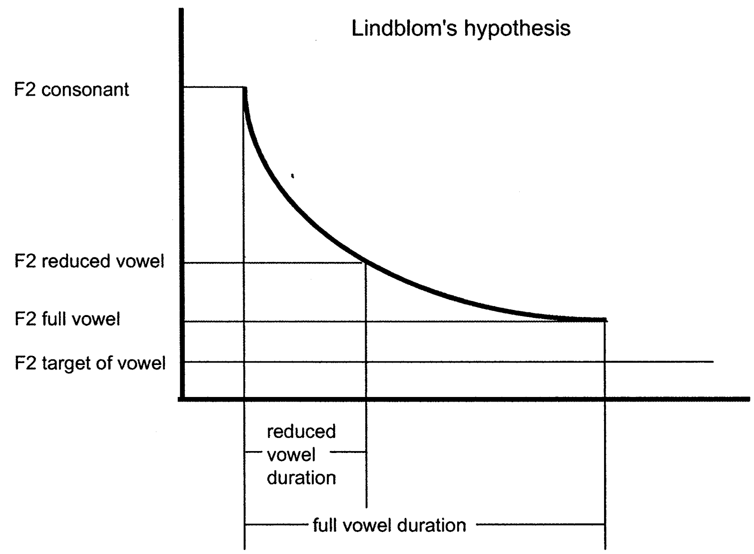

Figure 4.1 Graph shows Lindblom's hypothesis, and depicts the transition from consonant F2 to vowel F2. When a vowel is long enough, F2 reaches its target. Since a reduced vowel is short, F2 reduced vowel is far from the F2 target.

In Lindblom's hypothesis, there is only one transition from consonant F2 to vowel F2. When the duration of a vowel is long enough, then vowel F2 reaches its target. Because a reduced vowel is shorter than a full vowel, the F2 value of a reduced vowel moves less than the F2 value of a full vowel. This short movement of a reduced vowel results in vowel centralization.

Nord (1986) counter-argued that duration was not the sole determinant of formant displacements, using initial stressed vowels and final unstressed vowels in Swedish.

Van Bergem (1993) argued that vowel reduction was due to increased contextual assimilation based on F2 formant tracks. He had three models for the difference of the F2 formant tracks between reduced vowels and full vowels. His three models are reproduced in Figure 4.2.

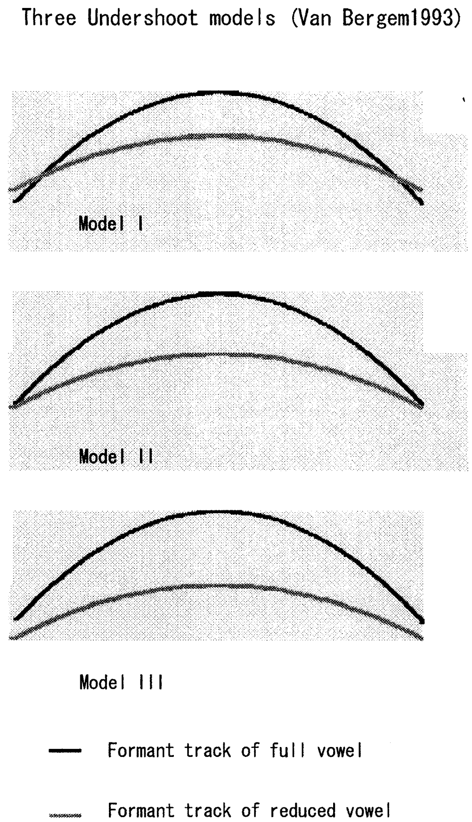

Figure 4.2 Van Bergem's three undershoot models.

In Model I, the F2 values at the onsets and the offsets of reduced vowels are closer (greater in the graph) to the targets than those of full vowels. In Model II, the F2 values of reduced vowels are the same as those of full vowels. In Model III, the F2 values of reduced vowels are farther from (less than) the targets than those of full vowels. The results of Van Bergem's research showed that the F2 values of reduced vowels at onset and offset were less than those of isolated vowels, which was Model III in the graph.

Moon and Lindblom (1994) argued that formant displacement depended on duration and the speaking style, comparing F2 values of front vowels in [w_l] contexts in citation speech and clear speech with those in [h_d] context. Their results seemed to indicate that duration was a determinant of formant displacement. However, it was not clear whether or not there was another factor. In my hypothesis, there is an independent lexical effect, in addition to the durational (i.e. undershoot) effect.

4.2 Hypothesis 3:

If durational reduction is the source of spectral reduction, i.e. vowel undershoot (Lindblom 1963), then differences in F1 values will be linked to differences in duration. In other words, there will be a significant correlation between duration and differences in formant values. If, on the other hand, lexical difference is the source of spectral reduction and independent from durational reduction, then differences in F1 values will be observed in the absence of differences in duration.

4.3 Materials

To test whether or not spectral vowel reduction is independent from durational reduction, I investigated correlations between duration and measurements of formants, formant trajectories, categorical durations and changes in formants. Since F1 and G1 of [a] show a strong lexical difference, I will focus on the F1 and G1 values for [a]. In this chapter, I use all measurements of [a] F1 and G1 since I am interested in the relation between DURATION and [a] F1 or G1.

4.3.1 Pearson's correlation coefficient

In statistics, the correlation coefficient (Kirk 1999) is used to test the relation between two variables. Pearson's correlation coefficient, which is signified by *r* (*rho*), is often used to measure the strength of the linear relationship between two variables. The range of Pearson's *r* is from –1 to 1. If Pearson's *r* is 1, there is a perfect positive correlation between two variables, such as x and y; all points expressed with [x, y] fall on a straight line from the bottom left corner to the top right corner in the graph. On the other hand, if Pearson's r is –1, there is a perfect negative correlation; all points fall on a straight line from the top left corner to the bottom right corner. If Pearson's r is 0, there is no correlation, and points are scattered in the graph. The p value shows the significance of the correlation. When p is less than 0.05, it is considered that there is a statistically significant correlation between two variables.

4.3.2 Formant trajectories

The coarticulatory effect of a consonant constriction on a neighboring vowel results in a displacement of the vowel's formants, commonly known as formant transitions. I have measured formants at five points to draw trajectories: onset, 1/4, mid-point, 3/4 and offset. When two vowels (content and function) occur in the same context, it may be assumed that differences in trajectories of formants for vowels (content and function) are dependent on duration. These trajectories show whether or not the change in formants is dependent on the duration.

Van Bergem tested whether the onsets and offsets of reduced vowels coincide with those of full vowels. The results showed that the F2 values at the onsets and offsets of reduced vowels were closer to the F2 values of the consonants. He concluded that spectral vowel reduction was due to increased contextual assimilation. His methods will be applied to the current study below. Furthermore, I will test whether F1 trajectory of content [a] is parallel to that of function [a].

4.3.3 Categorical duration

In Lindblom's hypothesis, duration was the sole determinant of formant displacement. In my hypothesis, lexical difference would also contribute to formant displacement to some degree. In order to test for possible difference between content and function formants of approximately the same duration, I used categorical duration. The measured durations of /a/ in both content and function words were arranged in ascending order, and then classified into three categories: short, middle and long. Each category contains the same number of tokens, but the number of tokens for the content category was not equal to that for the function category.

4.3.4 Δ Formants

Van Bergem observed that F2 values at the onsets of reduced vowels were farther from the target than those of full vowels. This raises the question whether formant trajectories of reduced vowels are parallel to those of full vowels. There are three possibilities: a. Since formants of function vowels at onset are further from the target than those of content vowels, then formants in reduced vowels move faster to reach the target than do those in full vowels. b. Formants of reduced vowels move in the same way as full vowels, and then the formant track of the reduced vowels is parallel to that of full vowels. c. The target is different for reduced vowels so formants of reduced vowels move slower than those in full vowels.

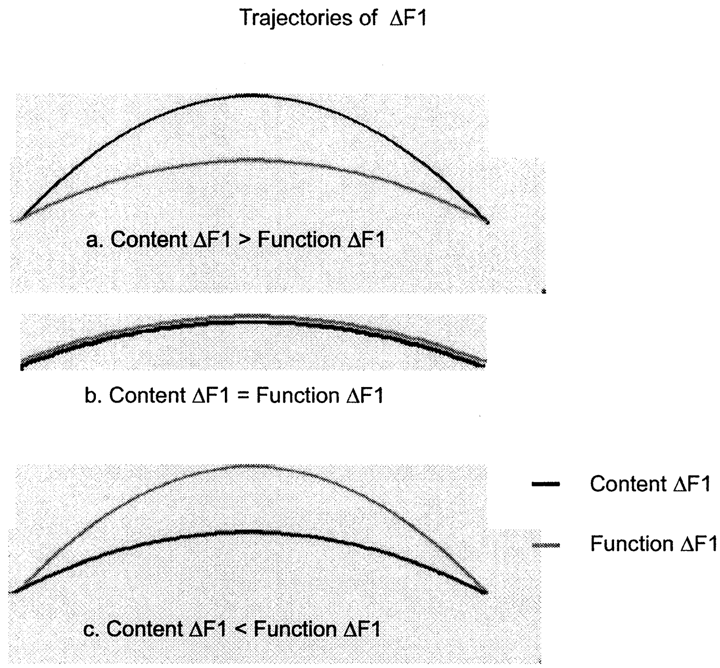

Figure 4.3 Three models for ΔF1 trajectories

In order to determine whether or not formant trajectories of reduced vowels are parallel to those of full vowels, Δformants was used. The equation of Δformants is as follows:

$$\Delta F1 = F1_{onset} - F1_{each\ point} \quad (4\text{-}1)$$

$$\Delta G1 = G1_{onset} - G1_{each\ point} \quad (4\text{-}2)$$

The scatter graph in Figure 4.4 illustrates the calculation of ΔF1 at mid-point. Based on the equation, ΔF1 at onset is always 0.

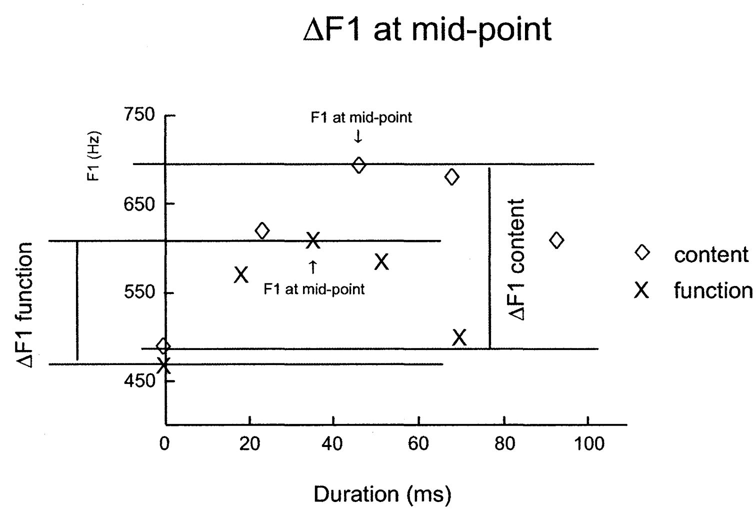

Figure 4.4 Scatter graph illustrating the calculation of ΔF1. This graph shows how to calculate ΔF1 at mid-point although ΔF1 at all five points are calculated.

The ΔF1 at five points is used for formant trajectories. In addition, ΔF1 at mid-point is used with the factor CATEGORICAL DURATION to observe the lexical difference. The ΔG1 at five points is also used for formant trajectories.

4.4 Results 3: Lexical vowel reduction

The results supported the vowel undershoot hypothesis. The calculated Pearson's *r* indicates that there were significant correlations between duration and F1 for [a]. However, the overall results of other tests indicate that there was a significant effect for LEXICAL on F1 for [a] in the absence of durational difference. First, I will discuss the correlation.

4.4.1 Correlation

The overall results indicate that for [a], there was a significant correlation between duration and formants.

4.4.1.1 Correlations between DURATION and FORMANTS

The results of the correlation between DURATION and F1 or G1 are summarized in Table 4.1.

Table 4.1 Correlation between DURATION and FORMANTS for [a]

		all	content	function
F1	*r*	0.363	0.183	0.315
	p	< 0.001	0.141	0.010
G1	*r*	0.535	0.420	0.448
	p	< 0.001	< 0.001	< 0.001

There were significant correlations between DURATION and G1 of [a]. However, Pearson's *r* for F1 of content [a] does not equal that for F1 of function [a]. The following scatter graph in Figure 4.5 illustrates DURATIONS and F1 values for [a].

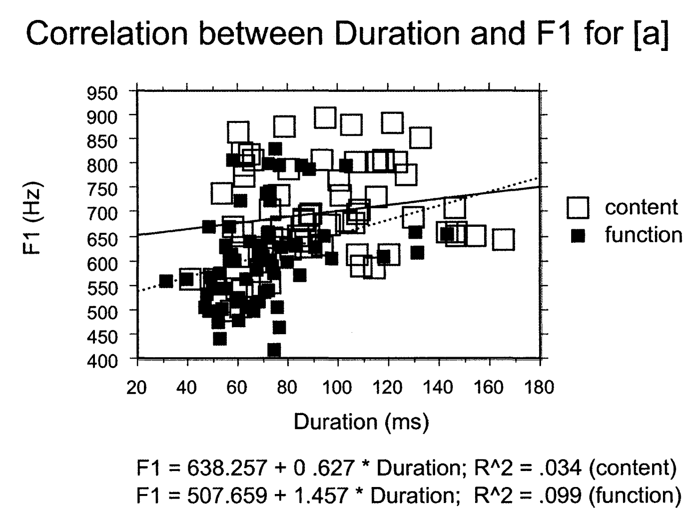

Figure 4.5 Scatter graph illustrating correlation between duration (horizontal axis) and F1 (vertical axis) for [a]. Unfilled: content, filled: function.

The scatter graph in Figure 4.5 illustrates the correlation between DURATION and F1 values for [a]. The points are diagonally distributed from the bottom left corner to the top right corner, indicating there was a positive correlation. The graph also shows that function [a]s form a group and content [a]s form another group. In addition, the durations of function [a]s were shorter, and F1 values of function [a]s were lower than those of content [a]s. These facts indicate that there was lexical difference. Furthermore, the F1 values of content [a]s were greater than those of function [a]s where the duration was short. Next I will introduce a scatter graph of strong correlations between duration and normalized formants (G1). The scatter graph in Figure 4.6 illustrates the strong positive correlation between DURATION and G1 for [a].

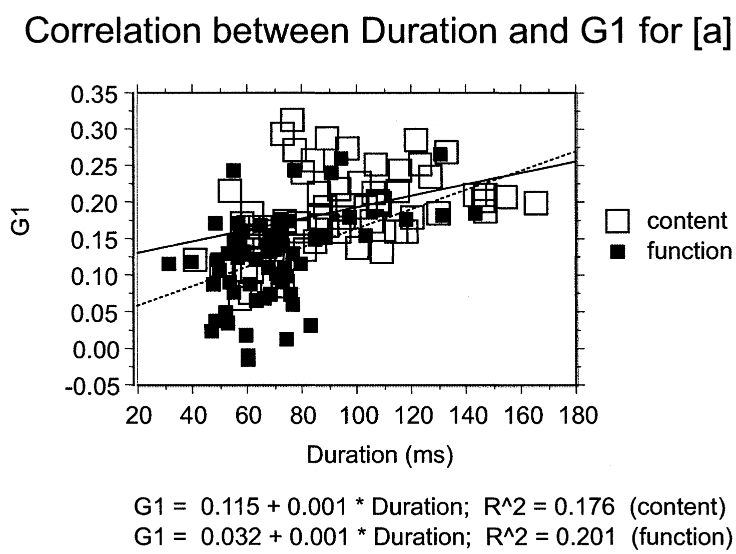

Figure 4.6 Scatter graph illustrating duration and G1 for [a]. Unfilled: content, filled: function.

The scatter graph in Figure 4.6 shows a diagonal distribution of points, which indicates a positive correlation between DURATION and G1. This graph also shows the difference between content [a]s and function [a]s. The content [a]s (unfilled) were densely concentrated in the region corresponding to long duration and high G1 while the function [a]s were densely concentrated in the region corresponding to short duration and low G1.

The results of the correlation and scatter graphs show that there were strong correlations between duration and F1 (and G1) values for [a]. In addition, there was a noticeable lexical difference for F1 (and G1) values of [a].

4.4.2 F1

4.4.2.1 F1 Trajectories

Van Bergem (1993) showed that the F2 values at the onsets of reduced vowels were closer to the F2 values of the preceding consonant than those of full vowels, and concluded that spectral vowel reduction was due to increased contextual assimilation. To replicate his research and to look at the potential role of a lexical effect in formant movement throughout the vowel, trajectories that consisted of five points of F1 of [a] were used. If DURATION is the only determinant of the formant trajectory, the trajectory of content vowels would be overlapped with the trajectory of function vowels at least at onset and offset. I calculated the F1 averages and DURATIONS at five points for content [a] and function [a] separately. The results are summarized in Table 4.2 and plotted in Figure 4.7.

Table 4.2 Mean F1 and DURATION at five points

LEXICAL		0	1/4	2/4	3/4	end
content	DURATION	0	23.11	46.22	69.33	92.44
	F1	485	622	696	686	607
function	DURATION	0	17.54	35.07	52.61	70.14
	F1	468	561	610	590	495

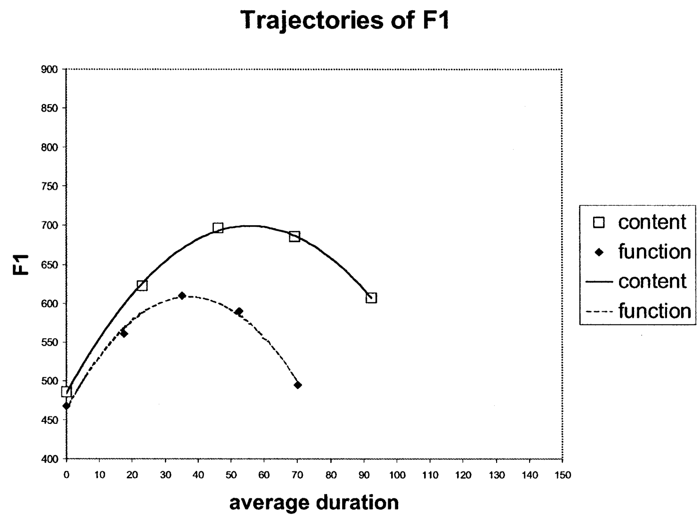

Figure 4.7 The trajectories of average F1 of [a]. The points are calculated based on the average durations of content [a] and function [a]. Unfilled: content, filled: function, solid line: content, dashed line: function.

Figure 4.10 shows that the average F1 of content [a] was greater than the average F1 of function [a] in all points. F1s of content [a]s at the beginning and at the end of vowels do not coincide with F1s of function [a]s; this accords with Van Bergem's (1993) finding for Dutch.

4.4.2.2 Calculated F1 values of function [a] with the equation of content [a]

In Lindblom's hypothesis, formant displacement was a function of DURATION. So I calculated the following equations based on DURATION and F1 values using a function provided in Microsoft Excel.

Content: $F1 = -0.0691(duration)^2 + 7.7077 * duration + 484.37$ **(4-3)**
Function: $F1 = -0.1034(duration)^2 + 7.736 * duration + 464.14$ **(4-4)**

These equations indicate that F1 values of function [a]s were not equal to those of content [a]s even when the durations of function [a]s were longer. I used the equation for content in (4-3) to calculate F1 values of function [a]; in other words, I inserted the duration of function vowels in (4-3). The results are summarized in Table 4.3.

Table 4.3 Comparison of calculated and measured F1 of function [a]

	0	1/4	2/4	3/4	end
DURATION (ms)	0	17.54	35.07	52.61	70.14
calculated F1	484	598	670	699	685
measured F1	468	561	610	590	495

There were differences between calculated F1 values and measured F1 values. At the mid-point, the difference was 60 Hz, 10 % of the F1 value. The scatter graph in Figure 4.8 illustrates the differences between calculated and measured F1 values.

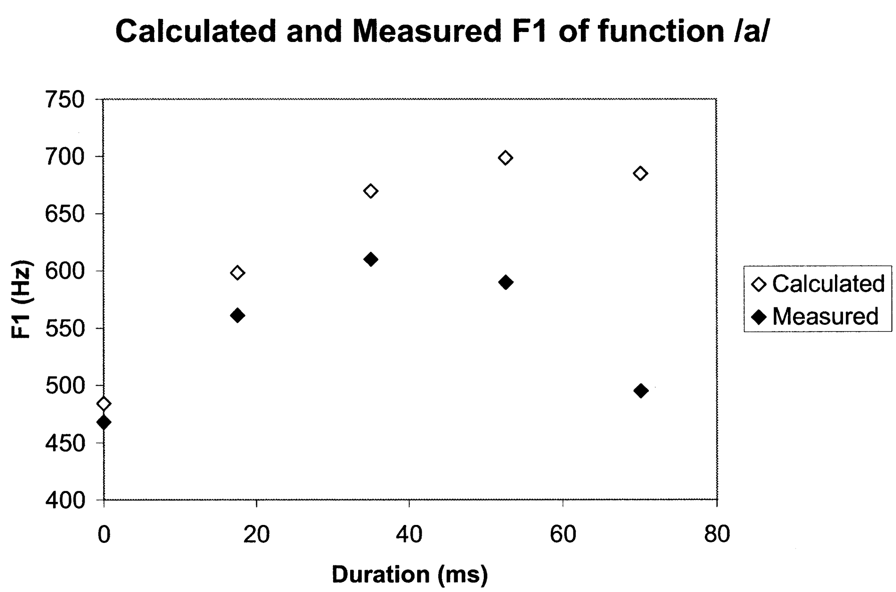

Figure 4.8 Scatter graph illustrating calculated F1 of function [a] based on the duration of function. Unfilled: calculated, filled: measured

The scatter graph in Figure 4.8 shows that the calculated F1 values were always greater than the measured F1 values in all points. If formant displacement were the function of duration, the calculated F1 should be equal to the measured F1. Thus, the formant displacement is not a function of duration. There is another factor involved, lexically motivated reduction.

4.4.2.3 F1 trajectories of [a] by individual subjects

Since the F1 averages and the duration averages were calculated over all subjects, the graphs might be over-generalized. The next scatter graph in Figure 4.9 illustrates the F1 average of [a]s produced by individual speakers.

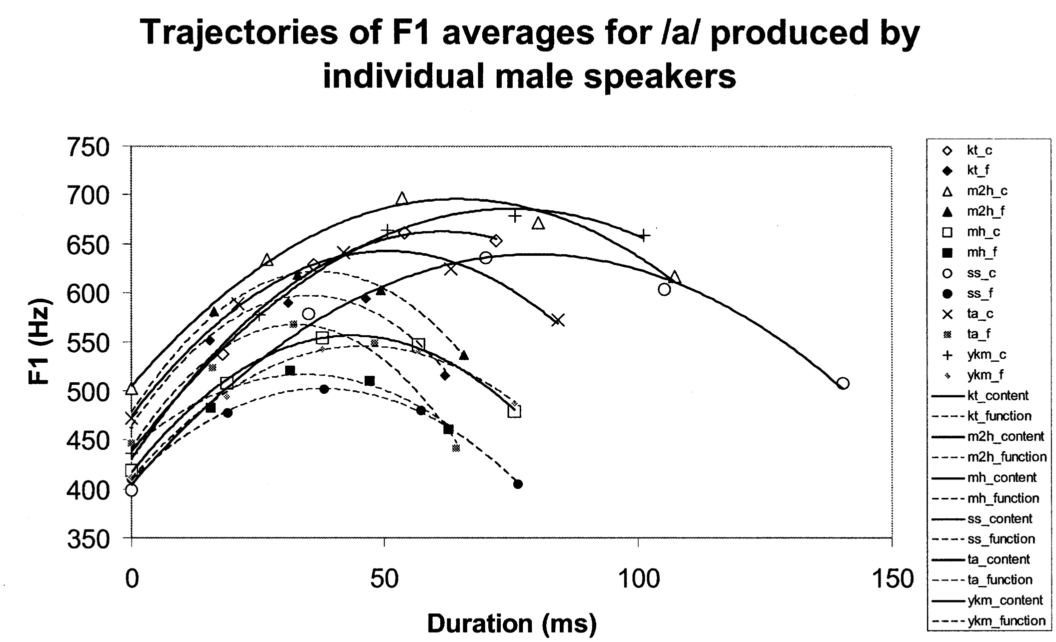

Figure 4.9 Scatter graph illustrating trajectories of F1 averages produced by individual male speakers. Solid lines: content, dashed lines: function, unfilled: content, filled: function

The graph in Figure 4.9 shows that the trajectories of content [a]s are more stretched than those of function [a]s. The trajectory of content [a] was almost always above the counter-part of function [a]. The next graph in Figure 4.10 illustrates the F1 averages of [a]s produced by individual female speakers.

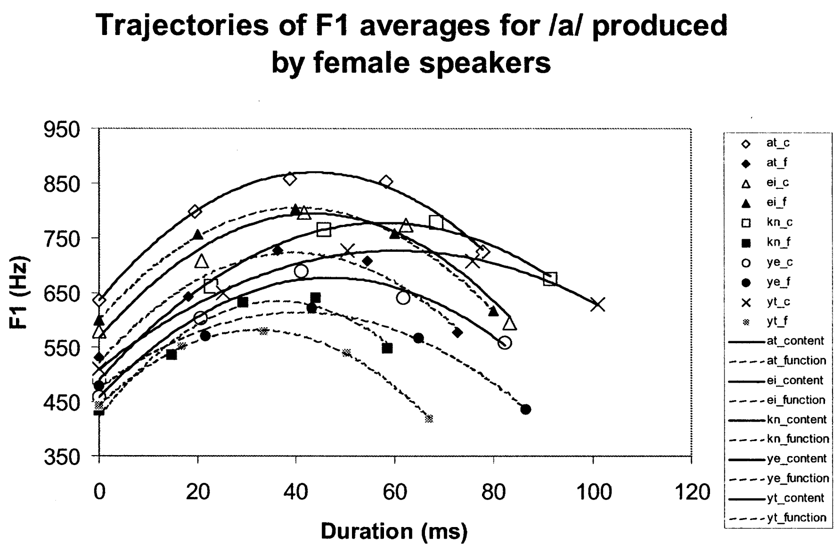

Figure 4.10 Scatter graph illustrating trajectories of F1 average produced by individual female speakers.

The scatter graph in Figure 4.10 shows that, with exceptions, the trajectories for content vowels were above those for function vowels. Both male and female speakers'[1] trajectories for content /a/ were consistently above the trajectories for function [a]. The only exception was the trajectory of content [a] produced by the subject "ei", which was below the counterpart of function [a] most of time. These graphs show that DURATION was not the only determinant of FORMANTS. In other words, lexical difference contributes to the formant displacement.

[1] The F1 values of content [a] produced by six subjects were greater than their counterpart of function [a] for all five points. The three other subjects showed the same tendency with the exception of F1 values at onset.

4.4.2.4 F1 average values of CATEGORICAL DURATION

To test whether or not there was a lexical difference in the absence of the durational difference, durations were classified[2] into 3 categories so that the F1 values at vowel mid-point could be compared. Because function vowels were short, the number of tokens for function in the short category was greater than that for content, while that for function in the long category was less than that for content. The F1 averages of individual categories are summarized in Table 4.4.

Table 4.4 Mean F1 in three CATEGORICAL DURATIONS

LEXICAL	CATEGORICAL DURATION	short	middle	long
content	# of tokens	16	15	35
	F1	675	675	715
function	# of tokens	28	29	9
	F1	576	624	669

The mid-point F1 values in individual CATEGORICAL DURATIONS were submitted to a factorial ANOVA. The results of a factorial ANOVA with F1 as a dependent variable, and CATEGORIZED DURATION and LEXICAL as independent variables show that there was a significant effect for LEXICAL on F1 of [a], $F(1, 126) = 12.639, p = 0.001$, and a significant effect for DURATION, $F(2,126) = 4.232, p = 0.017$, but no significant INTERACTION between the two, $F(2, 126) = 0.889, p = 0.413$. This means that while function [a] is more reduced than content [a], the reduction is not solely tied to duration. However, the general trend in the graph is for shorter durations to have lower F1 values, which means that there is a

[2] A function of SPSS, "categorize" transforms continuous variables into nominal variables based on percentile distribution. The duration is classified into 3 categories so each category contains 44 measurements.

significant effect for DURATION. The error bar graph in Figure 4.11 illustrates the interaction between F1 values and CATEGORICAL DURATION.

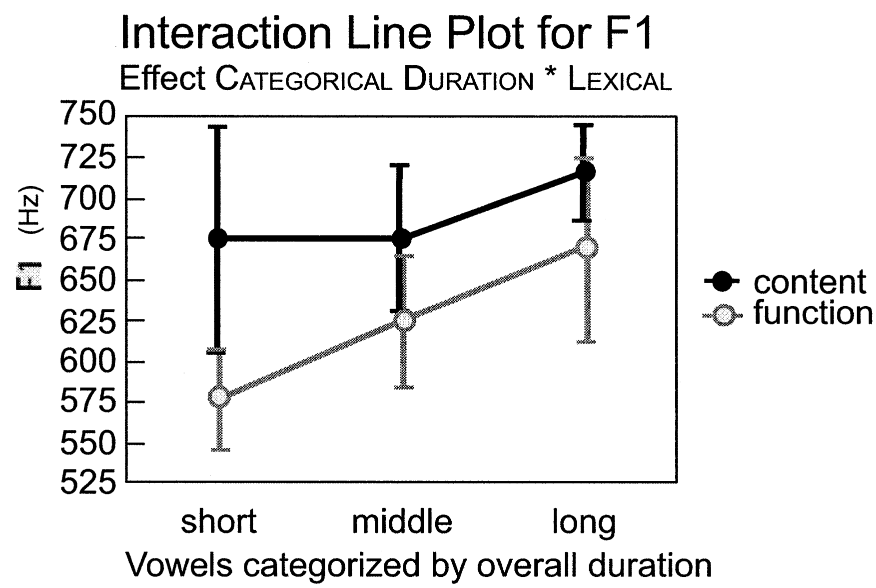

Figure 4.11 Error bar graph illustrates the interaction between LEXICAL and CATEGORICAL DURATION. Black: content, gray: function

Figure 4.11 shows that F1 values of content [a] are greater than those of function [a] in all three categories. Thus formant displacement is not tied to duration. If centralization is related to reduction more than undershoot, we expect there to be a different target for function [a] from content [a]. Next I present the results for the analysis of Δ F1s.

4.4.3 Δ F1

The trajectories of F1 show that the onsets of F1 values for content [a]s were greater than those for function [a]s. So to investigate the changes in F1 values, F1 values were converted to Δ F1. These Δ F1 values are taken to indicate how much the tongue moved.

4.4.3.1 Trajectories of ΔF1

The ΔF1 values are summarized in Table 4.5.

Table 4.5 Mean Δ F1 and DURATION at five points

	average ΔF1					average duration
LEXICAL	0	1/4	2/4	3/4	end	
content	0	137	211	200	121	92.44
function	0	93	142	123	28	70.14

Table 4.5 shows that the average ΔF1 of content [a] was greater than that of function [a] at all points. The ΔF1 values are plotted in Figure 4.12.

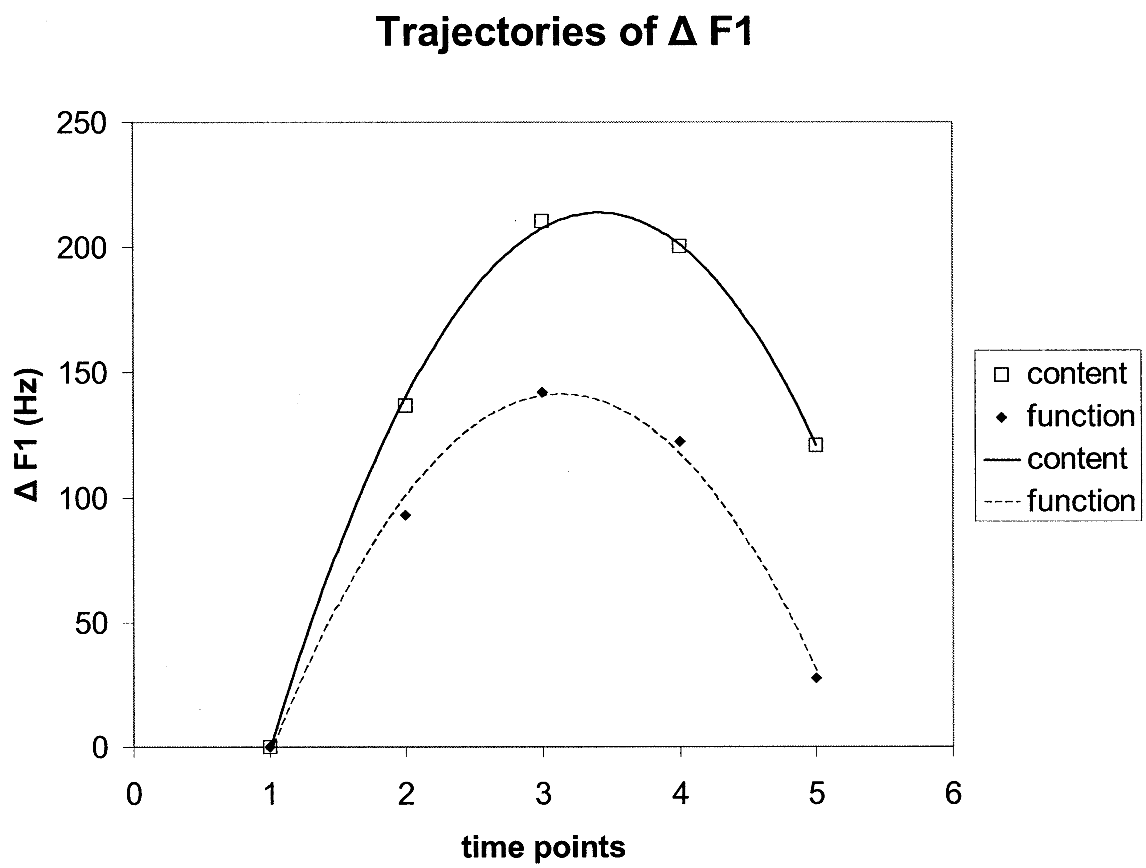

Figure 4.12 The scatter graph illustrating the trajectories of Δ F1. Unfilled: content, filled: function, solid line: content, dashed line: function.

Figure 4.12 shows that the slope of the trajectory for content Δ F1 at the beginning of the vowel is different from that of function Δ F1, and that the target of Δ F1 for function [a] seems to be different from that for content [a]. This graph clearly shows that duration is not the only determinant of formants. The graph shows that the average F1 of a function vowel does not reach the average F1 of a content vowel even when the duration of the function vowel is longer.

4.4.3.2 ΔF1 and CATEGORICAL DURATION

Again, the durations were classified into 3 categories (short, middle and long). If undershoot is responsible for F1 centralization in function [a], then different Δ F1 values should be restricted to different durations. If centralization is caused by reduction, then different Δ F1 values should be seen at the same duration. The averages of Δ F1 for individual categorical durations are summarized in Table 4.6.

Table 4.6 Mean ΔF1 in three CATEGORICAL DURATIONS

LEXICAL	CATEGORICAL DURATION	short	middle	long
content	# of tokens	16	15	35
	ΔF1	159	186	245
function	# of tokens	28	29	9
	ΔF1	110	157	195

Table 4.6 shows that the average Δ F1s for content [a]s were greater than those for function [a]s in all three categories. The average Δ F1s are plotted in Figure 4.13.

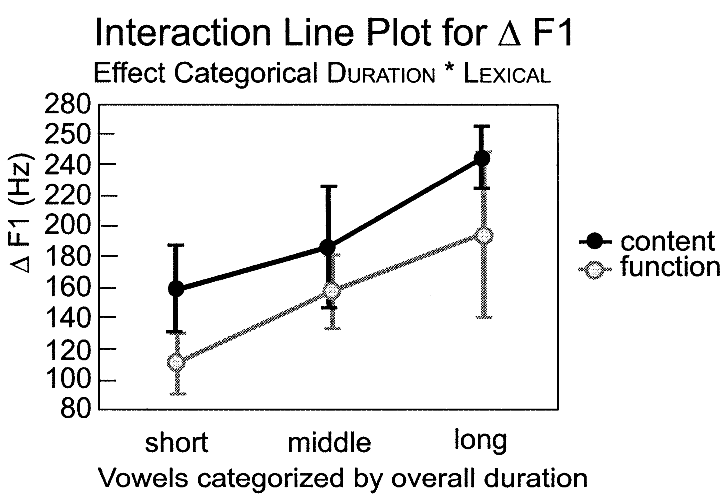

Figure 4.13 Error bar graph illustrating the interaction between LEXICAL and CATEGORICAL DURATION.

ΔF1 values for content [a] were greater than ΔF1 for function [a] in all three categorical durations. These lines were almost parallel, which indicates that ΔF1 is not tied to duration.

The results of a factorial ANOVA with ΔF1 as the dependent variable, and CATEGORIZED DURATION and LEXICAL as independent variables show that there was a significant effect for LEXICAL on F1 of [a], $F[1, 126] = 13.236, p < 0.001$, and a significant effect for DURATION, $F(2, 126) = 16.568$ $p < 0.001$. However, there was no significant INTERACTION, $F[2, 126] = 0.368$ $p = 0.693$. This means that ΔF1 is not tied to duration. Next I will report the results of the tests for LEXICAL effects that usedG1 as the dependent variable.

4.4.4 G1

4.4.4.1 G1 trajectories

I calculated five points along the trajectory G1. The results are summarized in Table 4.7.

Table 4.7 Mean G1 and DURATION at five points

LEXICAL		0	1/4	2/4	3/4	end
content	DURATION	0	23.11	46.22	69.33	92.44
	G1	0.0474	0.1459	0.1878	0.2147	0.1503
function	DURATION	0	17.54	35.07	52.61	70.14
	G1	0.0260	0.0985	0.1404	0.1673	0.1029

Table 4.7 shows that the average of content G1 for [a] was greater than that of the function counterparts at all points. These averages of G1 are plotted in Figure 4.14.

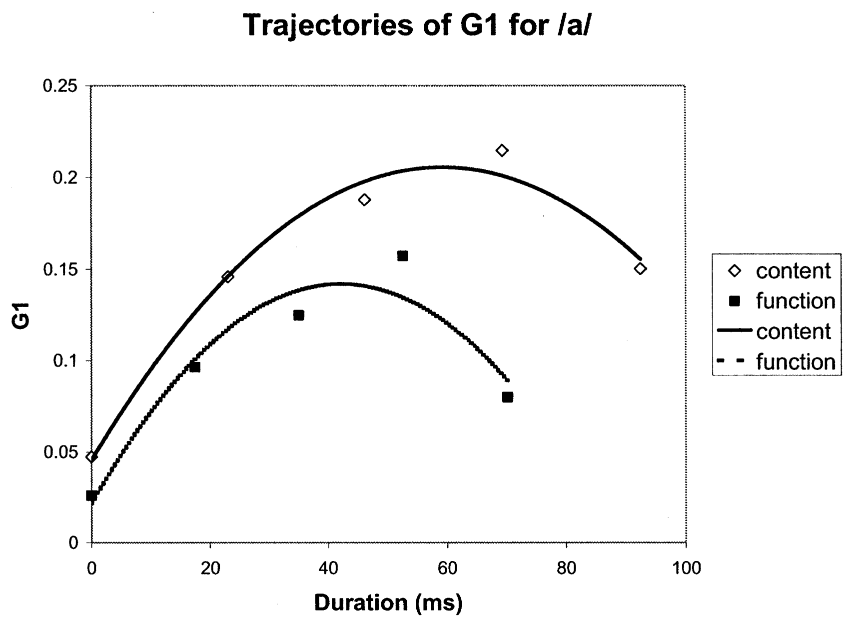

Figure 4.14 The trajectories of average G1 of [a]. The points are calculated based on the average durations of content [a] and function [a]. Unfilled: content, filled: function, solid line: content, dashed-line: function.

The trajectories of average G1 for [a] are similar to those for average F1 of [a]. The content G1 values at the onsets of vowels were higher than the function counterparts. These results indicate that the F1 trajectories were not skewed by outliers because G1 is normalized F1. Using an Excel function, I provide the equations for content and function.

Content: $G1 = 0.00005(duration)^2 + 0.0054 * duration + 0.0456$ **(4-5)**
Function: $G1 = 0.00007(duration)^2 + 0.0057 * duration + 0.0215$ **(4-6)**

These equations indicate that G1 values of function [a] were not equal to those of function counterparts. Since G1 values were normalized based on the differences among subjects, I will not introduce individual subjects' trajectories.

4.4.4.2 G1 average values of CATEGORICAL DURATION

To test whether or not there was a lexical difference in the absence of the durational difference, duration was classified into 3 categories so that the G1 values at vowel mid-point could be compared. The G1 averages of individual categories are summarized in Table 4. 8.

Table 4. 8 Mean G1 in three CATEGORICAL DURATIONS

LEXICAL	CATEGORICAL DURATION	short	middle	long
content	# of tokens	16	15	35
	F1	0.1399	0.1963	0.2060
function	# of tokens	28	29	9
	G1	0.1035	0.1218	0.2003

The mid-point G1 values in individual CATEGORICAL DURATIONS were submitted to a factorial ANOVA. The results of the factorial ANOVA with G1 as dependent variable, and CATEGORIZED DURATION and LEXICAL as independent variables show that there was a significant effect for LEXICAL on G1 of [a], $F(1, 126) = 16.017, p < 0.001$, a significant effect for DURATION, $F(2, 126) = 22.323, p < 0.001$, and a significant interaction between the two $F(2, 126) = 4.030, p = 0.020$. Although the INTERACTION between LEXICAL and CATEGORICAL DURATION on F1 was not significant, the INTERACTION on G1 was significant. The next error bar graph in Figure 4.15 illustrates the interaction between G1 and CATEGORICAL DURATION.

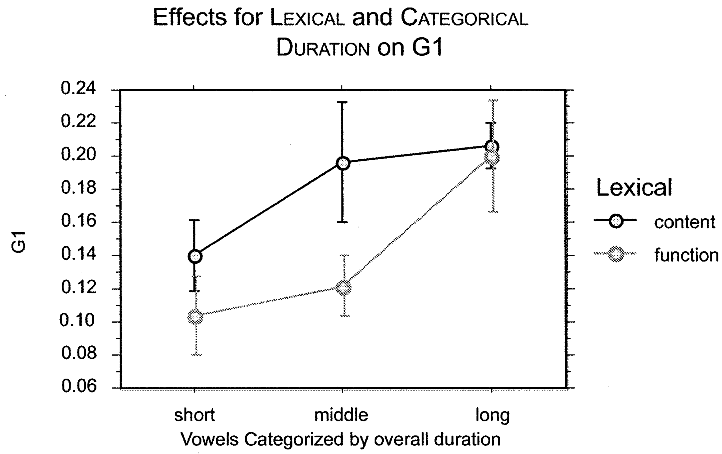

Figure 4.15 Error bar graph illustrates the interaction between LEXICAL and CATEGORICAL DURATION on G1.

The interaction graph shows that the content G1 in long durations is close to the function G1. The number of tokens in the content category was 35 while the number of tokens in the function category was 9. Although the number of tokens for function in the long category was small, the whisker is longer than that for the content one. Whiskers indicate the 95% confidence intervals. Confidence intervals are calculated based on means, standard deviations and number of tokens. A long whisker indicates that the standard deviation is large. The small number of tokens and the long whiskers indicate that it is possible that the average in the long category is influenced by some outliers. Next I will show the results for the analyses of ΔG1.

4.4.5 Δ G1

The trajectories of G1 show that the onsets of G1 values for content [a]s were greater than those for function [a]s. To investigate the changes of G1 values, G1 values were converted to ΔG1.

4.4.5.1 Trajectories of ΔG1

The ΔG1 values are summarized in Table 4.9.

Table 4.9 Mean Δ G1 and DURATION at five points

	average ΔG1					average duration
LEXICAL	0	1/4	2/4	3/4	end	
content	0	0.0839	0.1161	0.1955	0.1320	92.44
function	0	0.0822	0.1237	0.1073	0.0122	70.14

Table 4.9 shows that the average ΔG1 of content [a] was greater than that of function [a] at all points. The ΔG1 values are plotted in Figure 4.16.

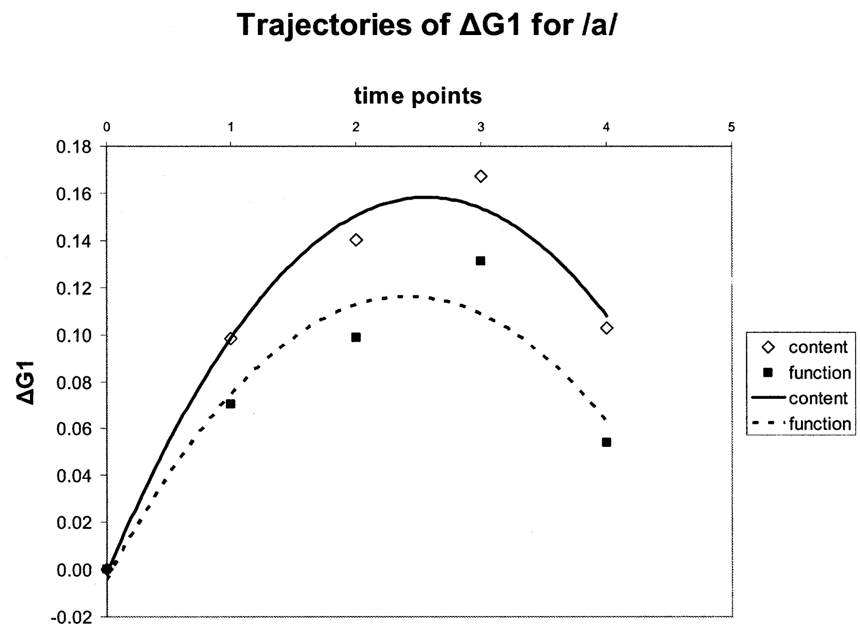

Figure 4.16 The scatter graph illustrating the trajectories of ΔG1. Unfilled square: content, filled square: function, solid line: content, dashed line: function.

The trajectories of ΔG1 look similar to the trajectories of ΔF1 although the trajectories of ΔG1 are more asymmetric than those of ΔF1. These trajectories indicate that the target for function /a/ seems to be different from that for content /a/.

4.4.5.2 ΔG1 and CATEGORICAL DURATION

The durations were again classified into 3 categories (short, middle, and long). If undershoot is responsible for G1 centralization in function [a], then different ΔG1 values should be restricted to different durations. If centralization is caused by reduction, then different ΔG1 values should be seen at the same duration. The averages of ΔG1 for individual categorical durations are summarized in Table 4.10.

Table 4.10 Mean ΔG1 in three CATEGORICAL DURATIONS

LEXICAL	CATEGORICAL DURATION	short	middle	long
content	# of tokens	16	15	35
	ΔF1	0.1166	0.1182	0.1608
function	# of tokens	28	29	9
	ΔF1	0.0694	0.1237	0.1095

Table 4.10 shows that the average ΔG1s for content [a]s were greater than the content counterparts in the short and long categories but not in the middle category. The results of a factorial ANOVA with ΔG1 as the dependent variable, and CATEGORIZED DURATION and LEXICAL as independent variables show that there were significant effects for LEXICAL, $F(1, 126) = 5.922$, $p = 0.016$, and for CATEGORICAL DURATION, $F(2, 126) = 3.799$, $p = 0.025$, but there was not a significant INTERACTION between the two, $F(2, 126) = 2.174$, $p = 0.118$. The next error bar graph illustrates the interaction between LEXICAL and CATEGORICAL DURATION.

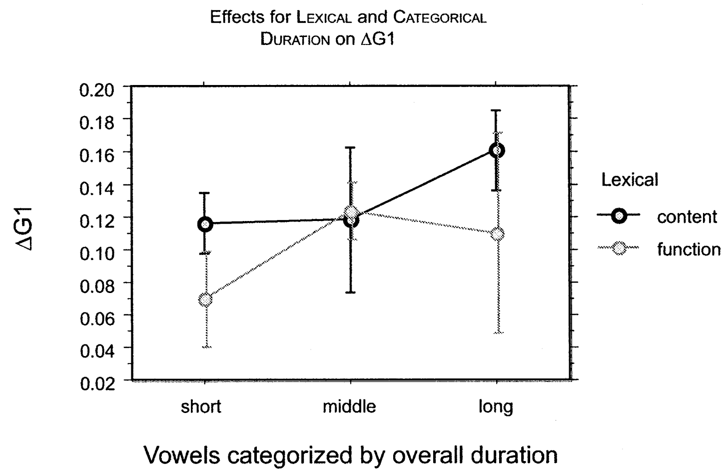

Figure 4.17 Error bar graph illustrating the interaction between LEXICAL and CATEGORICAL DURATION.

In Figure 4.7, the average of function ΔG1 in the middle category is greater than content counterparts. However, the whiskers for content ΔG1 in the middle category and the whiskers for function ΔG1 in the long category were long. Again, it is possible that an outlier causes the low average for content ΔG1 in the middle category.

In summary, there was a significant effect for LEXICAL and CATEGORICAL DURATION on G1 and ΔG1. Content G1 of /a/ was higher than its function counterparts in most points.

4.5 Discussion 3: Lexical vowel reduction

The results of the correlation tests indicate that there was a significant correlation between DURATION and F1 (and G1) for [a]. However, the correlation for content [a] was not equal to that for function [a]. Note that the correlation between content F1 and DURATION was not significant while the correlation between function F1 and DURATION was. For convenience, in the following section, I refer to F1 for discussion although the results of G1 also show a significant effect for LEXICAL.

The trajectories of F1 values for content [a] were different from those for function [a]. The trajectories for content [a] reached a peak later than those for function [a]. In addition, the calculated F1 values yielded by the content equation were different from the measured F1 values.

The results of my experiments agreed with Van Bergem's result for Dutch. The results of the factorial ANOVA of F1 and G1 show that there were significant effects for LEXICAL and DURATION but no significant interaction between the two. As Van Bergem pointed out, contextual assimilation for unreduced vowels was less than contextual assimilation for reduced vowels. In other words, there was a difference in the degree of contextual assimilation.

If reduction were the result of undershoot, then we would expect that at least the trajectories at the beginning of vowels should be overlapped. However, the trajectories of content [a] did not overlap with those of function [a]. If the target of content F1 and G1 for [a] was the same as the target for function [a], then it is possible that speakers might make an effort to cover the short duration so that F1 movement of function [a] toward the target might be faster than that of content [a]. However, the results showed they were not. The slope of F1 (F1 movement / duration) for the content category was steeper than that for the function category. Thus, the target of content F1 seemed higher than the target of function F1.

Furthermore, the results of the factorial ANOVA of Δ F1 showed that there was a significant lexical effect although the starting point for content vowels was the same as the starting point for function vowels. It indicated that the target for F1 of function [a] was different from the target for F1 of content [a]. Furthermore, the fact that no significant interactions were found between duration and F1 or DURATION Δ F1 indicated that F1 and ΔF1 are not tied to the duration.

Therefore, spectral vowel reduction was caused by lexical reduction in addition to vowel undershoot. The target for content [a] is different from that for function [a].

In the H & H hypothesis (Lindblom 1990), there are two constraints (output-oriented constraints and system-oriented constraints), which exist in a state of a tug-of-war. An example that the output-oriented constraint wins is the Lombard effect, which is the phenomenon in which a speaker clearly enunciates words to help a listener understand in a noisy environment. An example in which the system-oriented constraint wins is vowel reduction. The frequency of the Japanese function particles is high. In addition, some particles such as the subject marker /ga/ are redundant since sentences without them are grammatical. Both facts indicate that Japanese function particles are in the situation in which the system-oriented constraint wins. In this case, then, speakers do not have to pronounce these particles clearly. Consequently, the target formants for function vowels for [a] are different from those for content vowels.

5 Conclusion

In the first two sections of this dissertation, I discussed the vowel reduction in Japanese function particles. I had two questions: First, were the vowels in function particles shorter than their counterpart vowels in content words? Second, were they more centralized? In the third section, I investigated whether the spectral vowel reduction was determined only by the duration.

5.1 Durational vowel reduction

In Chapter 2, I introduced two prior studies that had focused on Japanese vowel reduction in the durational dimension: one conducted by Campbell and the other conducted by Ueyama. The first study by Campbell seemed to show that vowels in function words were of longer duration than were vowels in content words, whereas the second study by Ueyama indicated the opposite result. Thus, the results of the two appeared to be contradictory. I also noted, however, that both experiments failed to control for other factors which could affect the vowel length; in addition, the study by Campbell was based on the read speech produced by a male speaker while the study by Ueyama was based on the read speech produced by two female speakers.

I then introduced my own research, which was specifically designed to control for the factors affecting the duration of phones. Then I introduced my methodology for this experiment, which incorporated controls for the number of syllables, for the rate of speech, and for adjacent phones.

I then looked at several aspects of my results. The overall results indicated that vowels in function words were shorter than those in content words. The results showed that

[a] and [o] were clearly longer duration in content words than in function words, while [e] was not. The results of the Wilcoxon signed-rank test agreed with the results of the ANOVA. I next examined the possibility of a token factor. There was a significant effect only for [e]. I hypothesized that the duration of function [e] in [hitode] was longer than expected because of phrase final lengthening. Therefore in future studies, greater control over intonational phrasing might result in more uniform effects.

5.2 Spectral vowel reduction

Chapter 3 was devoted to answering the second question: Were vowels in function words more centralized than vowel in content words? First I introduced prior studies that have focused on spectral vowel reduction. Lindblom (1963) showed that spectral vowel reduction was due to durational vowel reduction. A second study conducted by Nord (1986) showed that the stress factor affected the formants of vowels. The research conducted by Van Bergem (1993) indicated that spectral vowel reduction was due to the increased contextual assimilation. The last study conducted by Moon and Lindblom (1984) indicated that variation in formant frequencies was due to speaking style. In these studies, it is clear that the duration of vowels influence the formants. I also introduced the study conducted by Keating and Huffman (1984). The results indicated that Japanese vowels in prose were more collapsed toward the center of vowel space than were those in the word-list readings.

I then introduced my own research, in which the tokens were the same as in experiment 1. I described the asymmetrical distribution of Japanese vowels. Then I introduced the normalization method developed by Nearey (1978), and explained the reason that modification was necessary for Japanese. I introduced the modified normalization, which was adjusted to the unevenly distributed Japanese vowel system.

I then looked at the results in several aspects. The overall results of the distance from the center of the vowel space indicated that Japanese vowels in function words were more central than were those in content words. There were significant effects for LEXICAL on F1 and F2 values of [a], and G2 values of [e]. In addition, there was a near-significant effect for LEXICAL on F2 of [e]. Furthermore, the results of the Wilcoxon signed-rank test indicate that there was a significant effect for LEXICAL on F2 of [e] and [o], and on G2 of all vowels.

I next examined the possibility of a token factor. The results indicate that there was a significant effect for TOKEN on F1 for [a] and [e], and on F2 for all vowels. Then I reasoned these results with coarticulation effects.

5.3 Vowel reduction and undershoot

Chapter 4 was devoted to Question 3: Was the spectral vowel reduction determined only by the duration? Overall results of Pearson's correlation indicate that there was an effect for DURATION. However, there was an independent effect for LEXICAL for [a] as well indicating that the undershoot effect was not the only factor at work in spectral vowel reduction. In my conclusion, the target formant for function vowels seemed different from the target for content vowels. Thus, the answer to the third question was that there was a significant effect for LEXICAL.

The results of my research indicate that there is a lexical effect in addition to a vowel undershoot-effect. In previous studies, this lexical effect has not been observed. It is possible that researchers overlooked the lexical effects. I believe that researchers would find a lexical effect if they re-examined the data using categorical duration.

5.4 Improving my experiments

If I were doing the research again, I would make two improvements in the experiments. First, I would like to record words in a word-list as did Van Bergem and other researchers. Possible words in the word-list are the same content words in the sentence tokens and nonsense words in the context of [d_de][1] such as [dade]. The researchers in previous studies used the formants of the words in the word-lists as the target. Second, I would like to change tokens because some tokens led to unexpected results. For example, [de] in [hitode] was the grammatically copula, which indicated the end of a short sentence; consequently it led to contradictory results. Thus, I would instead use the locative particle, [de]. In another token, [hato], the following vowel, [ɯ], tended to be devoiced since it is a high vowel and its adjacent phones were voiceless consonants. Therefore, I would instead use words whose first vowel is [a], [e] or [o] in the words that follow the target words. When I designed the experiments, I did not realize how coarticulation effects work. I also had a hard time finding pairs of words (content and function). To observe the effect for LEXICAL, I would optimize the F2 distance of vowels such as for [e], using [o] as previous and following vowels. In addition, I would use [w] as the following consonants since [w][2] greatly lowers F2 values (Lehiste 1964).

---

[1] Moon and Lindblom (1994) used front vowels in the context of [h_d]. In Japanese, [h] changes to [ç] or [ɸ] before [i] or [ɯ], respectively so it is inappropriate to use vowels in the context of [h_d].

[2] In Japanese, [l] is not a phoneme and [ɾ] is a flap so probably [w] is supposed to lower F2 values most.

5.5 Future research

I would like to investigate the different coarticulation effect in English. In English, most vowels in function words are reduced. I summarized examples of vowel reduction in English in Table 5.1.

Table 5.1 Vowel reduction in English

function word	pronunciation with focus	reduced pronunciation
a	eʲ	ə
and	ænd	ənd, nd
the	ði:	ðə
of	ʌv	əv
would	wʊd	wəd

Table 5.1 illustrates the pattern of vowel reduction in English function words. If coarticulation is the only reason for the vowel centralization, then these function vowels, especially the indefinite article “a”, will not be centralized in the context of [w_l].

Another possible research topic is the coarticulation effect in Japanese. Previous research on coarticulation effects conducted by Kondo and Arai (1998) showed that there was a significant coarticulation effect of the following vowel (anticipatory effect) on F1 and F2 for vowel [a]. However, there was a significant effect on F2 but not on F1 for [i]. Furthermore, the significance level for [a] was greater than that for [i]. Thus, it is possible that F1 and F2 vary extensively for [a] in Japanese but not for other vowels. Unfortunately, not many studies on coarticulation effects in Japanese have been conducted. Thus, research on coarticulation effects in Japanese needs to be done.

Function [o] in the token “hato” showed no lexical effect. The following vowel of the target [o], [ɯ], is devoiced. It seems that this devoiced vowel has little V-to-V anticipatory effect. The research of the coarticulation effect of devoiced vowels on the consonant [ʃ] (Beckman and Shoji 1984) showed that a difference of the following devoiced vowels ([i̥] and [ɯ̥]) led to a difference of the lower edge formant value of the fricative [ʃ]. However, no research on the V-to-V coarticulation effect of devoiced vowels has yet been conducted. Therefore, it is possible that the voicing of a vowel influences the formants of preceding vowels.

The results indicate that spectral vowel reduction is observable in Japanese. This opens the door to various types of research. For example, Japanese motherese should be investigated. Previous research conducted by Masataka (2002) showed that Japanese motherese has wide excursion of pitch as does motherese in other languages. In Masataka's research, young Japanese females told stories based on a children's picture book to a child (not her own child), young female adults and elderly women. The results showed that the range of pitch in child-directed speech was greater than in adult-directed speech. In addition, when the subjects talked to elderly women, the pitch range was also large. Since there is difference in pitch, it is possible that there is difference of duration and formants between child-directed speech and adult-directed speech. Are long vowels lengthened more than short vowels? Are geminate consonants lengthened more than single consonants? Are vowels in child directed speech more expanded than adult directed speech? Are the targets for vowels in child-direct speech different from those in adult-direct speech? I believe all of these questions raised by the results of my research would lead to fruitful avenues of further study.

References

Balota, D. A., Boland, J. E. & Shields, L. W. 1989. Priming in pronunciation: Beyond pattern recognition and onset latency. Journal of Memory and Language, **28**, 14-36.

Beckman, M. E. & Edwards, J. 1990. Lengthening and shortenings and the nature of prosodic constituency. In *Paper in laboratory phonology 1*: between the grammar and physics of speech (J. Kingston & M. E. Beckman editors) Cambridge: Cambridge University Press.

Bolinger, D and Sears, D. 1981. *Aspects of Language* third ed. New York: Harcourt Brace Jovanovich.

Bybee, J. L. 1994. A view of phonology from a cognitive perspective. Cognitive Linguistics, **5** (4), 285-305.

Campbell, N. 1999. Segmental Elasticity and Timing in Japanese Speech. *Speech Perception, Production and Linguistic Structure.* Tokyo: Ohmsha.

Han, M. 1964. Acoustic Manifestation of Mora Timing in Japanese. *Journal of the Acoustical Society of America*, **96**, 73-81.

Hunnicut, S. 1985. Intelligibility vs. redundancy – conditions of dependency. Language and Speech, **28**, 47-56.

Lehiste, I. 1964. *Acoustical Characteristics of Selected English Consonants*, Bloomington, Indiana University.

Kaiki, N., Tkeda, K., and Sagisaka, Y. 1990. The control of segmental duration in speech synthesis using linguistic properties. *Proceedings of ESCA Tutorial workshop on Speech Synthesis*, Autrans, France. 18-22.

Kirk, R. E. 1999. *Statistics: An Introduction* (4th ed.). Fort Worth: Harcourt Brace College Publishers.

Keating, P and Huffman, M. 1984. Vowel Variation in Japanese. *Phonetica*, **41**, 191-207.

Klatt, D. H. 1976. Linguistics Uses of Segmental Duration in English – Acoustic and Perceptual Evidence. *Journal of Acoustical Society of America*, **59**, 1208-1221.

Kuriyagawa, F. and Sawashima, M. 1985. Vowel Duration in Japanese /hu**ku/ and /huku**/ (II). *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, **19**, 139-155.

Kuriyagawa, F. and Sawashima, M. 1986. Vowel Duration in Japanese /tsuXku/ and /tsukuX/. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, **20**, 119-130.

Lieberman, P. 1963. Some effects of semantic and grammatical context on the production and perception of speech. Language and Speech, **6**, 172-187.

Lieberman, P and Blumstein, S. 1988. *Speech Physiology, Speech Perception, and Acoustic Phonetics*. Cambridge: Cambridge University Press.

Lindblom, B. 1963. Spectrographic Study of Vowel Reduction. *Journal of Acoustical Society of America*,. **35**, 1773-1781.

Lindblom, B. 1990. Explaining Phonetic Variation: A Sketch of the H&H Theory. In W. J. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modeling*. Kluwer Academic Publishers: Dordrecht, 403-439.

Martin, S. 1988. *A Reference Grammar of Japanese*. Rutland, VT: C. E. Tuttle Co.

Masataka, N. 2002 Pitch modification when interacting with elders: Japanese women with and without experience with infants. *Journal of Child Language,* **29,** 939-951.

Moon, S.J, and Lindblom, B. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *Journal of Acoustical Society of America*, **96**, 40-55.

Nearey, T. M. 1978. *Phonetic features systems for vowels.* Indiana University Linguistics Club: Bloomington IN.

Nearey, T. M. 1989. Static, Dynamic, and Relational Properties in Vowel Perception. Journal of Acoustic Society of America, 85 2088-2113.

Nearey, T. M. 1992. Applications of Generalized Linear Modeling to Vowel Data. *Proceedings of International Conference on Spoken Language Processing,* 583-586.

Nord, L. 1986. Acoustic Studies of Vowel Reduction in Swedish. *Quarterly Progress and status report*, **4**, 19-36.

Otake, T., and Cutler, A. 1999. Pitch Accent in Spoken-word Recognition in Japanese. *Journal of the Acoustical Society of America*, **105**, 1877-1888.

Port, R. 1981. Linguistic timing factors in combination. *Journal of Acoustic Society of America*, **69**, 262-274.

Port, R, Al-Ani, S. and Maeda, S. 1980. Temporal Compensation. *Phonetica*, **37**, 235-252.

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.

Ueyama, M. 1999. An Experimental Study of Vowel Duration in Phrase-Final Contexts in Japanese. *University of California Working Papers in Phonetics,* 174-182.

Turk, A. E., and Shattuck-Hufnagel. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics*, **28**, 397-440.

Vance, T. J. 1986. *An Introduction to Japanese Phonology*. New York: State University of New York.

Van Bergem, D. R. 1993. Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication*, **12**, 1-23.

Wright, R 2003 Lexical competition and reduction. *Papers in Laboratory Phonology VI.* 75-87. Cambridge: Cambridge University Press.

Zipf, G. K. 1935. *The psycho-biology of Language.* Boston: Houghton Mifflin.

Appendix A: Measurements

A Duration

Table A.1 Duration for [kiga] (ms)

subject	content	function	subject	content	function
kt	61.64	48.81	at	78.65	60.45
kt	40.76	39.22	at	60.19	83.24
kt	59.69	56.99	at	61.47	76.32
m2h	88.24	68.35	ei	93.40	85.10
m2h	108.38	68.30	ei	117.94	88.00
m2h	78.50	57.34	ei	99.85	102.97
mh	57.30	51.49	kn	80.22	54.78
mh	61.07	75.92	kn	53.68	69.01
mh	58.97	66.20	kn	57.93	67.65
ss	113.84	75.98	ye	88.66	94.32
ss	146.41	74.10	ye	95.94	90.44
ss	165.23	52.63	ye	83.58	130.13
ta	69.66	68.91	yt	84.65	73.53
ta	102.65	64.63	yt	72.36	74.40
ta	73.03	72.54	yt	80.88	79.51
ykm	92.98	58.95			
ykm	107.08	46.87			
ykm	106.16	60.09			

Table A.2 Duration for [jinbutsuga] (ms)

subject	content	function	subject	content	function
kt	94.99	31.37	at	105.36	72.73
kt	86.94	52.60	at	95.15	70.93
kt	89.00	143.23	at	66.18	72.85
m2h	107.39	58.32	ei	61.89	74.87
m2h	146.20	69.90	ei	64.14	57.79
m2h	114.87	72.07	ei	62.32	71.51
mh	68.89	62.61	kn	123.18	56.20
mh	61.12	49.94	kn	107.00	56.13
mh	146.45	70.22	kn	126.41	47.94
ss	119.70	117.51	ye	76.38	71.60
ss	154.05	84.66	ye	72.85	55.14
ss	142.63	53.13	ye	76.82	77.11
ta	109.17	47.68	yt	121.35	71.81
ta	50.05	63.31	yt	132.26	48.14
ta	100.87	68.13	yt	115.37	54.74
ykm	86.13	97.34			
ykm	129.46	59.88			
ykm	85.61	130.75			

Table A.3 Duration for [hitode] (ms)

subject	content	function	subject	content	function
kt	84.77	50.36	at	88.87	101.37
kt	50.91	93.06	at	73.99	88.74
kt	88.46	68.88	at	114.93	112.58
m2h	56.67	125.43	ei	54.62	110.42
m2h	56.88	122.98	ei	72.46	57.95
m2h	76.82	161.41	ei	64.27	127.90
mh	52.61	121.70	ye	90.31	61.87
mh	56.93	116.13	ye	94.17	36.00
mh	108.37	122.07	ye	91.78	64.58
ss	122.17	104.71	yt	156.89	99.24
ss	132.86	143.49	yt	122.73	80.93
ss	163.94	174.53	yt	106.17	107.42
ta	65.87	114.04			
ta	50.07	110.66			
ta	84.94	105.23			
ykm	101.15	61.13			
ykm	81.11	121.55			
ykm	86.69	97.15			

Table A.4 Duration for /tade/ (ms)

subject	content	function	subject	content	function
kt	64.55	92.90	at	64.01	101.11
kt	59.33	79.13	at	75.37	62.26
kt	76.29	67.49	at	76.96	74.55
m2h	83.98	70.09	ei	77.73	75.27
m2h	106.20	71.76	ei	87.34	71.45
m2h	78.87	67.39	ei	74.38	77.86
mh	92.43	66.12	kn	64.51	105.35
mh	76.68	61.34	kn	76.14	73.94
mh	78.97	50.61	kn	44.89	102.59
ss	85.78	75.58	ye	104.12	114.19
ss	83.01	54.63	ye	79.28	78.30
ss	56.18	81.40	ye	75.14	73.70
ta	48.21	92.24	yt	121.86	74.26
ta	69.80	64.18	yt	70.85	54.92
ta	51.12	52.00	yt	78.66	46.99
ykm	53.76	57.86			
ykm	65.09	59.97			
ykm	51.88	59.94			

Table A.5 Duration for [hato] (ms)

subject	content	function	subject	content	function
kt	41.80	64.80	at	115.69	52.72
kt	48.18	65.34	at	71.42	66.46
kt	43.22	52.83	at	73.86	67.52
m2h	72.80	65.81	ei	54.40	59.01
m2h	67.54	68.80	ei	52.51	46.53
m2h	43.07	93.53	ei	74.02	84.33
mh	37.25	41.37	kn	87.02	64.81
mh	40.61	26.94	kn	86.75	40.98
mh	20.60	114.12	kn	80.36	36.21
ss	72.45	58.26	ye	32.49	42.91
ss	96.90	72.60	ye	39.29	28.88
ss	81.06	51.40	ye	54.68	58.66
ta	60.67	43.88	yt	114.25	55.22
ta	34.74	51.07	yt	77.12	56.96
ta	77.48	61.49	yt	62.56	121.88
ykm	153.11	84.00			
ykm	79.42	88.77			
ykm	98.78	112.99			

Table A.6 Duration for [kogoto] (ms)

subject	content	function	subject	content	function
kt	88.82	58.81	at	64.08	64.39
kt	28.07	91.18	at	39.81	50.06
kt	83.42	42.63	at	28.85	85.82
m2h	73.56	74.86	ei	49.17	42.28
m2h	68.52	59.16	ei	37.29	48.23
m2h	82.08	79.25	ei	49.03	41.04
mh	47.44	33.90	kn	157.37	49.53
mh	44.42	52.77	kn	78.60	42.04
mh	57.14	21.48	kn	46.88	50.18
ss	92.49	67.49	ye	46.37	47.57
ss	116.56	65.19	ye	32.98	44.98
ss	145.97	57.60	ye	39.76	50.53
ta	90.32	29.22	yt	78.76	33.82
ta	45.10	109.04	yt	102.64	49.21
ta	160.43	43.89	yt	77.89	48.86
ykm	127.54	110.17			
ykm	132.64	16.83			
ykm	125.89	25.24			

B Formants

Table A.7 Formants for [kiga] produced by male speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
kt	658	1399	568	1383
kt	566	1515	564	1385
kt	610	1437	615	1430
m2h	694	1381	605	1610
m2h	705	1411	634	1495
m2h	643	1404	614	1496
mh	498	1684	475	1706
mh	507	1674	506	1703
mh	545	1605	498	1720
ss	586	1386	465	1467
ss	659	1269	418	1412
ss	644	1278	441	1451
ta	629	1428	608	1416
ta	682	1209	640	1446
ta	648	1410	601	1368
ykm	631	1268	517	1436
ykm	616	1247	508	1431
ykm	677	1235	482	1499

Table A.8 Formants for [kiga] produced by female speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
at	877	1805	724	1864
at	867	1796	635	1763
at	826	1845	798	1681
ei	807	1698	799	1863
ei	810	1719	790	1684
ei	767	1551	796	1670
kn	788	1703	632	1261
kn	742	1880	632	1598
kn	669	1566	585	1408
ye	698	1492	653	1531
ye	676	1592	627	1504
ye	651	1469	661	1485
yt	641	1870	591	1872
yt	557	1977	577	1904
yt	625	1902	600	1961

Table A.9 Formants for [jinbutsuga] produced by male speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
kt	676	1276	559	1375
kt	631	1340	577	1380
kt	630	1307	656	1310
m2h	698	1321	603	1429
m2h	711	1363	609	1601
m2h	733	1321	647	1443
mh	580	1477	563	1427
mh	540	1473	545	1446
mh	654	1441	537	1440
ss	614	1228	611	1218
ss	653	1197	572	1239
ss	658	1204	502	1290
ta	590	1274	533	1429
ta	563	1377	507	1443
ta	734	1200	518	1372
ykm	682	1239	606	1266
ykm	691	1239	525	1352
ykm	687	1225	619	1206

Table A.10 Formants for [jinbutsuga] produced by female speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
at	882	1666	744	1636
at	895	1614	739	1569
at	808	1657	726	1637
ei	799	1635	833	1650
ei	820	1633	807	1626
ei	773	1618	799	1614
kn	806	1472	671	1600
kn	807	1519	603	1495
kn	777	1342	672	1637
ye	737	1398	539	1382
ye	706	1425	629	1496
ye	672	1465	628	1493
yt	886	1587	661	1645
yt	854	1662	500	1620
yt	806	1671	546	1611

Table A.11 Formants for [hitode] produced by male speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
kt	449	1679	482	1719
kt	489	1631	504	1722
kt	493	1883	524	1785
m2h	495	1800	555	1932
m2h	502	1792	534	1803
m2h	519	1715	547	1880
mh	440	1762	529	1699
mh	457	1666	540	1753
mh	501	1682	496	1631
ss	473	1781	493	1710
ss	479	1780	498	1693
ss	464	1841	443	1763
ta	460	1689	556	1653
ta	460	1687	546	1671
ta	491	1721	555	1646
ykm	497	1869	473	1779
ykm	493	1804	487	1828
ykm	455	1766	470	1807

Table A.12 Formants for [hitode] produced by female speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
at	682	2402	653	2352
at	575	2301	570	2331
at	628	2300	609	2357
ei	611	2294	642	2278
ei	623	2281	657	2274
ei	639	2138	605	2295
ye	479	2162	505	2069
ye	602	2065	516	1912
ye	631	2062	511	2005
yt	653	2262	622	2030
yt	652	2135	585	2025
yt	575	2132	500	2099

Table A.13 Formants for [tade] produced by male speakers (Hz)

subject	content		function	
	F1	F2	F1	F2
kt	407	1910	464	1914
kt	435	1698	477	1710
kt	450	1993	518	1815
m2h	502	1943	525	1951
m2h	545	1815	485	1934
m2h	532	1934	497	2014
mh	502	1834	401	1699
mh	492	1780	469	1722
mh	468	1817	469	1774
ss	386	1909	444	1779
ss	373	1844	373	1969
ss	363	1842	412	1838
ta	486	1814	493	1672
ta	468	1806	450	1679
ta	477	1929	447	1741
ykm	412	1926	404	1946
ykm	433	1869	410	1911
ykm	423	1940	423	1919

Table A.14 Formants for [tade] produced by female speakers (Hz)

subject	content		function	
	F1	F2	F1	F2
at	689	2476	680	2341
at	616	2383	707	2460
at	696	2466	734	2446
ei	574	2340	552	2386
ei	601	2345	576	2295
ei	567	2306	550	2367
kn	553	2247	549	2254
kn	466	2264	497	2086
kn	515	2176	510	2184
ye	383	2263	385	2229
ye	491	2005	529	2234
ye	389	2288	517	2209
yt	639	2375	467	2121
yt	541	2151	456	2207
yt	454	2095	445	2223

Table A.15 Formants for [hato] produced by male speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
kt	434	960	427	906
kt	498	963	425	969
kt	450	927	376	1008
m2h	491	977	482	983
m2h	453	1064	505	975
m2h	460	1143	499	938
mh	425	1192	465	1328
mh	457	1185	445	1290
mh	419	1292	459	880
ss	430	984	359	1067
ss	497	930	354	970
ss	459	965	410	1135
ta	479	992	466	1063
ta	451	1008	479	1035
ta	518	959	461	1067
ykm	470	734	502	955
ykm	485	899	446	944
ykm	463	863	444	901

Table A.16 Formants for [hato] produced by female speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
at	579	1223	583	1510
at	574	1275	529	1295
at	598	1230	563	1282
ei	603	1311	604	1321
ei	619	1354	623	1246
ei	564	1276	592	1121
kn	536	1157	463	1094
kn	458	1235	522	1300
kn	493	1204	537	1366
ye	554	1162	472	1149
ye	503	1283	448	1170
ye	537	1188	552	1216
yt	581	1328	539	1351
yt	487	1451	595	1102
yt	474	1500	574	1269

Table A.17 Formants for [kogoto] produced by male speakers (Hz)

	content		function	
subject	F1	F2	F1	F2
kt	458	880	517	965
kt	393	948	462	847
kt	452	837	501	927
m2h	492	958	489	925
m2h	485	931	495	1009
m2h	496	948	515	972
mh	451	1164	458	1244
mh	4654	1100	480	1105
mh	466	1114	431	1341
ss	455	994	435	1033
ss	497	953	424	990
ss	433	897	399	986
ta	549	870	447	946
ta	491	1052	520	796
ta	493	879	488	951
ykm	479	802	406	688
ykm	472	885	405	921
ykm	489	882	434	848

Table A.18 Formants for [kogoto] produced by female speakers (Hz)

at	645	1291	559	1258
at	555	1330	568	1263
at	621	1294	576	1309
ei	705	1220	706	1127
ei	660	1228	693	1195
ei	751	1301	684	1275
kn	525	988	464	1087
kn	494	11793	442	1151
kn	487	1133	464	1210
ye	439	1047	556	1187
ye	539	1058	515	1149
ye	446	1095	518	1179
yt	508	1077	472	1250
yt	528	1092	482	1360
yt	484	1081	480	1341

C Distance

Table A.19 Distance for [kiga]

subject	content	function	subject	content	function
kt	0.194995	0.133004	at	0.215664	0.145714
kt	0.151041	0.130289	at	0.213184	0.143146
kt	0.168251	0.170163	at	0.189110	0.207427
m2h	0.197931	0.172807	ei	0.160852	0.162380
m2h	0.206808	0.172552	ei	0.162499	0.151638
m2h	0.167704	0.160226	ei	0.143489	0.154715
mh	0.137394	0.133283	kn	0.243728	0.177757
mh	0.139008	0.144912	kn	0.227414	0.146115
mh	0.146062	0.145094	kn	0.171289	0.124453
ss	0.189134	0.103884	ye	0.223156	0.195206
ss	0.236951	0.057874	ye	0.211477	0.177175
ss	0.226935	0.082830	ye	0.193423	0.200089
ta	0.170369	0.155163	yt	0.172051	0.143832
ta	0.196453	0.179179	yt	0.143015	0.140759
ta	0.181188	0.145916	yt	0.167191	0.162080
ykm	0.184963	0.114944			
ykm	0.174120	0.107061			
ykm	0.215721	0.103146			

Table A.20 Distance for [jinbutsuga]

subject	content	function	subject	content	function
kt	0.198951	0.125764	at	0.239902	0.199322
kt	0.172322	0.139135	at	0.254153	0.213936
kt	0.170152	0.187415	at	0.215567	0.193986
m2h	0.197474	0.144391	ei	0.156947	0.174997
m2h	0.207081	0.173180	ei	0.168168	0.161675
m2h	0.218329	0.174485	ei	0.143465	0.157777
mh	0.148824	0.131218	kn	0.253944	0.172084
mh	0.120666	0.120322	kn	0.252924	0.128653
mh	0.194347	0.113854	kn	0.247401	0.173416
ss	0.207123	0.205342	ye	0.248401	0.115838
ss	0.234401	0.175778	ye	0.229206	0.178292
ss	0.237897	0.118618	ye	0.207341	0.177676
ta	0.133220	0.105393	yt	0.287521	0.163497
ta	0.119966	0.090481	yt	0.273597	0.047537
ta	0.228783	0.085636	yt	0.249223	0.081110
ykm	0.218826	0.167549			
ykm	0.224459	0.110031			
ykm	0.222135	0.177317			

Table A.21 Distance for [hitode]

subject	content	function	subject	content	function
kt	0.136468	0.154037	at	0.063494	0.046613
kt	0.135223	0.162246	at	0.026137	0.023006
kt	0.194847	0.184127	at	0.036736	0.019363
m2h	0.164172	0.211438	ei	0.139602	0.144425
m2h	0.164418	0.177241	ei	0.139957	0.148253
m2h	0.152061	0.197960	ei	0.119143	0.138573
mh	0.138949	0.155006	ye	0.173796	0.165896
mh	0.118240	0.171250	ye	0.214165	0.143360
mh	0.138044	0.124470	ye	0.229043	0.157290
ss	0.170569	0.167604	yt	0.231889	0.183416
ss	0.173743	0.167385	yt	0.213683	0.164131
ss	0.178612	0.152874	yt	0.177781	0.145941
ta	0.131887	0.160941			
ta	0.131249	0.159193			
ta	0.147260	0.158876			
ykm	0.191329	0.163067			
ykm	0.176115	0.178807			
ykm	0.154807	0.168518			

Table A.22 Distance for [tade]

subject	content	function	subject	content	function
kt	0.192343	0.194952	at	0.068563	0.064329
kt	0.140079	0.150655	at	0.020912	0.079433
kt	0.210648	0.188157	at	0.072557	0.095118
m2h	0.197922	0.205560	ei	0.142765	0.150809
m2h	0.183908	0.192253	ei	0.146902	0.134608
m2h	0.204199	0.211815	ei	0.136098	0.147337
mh	0.171856	0.124870	kn	0.173701	0.173084
mh	0.156297	0.135304	kn	0.153447	0.124388
mh	0.157330	0.147577	kn	0.147265	0.147044
ss	0.173298	0.156718	ye	0.186627	0.179630
ss	0.158554	0.187108	ye	0.148068	0.204945
ss	0.159152	0.160009	ye	0.190073	0.195908
ta	0.167551	0.136733	yt	0.242565	0.145349
ta	0.161773	0.127691	yt	0.167945	0.162395
ta	0.191315	0.143084	yt	0.139936	0.165975
ykm	0.186417	0.191084			
ykm	0.174543	0.182939			
ykm	0.189732	0.185217			

Table A.23 Distance for [hato]

subject	content	function	subject	content	function
kt	0.107703	0.132741	at	0.298309	0.206793
kt	0.125245	0.103603	at	0.280414	0.277338
kt	0.124638	0.103023	at	0.295828	0.278454
m2h	0.116659	0.111072	ei	0.114499	0.111706
m2h	0.071472	0.122145	ei	0.105902	0.140053
m2h	0.042513	0.135676	ei	0.121105	0.179156
mh	0.031674	0.042013	kn	0.157548	0.163378
mh	0.046591	0.020864	kn	0.110651	0.108333
mh	0.006389	0.167068	kn	0.127286	0.100580
ss	0.125576	0.083914	ye	0.163059	0.123837
ss	0.180034	0.125079	ye	0.103020	0.108310
ss	0.146601	0.061064	ye	0.146176	0.149191
ta	0.110289	0.077828	yt	0.119229	0.087252
ta	0.096288	0.093473	yt	0.033162	0.180048
ta	0.139241	0.074671	yt	0.016218	0.125789
ykm	0.239589	0.145958			
ykm	0.160919	0.127994			
ykm	0.170013	0.146816			

Table A.24 Distance for [kogoto]

subject	content	function	subject	content	function
kt	0.148769	0.133521	at	0.277612	0.287036
kt	0.118847	0.165635	at	0.263038	0.284654
kt	0.169118	0.140717	at	0.274609	0.269093
m2h	0.124603	0.138166	ei	0.173932	0.202702
m2h	0.134583	0.104840	ei	0.156098	0.176820
m2h	0.130051	0.127498	ei	0.171484	0.150365
mh	0.049341	0.035511	kn	0.217246	0.166145
mh	0.077293	0.083591	kn	0.136496	0.141358
mh	0.073317	0.020683	kn	0.151615	0.119857
ss	0.133916	0.108937	ye	0.153471	0.157828
ss	0.172126	0.120636	ye	0.184690	0.144502
ss	0.164138	0.115316	ye	0.135389	0.156694
ta	0.188361	0.122904	yt	0.155984	0.085538
ta	0.092843	0.212087	yt	0.156110	0.052885
ta	0.163934	0.130255	yt	0.149579	0.057720
ykm	0.204604	0.261043			
ykm	0.162256	0.134170			
ykm	0.169907	0.171314			

Appendix B: Formant plot for individual speakers

In the following figures, the letters, “a”, “e” and “o” represent content vowels while the mark, “.” represents function vowels.

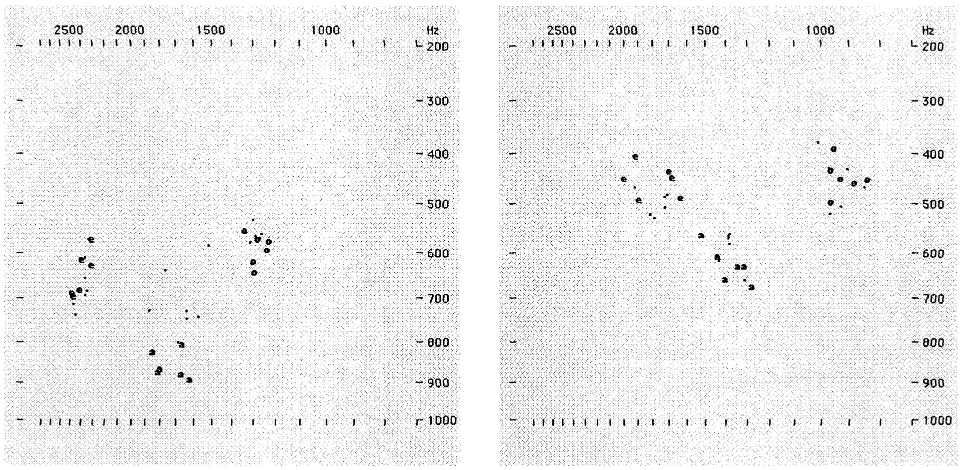

Figure A.1 Formant plot illustrates F1 (vertical axis) and F2 (horizontal axis) of Subjects “at” (left) and “kt” (right)

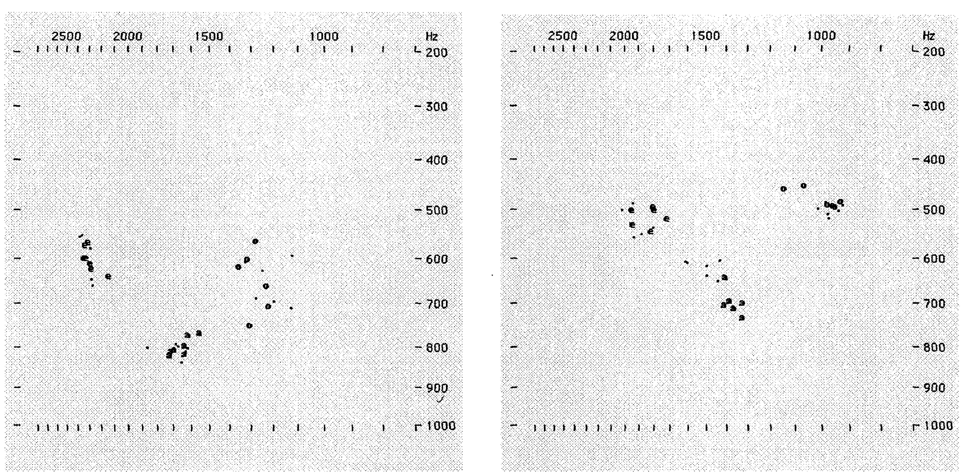

Figure A.2 Formant plot illustrates F1 (vertical axis) and F2 (horizontal axis) of subjects “ei” (left) and “m2h” (right)

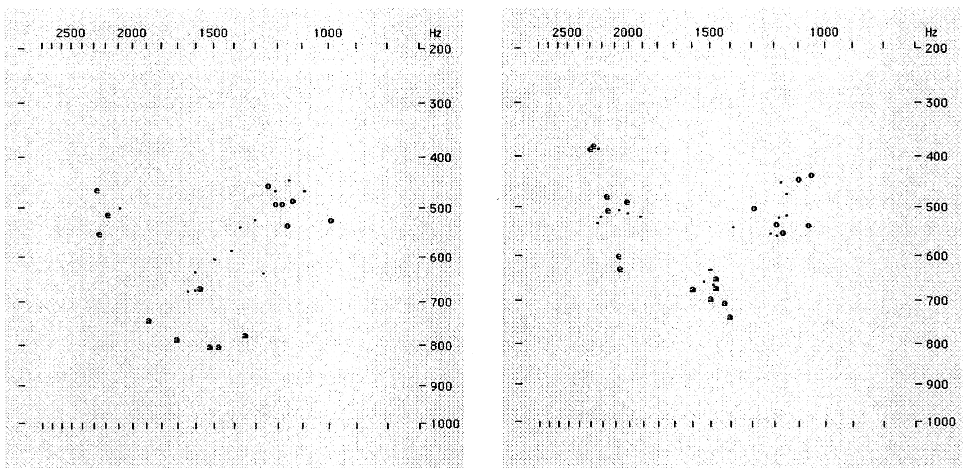

Figure A.3 Formant plot illustrates F1 (vertical axis) and F2 (horizontal axis) of subjects "kn" (left) and "mh" (right)

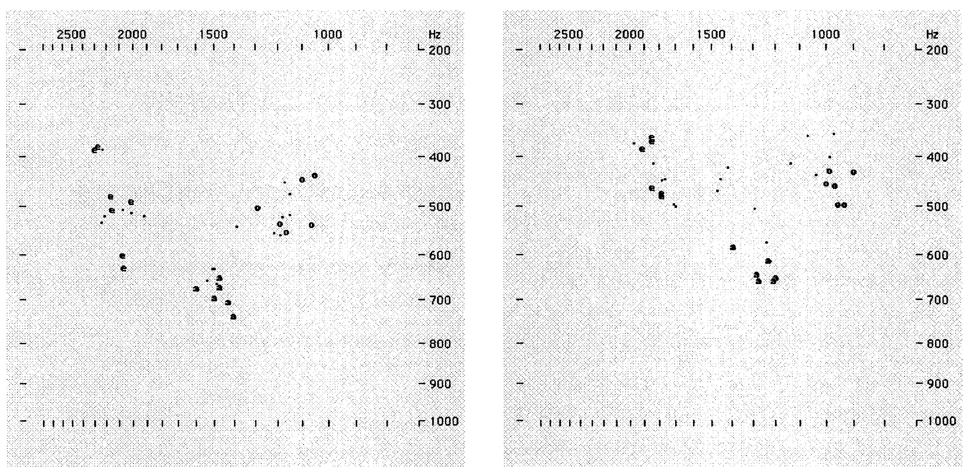

Figure A.4 Formant plot illustrates F1 (vertical axis) and F2 (horizontal axis) of subjects "ye" (left) and "ss" (right) .

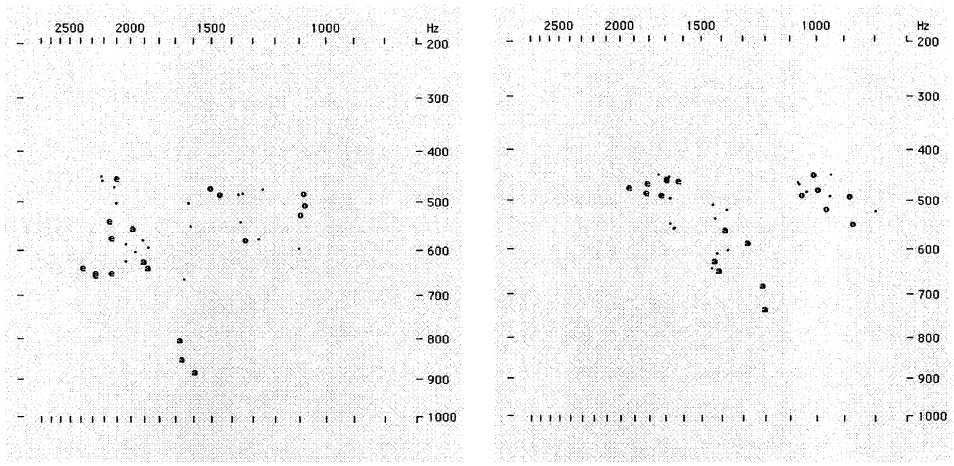

Figure A.5 Formant plot illustrates F1 (vertical axis) and F2 (horizontal axis) of subjects "yt (left) and "ta" (right). F2 of /a/ in the token /kiga/ produced by "yt" seems to be too low. However, a spectrum slice indicates it is correct.

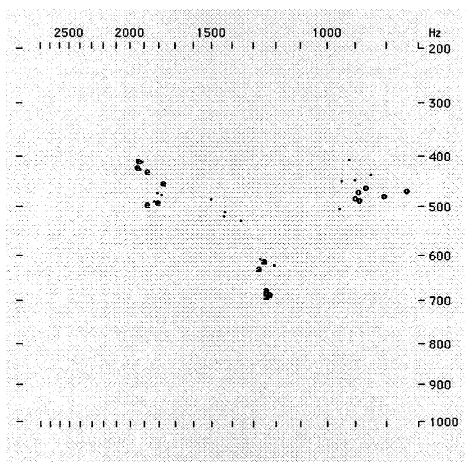

Figure A.6 Formant plot illustrating "ykm".

VITA

Setsuko Shirai was born in Kobe, Hyogo prefecture, Japan, on October 4, 1950; the first child to Shigeo Kotani and Masako Kotani. She moved with her family to Tokyo at age 10, where she lived until at the age of 37. Although she took English classes in schools, she could not understand English conversations in movies or dramas. At the age of 38, she moved with her husband to Mito, Ibaraki prefecture. She was a housewife and did not have friends. She was board so she started listening to TV News in English and studying English as a hobby. Later, she learned how to teach Japanese in non-academic classes. She came to the USA in 1994 in the internship program, in which she taught Origami in elementary schools in Spokane and helped a Japanese language teacher in Shoreline community collage. She started studying Linguistics at the University of Washington in 1995 as a non-matriculated student.

Education

Doctor of Philosophy, Linguistics, University of Washington, expected December, 2004.
Master of Arts, Linguistics, University of Washington, 1999.
Bachelor of Science in Electrical Engineering, 1973.

Publications

2004. "Lexical Effects in Japanese Vowel Reduction." Unpublished doctoral dissertation, Linguistics. University of Washington.

2004 "Spectral Vowel Reduction in Japanese." University of Washington Working Papers in Linguistics 23.

2002 "The Duration of Function Words in Japanese." University of Washington Working Paper in Linguistics 21.

2001 "Gemination in loans from English to Japanese." University of British Columbia Working Papers in Linguistics

Other Experience

Research Assistant (UW)	2003
Linguistic Tester, Japanese Speech Recognition at Microsoft Corporation	1999-2000