國 立 交 通 大 學

語言與文化研究所

碩 士 論 文

自發性國語語料中的焦點

Focus Condition in Spontaneous Taiwanese Mandarin

指導教授： 潘荷仙　教授

研 究 生：黃盈惠

中 華 民 國　九 十 三　年 八 月

自發性國語語料中的焦點

Focus Condition in Spontaneous Taiwanese Mandarin

研 究 生： 黃盈惠　　Student: Ying-Hui Huang

指導教授： 潘荷仙 博士　Advisor : Dr. Ho-Hsien Pan

國 立 交 通 大 學

語言與文化研究所

碩 士 論 文

A Thesis
Submitted to Department of Foreign Languages and Literatures
Graduate Institute of Linguistics and Cultural Studies
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Graduate Institute of Linguistics and Cultural Studies

August 2004

Hsinchu, Taiwan, Republic of China

中華民國九十三年八月

摘要

以焦點為主題的研究，在語言學的各個領域都有相關討論。然而，針對國語中的焦點，大部分研究所使用的語料為大陸地區的普通話，即使有少數針對台灣地區國語所做的研究 (Hsiung, 2002)，仍以朗讀語料 (read speech) 為分析的重點，且多為發音實驗 (production experiment)。因此，本研究以自發性語料 (spontaneous speech)，針對當前台灣地區國語中焦點的現象進行探討。

在實驗一的發音實驗 (production experiment) 中，十六個雙音節名詞涵蓋國語中可能的聲調組合，分別位在句首、句中及句尾的位置，承受窄焦點 (narrow focus) 或是沒有承受焦點 (neutral focus)，由施測者發問，受試者根據所聽到的問題及圖片為引導回答。在實驗二的感知實驗 (perception experiment) 中，使用受試者 CBY 的發音語料為刺激項，將每一答句可能對應的兩個問題以 ABX 順序錄製，受試者根據所聽到的肯定句，反推該句所回答的問句為何。

研究結果發現，在本實驗自發性國語的發音語料中，音長的增加及基頻範圍的增加皆是表現焦點的方式，然而基頻範圍的表現不如音長來的明顯，因此，音長的增加為表現焦點的主要方式。同時，受試者 CBY 和 CKC 的資料顯示：由母音 /i/，/u/，/a/的第一共振峰 (F1) 和第二共振峰 (F2) 所構成的母音空間 (vowel space)，在承受窄焦點時，有擴張 (enlargement) 的現象。本研究同時發現，在自發性國語的感知方面，音長是受試者判斷焦點最主要的依據。

Abstract

Two experiments are conducted to investigate the effect of focus in Taiwanese Mandarin. In production experiment, two types of sentences are used. There are 16 target nouns with all tonal combinations in Mandarin; each target noun consists of two syllables. Two focus conditions are manipulated: narrow focus on target noun, neutral focus on target noun. Three positions of target nouns are controlled: initial, middle and final. Three subjects participated and produced 288 sentences by answering precursor questions.

In perception experiment, the production data of subject CBY is used as stimulus without synthesis or manipulation. The stimulus is two precursor questions and one answer as a set recorded in ABX order. There are 288 questions; subjects are tested and instructed to circle the precursor question that the utterance responds to.

The results of two experiments show that in production, both duration and F0 range expansion are salient acoustic cues in spontaneous Taiwanese Mandarin, however, the parameter, F0 range expansion does not seem to be as important as duration. Besides, the vowel space formed by the F1 and F2 values of /i/, /u/, /a/ in target nouns tends to enlarge under narrow focus except for one subject. In perception, duration is the most important cue subjects use to perceive focus condition.

Acknowledgements

Firstly, I would like to express my sincere gratitude and appreciation to my thesis advisor, Ho-hsien Pan, for her teaching, advice, guidance and encouragement. This thesis would have never existed without her never-ending support. I have benefited and learned much since the first time I took her course. She introduced me into the field of phonetics, including theory of phonetics, experimental design, data analysis and also the attitude a phonetician should have when doing research.

It is my pleasure to have H. Samuel Wang and Hua-li Jian as the members of my thesis committee. They have been generous and gave me advisements and suggestions on this study with their profound linguistic knowledge.

I would also like to thank all professors in Graduate Institute of Linguistics and Cultural Studies of NCTU for their teaching to broaden my linguistic background, the classmates and faculty members of the department who have been helpful during the past three years. I am also grateful to my roommate, Alice Chang, for her encouragement and companionship.

Lastly, I would like to show the deep gratitude to my parents, brothers and my pet. Without their persistent, never-ending support, encouragement and love, I could have never achieved this goal in my life.

Contents

List of Tables

List of Figures

Chapter Ⅰ

Introduction

1.1 Motivation

Focus has been a popular topic and attracted attention for many years. Research on focus displays a wide spectrum of topics, including pragmatics, discourse, semantics, syntax, phonology to phonetics.

As to phonetic experiments, many research was done to investigate the effect of focus, however, none of the previous research of Mandarin used spontaneous speech. Most of the experiments tried to elicit data as natural as possible; subjects read sentences from a reading list with focus condition and location indicated to remind them of the designed focus pattern for answers. The acoustic characteristics could be well preserved in speech when reading, but it may not be the natural speech people use in daily conversation. Also, there was no previous research on the perception of focus in Mandarin. Moreover, besides Hsiung's (2002) study on Taiwanese Mandarin, most studies on focus in Mandarin used Pekinese Mandarin. There are differences between Taiwanese and Pekinese Mandarin. Therefore, to follow up on Hsiung's study, this author's study uses spontaneous speech in Taiwanese Mandarin. In addition, the production and perception of focus condition in spontaneous Mandarin in Taiwan will be conducted in this study.

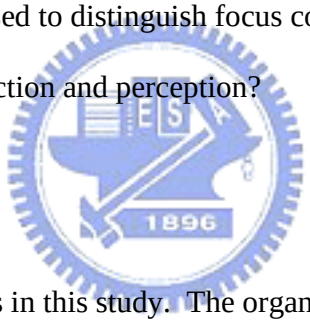
1.2 Aim of this study

There are two objectives of this study; first, the author wants to examine the phenomena of production data under different focus conditions in spontaneous Taiwanese

Mandarin, and compare if the acoustic cues for production used in spontaneous speech are the same as read speech or not. Second, the perceptual cues used by native Taiwanese Mandarin speakers to perceive focus in spontaneous speech in Taiwanese Mandarin will be investigated.

Therefore, two questions that will be discussed in this paper are:

1. What are the effect of narrow focus, that is syllables carrying new information, and neutral focus, that is syllables carrying given information, upon vowel quality, fundamental frequency (F0) expansion and duration in production data? What are the most salient cue used to distinguish focus conditions in perception?
2. Are the acoustical cues used to distinguish focus conditions in read versus spontaneous speech the same in production and perception?



1.3 Outline

There are five chapters in this study. The organization for this paper is as follows. In addition to the general introduction of motivation, aim of this study and outline of this study in Chapter Ⅰ, previous studies of focus on different linguistic aspects are introduced in Chapter Ⅱ. In Chapter Ⅲ, the method and results of production experiment are reported. Chapter Ⅳ introduces the method and results of perception experiment. Chapter Ⅴ is devoted to the discussion of the results from both production and perception experiments. Results of production and perception experiments will be used to address the questions of this study. Suggestions for further studies will also be introduced in the last chapter.

Chapter Ⅱ

Literature Review

2.1 Definition of focus conditions

There are many different definitions of the word, focus. Among all the definitions, three focus conditions are related to this study; they are narrow focus, broad focus and neutral focus separately. When the focus is placed on a single word or certain phrase, it is narrow focus. When the focus is placed on a whole utterance, it is neutral focus. When a word or a phrase does not bare narrow focus or broad focus in an utterance, it bares neutral focus. In this study, two types focus conditions, narrow focus, which carries new information and neutral focus, which carries given information will be discussed.

2.2 The function of focus

Focus has been studied and discussed in different linguistic aspects. Pragmatically, broad focus and narrow focus has been studied. According to Bolinger (1972) and Halliday (1967), the word that carried new or important information became the focus of that utterance, and attracted narrow focus or contrastive focus, while the information that the word attracted broad focus carried was not as important as the word attracted narrow focus carried. From a discourse perspective, Jakendoff (1972) introduced the notion of "focus" and "presupposition". In his definition, "focus of a sentence" meant that the speaker assumes that the information in the sentence is not shared by him and the hearer, while "presupposition of a sentence" meant that the speaker assumes that the information

in the sentence is shared by him and the hearer. Take the following two sentences for example,

(1) Is it PETER who plays guitar?

(2) No, it is BILL who plays guitar.

In sentence (1), the presupposition is that someone plays guitar. "Peter" is in narrow focus. It means that the speaker assumes the status of the constituent (or the topic of a discussion) in question to be known by the hearer, but obviously is not known by the speaker himself. In sentence (2), the presupposition is also that someone plays guitar, and "Bill" is in narrow focus. Therefore, the new information is conveyed to inform the speaker that it is "Bill" but not "Peter" who plays guitar.

In addition to pragmatics, semantically, Bolinger (1972) explained focus from the view of information status, and used pitch accent to signal focus. He reported that the word that received special status in the information structure of a sentence received pitch accent and became a focused word. He claimed that the importance of each word in a sentence depended not only on the previous context, but also on the context after. There was no systematic relationship between the structure of a sentence and its accents, which meant that sentence accents were determined by focus but not by syntax. Bolinger (1972) also believed that a speaker could choose any part of the sentence to be the point of information focus and determine the placement of accents for that sentence. Consider the sentence, "John bought a necklace for Mary." if it is the response to the question, "Who bought a necklace for Mary?" "John" was in narrow focus. However, in the question

"What did John buy for Mary?" The focus was "a necklace." He claimed that focus placement was determined purely by semantics.

However, focus is not only a pragmatic and semantic issue, but also a syntactic and phonological issue. Jakendoff (1972) reported that the rule "focus assignment" reflected the division of a sentence into focus and presupposition in syntactic structure. Selkirk (1984, 1994) described a set of projection principles to govern the relationship between focus and pitch accents. In her rules, an accented word was F-marked (Focus-marked), and the F-marking of higher constituent was projected according to the following rules:

(3) F-marking of the head of a phrase licenses F-marking of the entire phrase;
(4) F-marking of an internal argument of a head licenses the F-marking of the head.

Consider the following question and answer:

(5) What did Laura do?
(6) Laura bought a book about GARDENS.

For instance, in sentence (6), "gardens" was F-marked because it was accented. The F-marking of "gardens" licensed F-marking of the head of the prepositional phrase, "about", according to rule (4). The F-marking of "about" licensed F-marking of the entire prepositional phrase "about gardens" according to rule (3).

As opposed to Selkirk's focus projection principles, Gussenhoven (1983,1992,1994) reported focus-accent rules. His rules based on surface positions of the constituents that

expressed certain semantic roles. According to his rules, nonfocused constituents were unaccented, while all focused constituents had to be accented, including predicates, arguments, and modifiers.

Though Selkirk's focus projection and Gussenhoven's sentence accent assignment rules were different, it has been proven that focus is not only a pragmatic and semantic issue, but also a syntactic and phonological issue. In other words, in addition to syntax, focus can be realized phonologically.

The definition of focus may easily cause confusion phonologically because it can be applied to two levels – word level and sentence level. On the level of word, stress is stored and fixed in mental lexicon, i.e. word stress in English; it is invariant under different linguistic environments. However, on the level of sentences stress, it changes under different semantic and syntactic environments.

In addition, broad focus and narrow focus can be divided depending on the range they cover. Broad focus deals with the whole intonation-group, but narrow focus deals with only a certain grammatical constituent within a sentence (Crudentten, 1997; Ladd, 1980). For instance, in sentence (10), (11), (12) below, "vase" is under narrow focus in sentence (12) if it follows sentence (10). But if sentence (12) responds to sentence (11), the focus is on the entire verb phrase, "broke the vase", and it is broad focus.

(10) What did Mark break?

(11) What did Mark do?

(12) Mark broke the vase.

Therefore, narrow focus means that a certain constituent attracts focus; broad focus means that the entire sentence attracts focus, and neutral focus refers to the constituent that does not attract focus in the utterance.

2.3 Phonetic realization of focus

In addition to research of focus in the fields of pragmatics, semantics, syntax, and phonology as previously mentioned, many phonetic studies on focus have been conducted. For example, there has been research on focus effect upon vowel quality, fundamental frequency (F0) range expansion and duration in English (Weismer and Ingrisano, 1979; Cooper and Eady, 1985; Eady and Cooper, 1986; Eady et al., 1986) and other languages, such as Japanese (Maekawa, 1996), Arabic (de Jong, 2001), and Swedish (Heldner and Strangert, 2001).

Acoustic research on the effect of contrastive focus upon vowel quality in Japanese (Maekawa, 1996) reported that the vowel space - the space formed by the F1 and F2 values of vowels - of target vowels under direct focus became more peripheral, while the vowel space of vowels outside the domain of focus became less peripheral. In Arabic, lexical focus generally has the effect of increasing F1 of vowels (de Jong, 2001).

As for F0, it is recognized in Japanese that prosodic information depends primarily upon F0. The fundamental frequency (F0) could differ depending on the prosodic location of vowel in question. Maekawa (1996) found that F0 value increased due to the influence of focus. He observed that the effect of focus upon the parameter, F0 value, was omni-directional, meaning that focus influenced not only the preceding constituents but also the following ones. Studies in English showed that the fundamental frequency

(F0) range expands when a syllable is under stress. In English, the effect of narrow focus is to raise the F0 of the focused word, or to lower the F0 of post-focus words (Cooper, Eady, and Muller, 1985; Eady and Cooper, 1986; Eady et al., 1986).

Effect of focus on the parameter duration can be observed at various levels, including utterance, accentual phrase and individual segment. In Japanese, Maekawa says, "When a target phrase was focused, durations of the preceding and/or following phrases were reduced, while the duration of the target phrase stayed nearly constant" (1996). However, the situation of duration of target phrase under focus in English was different from Japanese. In English, the durational change caused by focus is an increase in duration (Cooper, Eady, and Muller, 1985; Eady and Cooper, 1986; Eady et al., 1986). In Swedish, Heldner and Strangert (2001) investigated the amount and domain of lengthening related to focal accent. They measured word duration, syllable duration, and segment duration in sentences read with the focus in different positions. It was found that the duration of words with focal accents was longer than the duration of non-focal words, and lengthening occurred within stressed syllables. Besides, by analyzing the internal structure of stressed syllables, they found that the phonologically long segments were lengthened under focus, while the phonologically short segments were not affected. Therefore, the domain of focal accent lengthening was restricted to stressed syllables in Swedish. De Jong and Zawaydeh (2002) examined how stress and focus affect the durational correlation of phonemic contrasts in Arabic vowels, and found that stress increased durations and reduced undershoot.

2.4 Mandarin

Though the effect of focus on vowel quality, F0 and duration has been found in English, Japanese, Arabic and Swedish, there have been few studies on the effect of focus in tone languages, such as Mandarin. In a tone language, F0 is used to distinguish lexical tones; for example, in Mandarin the syllable [mɑ] is phonemically different when F0 varies. For the First tone, [mɑ˥] means "mother"; for the Second tone, [mɑ˧˥] means "hemp"; [mɑ˨˩˦] means "horse" for the Third tone; and [mɑ˥˩] means "scold" for the Fourth tone.

In Mandarin, as in English, fundamental frequency (F0), intensity and duration are considered to be the three acoustic parameters affected under the focus condition. Jin (1996) examined the sentence stress (focus) in Mandarin Chinese by comparing effects of broad focus and narrow focus. Three acoustic parameters - F0, intensity and duration - of words under different stress (focus) conditions were recorded, and how focus affected these parameters was investigated. He found that duration was related to stress. A syllable became longer when it was under narrow focus. When a syllable was not under narrow focus, its relative position to stressed syllables influenced its duration: the pre-focused syllable became shorter if it was farther away from the syllable under narrow focus. Xu (1999) found that the duration of a syllable increased under narrow focus, regardless of its position in the utterance.

Hsiung (2002) investigated the effects of focus in Taiwanese Mandarin, comparing the effects of broad focus and narrow focus. She reported that duration was related to focus: the duration of target syllable became longer when under narrow focus. Duration is the most salient acoustic parameter for sentence focus in Taiwanese Mandarin.

As to the parameter F0, in Jin's (1996) research, he found that F0 was related to sentence stress. When a syllable was under narrow focus, its pitch range expanded, followed by a F0 reduction during the following syllable. Xu (1999) found that F0 range expands in different focus conditions. The F0 high points became higher, and the F0 low points became lower in non-final narrow focused words. The F0 values of post-focus words were lowered, but the F0 values of pre-focus words were unchanged. However, in her study on Taiwanese Mandarin, Hsiung (2002) reported that speakers of Taiwanese Mandarin did not always expand fundamental frequency (F0) range significantly under narrow focus, as speakers of Mandarin spoken in China did in Jin's (1996) and Xu's (1999) research.

2.5 Perception of focus

Besides production experiments, perceptual experiments play another important role in understanding the effect of focus upon different parameters because they help to examine if the acoustical cues used to distinguish focus condition are the same in production and perception. In English, perceptual experiments showed that pitch (fundamental frequency) is the most efficacious parameter, and intensity is the least. But the importance of duration varies across languages. If duration is used to cause phonemic contrast in a certain language, intensity will become the second most important parameter. If duration is not used to cause phonemic contrast, it is the second most important parameter in that language (Fry, 1955, 1958). However, there have been few studies on perception of focus in Mandarin.

In Jin's (1996) study, he considered four parameters - pitch range, top line of F0 range, duration and intensity - in perceptual experiments, and reported on the relative importance of different parameters on affecting subjects' perceptual judgments. He found that in Mandarin, pitch range was the most important cue for identifying narrow focus among the four parameters, duration was the second important cue, and intensity was the least important cue.

Chapter III

Production Experiment

3.1 Introduction

In this chapter, the method and results in the production experiment will be introduced. In the first section, method, experimental design and the process will be discussed in detail, including the linguistic background of subjects, the design and content of corpus, the instrumentation used, the recording procedure followed and the tools used in data analysis. Second, results of two different focus conditions carrying given and new information upon vowel quality, duration and F0 range will be reported, using vowel-space figures, bar-chart figures and statistical tables. At the conclusion of this chapter, a summary will be made to present what has been investigated in the production experiment.

3.2 Method

3.2.1 Subjects

Two female native Mandarin speakers, who live in Taipei City, and one male native Mandarin speaker, who live in Taipei County, participated in the production experiment. They were all born in Taipei area and have been brought up in Mandarin-speaking families. None of them can speak either Taiwanese or Hakka (the two most widely-used dialects in Taiwan). All of them were aged between 20-25 years old students in National Chiao Tung University at the time of recording. Analyses in the following sections and chapters in this study coded the two female subjects as Subject CBY and Subject CKC, and the one male subject as Subject RLY.

3.2.2 Corpus

The factors manipulated in the experiment were segmental composition of syllables, lexical tones, positions and focus conditions. All target nouns were bi-syllable nouns, with sonorant onset, such as /m, n, l, w, y, r/ and closed off by nasals or vowels. By using only resonant consonants and vowels, one can observe F0 transition between syllables. To control for lexical tones, both syllables of bisyllabic nouns contained four Mandarin lexical tones, which were First tone (high-level tone), Second tone (high-rising tone), Third tone (low-rising tone) and Fourth tone (falling tone). By matching the four tones in the words' first syllable with four tones in the words' second syllable, 16 combinations of tones were obtained for bisyllabic nouns (shown in Table3.1). In Tone3-Tone3 combination of target nouns, the third tone sandi in Mandarin occurred, therefore, the first syllable Tone3 [maˇ] was generalized as Tone2 [maˊ].

Table 3.1 16 target nouns of all possible tonal combination in Mandarin

	Tone1 (High-level tone)	**Tone2 (High-rising tone)**	**Tone3 (Low-rising tone)**	**Tone4 (Falling tone)**
Tone1 (High-level tone)	[mauˉ miˉ] "kitty"	[wuˉ meɪˊ] "dried plum"	[yɪŋˉ wuˇ] "parrot"	[laˉ miɛnˋ] "noodles"
Tone2 (High-rising tone)	[nioˊ waˉ] "frog"	[nɪnˊ məŋˊ] "lemon"	[lɔŋˊ yænˇ] "longan"	[lɪɛnˊ wuˋ] "bell fruit"
Tone3 (Low tone)	[lauˇ yɪŋˉ] "eagle"	[naiˇ yoˊ] "butter"	[maˊ yiˇ] "ant"	[ruˇ luoˋ] "cheese"
Tone4 (Falling tone)	[yɛnˋ woˉ] "bird's nest"	[yiˋ rənˊ] "beadlike grains"	[yuˋ miˇ] "corn"	[laˋ roˋ] "cured meat"

Two types of utterances were designed as follow. In the first type of utterance, the target noun was either in initial position or final position of the utterance, as shown in Table 3.2. In the second type, the target noun appeared in middle position of the utterance (Table 3.2). By placing 16 target nouns in initial, middle, and final positions, 48 sentences were derived. Table 3.2 shows the positions of target nouns. For example, for nouns in initial and final position of an utterance ('XX' represents target nouns), the structure of the sentences was "'XX' [tə˥] [yo˨ piɛn˥]/[tsʊo˧ piɛn˥]/[çaŋ˨ miɛn˨]/[çia˨ miɛn˨] [çɪ˨] 'XX.'" 'XX' was in the right/left/top/bottom side of 'XX.' For nouns in the middle position of an utterance, the structure of the sentence was "'XX' [tsai˨] 'XX' [tə˥] [yo˨ piɛn˥]/[tsʊo˧ piɛn˥]/[çaŋ˨ miɛn˨]/[çia˨ miɛn˨]" 'XX' was in the right side of 'XX.' The initial nouns were not analyzed in this sentence structure.

Table 3.2 Positions of target nouns (Target syllables are underlined).

Position of target syllables	**Answers**
Initial	[<u>mau˥ mi˥</u> tə˥ yo˨ piɛn˥ çɪ˨ ma˦ yi˧] "Ant is on the right side of kitty."
Middle	[ma˦ yi˧ tsai˨ <u>mau˥ mi˥</u> tə˥ yo˨ piɛn˥] "Ant is on the right side of kitty."
Final	[ma˦ yi˧ tə˥ tsuo˧ piɛn˥ çɪ˨ <u>mau˥ mi˥</u>] "Kitty is on the left side of ant."

In addition, the control for target nouns under narrow focus carrying new information, or under neutral focus carrying given information was also required in this experiment. Each utterance would respond to two precursor questions. Take the utterance [mau˥ mi˥ tə˥ yo˨ piɛn˥ çɪ˨ ma˦ yi˧] "Ant is on the right side of kitty", for

example. It can be the answer to two different questions, i.e.

[mau˧ mi˧ tə˧ yoˬ piɛn˧ çɪˬ çəˊ mə˧] “What is on the right side of kitty?” and

[çəˊ mə˧ tə˧ yoˬ piɛn˧ çɪˬ maˊ yiˇ] “Ant is on the right side of what?” In response to the first question, the target noun, [mau˧ mi˧] “kitty” was in neutral focus, carrying given information in utterance, while in response to the second question, the target noun, [mau˧ mi˧] “kitty” was in narrow focus, carrying new information (shown in Table 3.3).

Table 3.3 Focus patterns (target nouns are underlined).

Focus condition	Answers	Precursor questions
Neutral focus (given information)	[mau˧ mi˧ tə˧ yoˬ piɛn˧ çɪˬ maˊ yiˇ] “Kitty is on the right side of ant.”	[mau˧ mi˧ tə˧ yoˬ piɛn˧ çɪˬ çəˊ mə˧] “What is on the right side of kitty?”
Narrow focus (new information)	[mau˧ mi˧ tə˧ yoˬ piɛn˧ çɪˬ maˊ yiˇ] “Kitty is on the right side of ant.”	[çəˊ mə˧ tə˧ yoˬ piɛn˧ çɪˬ maˊ yiˇ] “Ant is on the right side of what?”

Take the utterance [maˊ yiˇ tsaiˬ mau˧ mi˧ tə˧ yoˬ piɛn˧] “Ant is on the right side of kitty”, for example; it can answer to two different questions, i.e.

[maˊ yiˇ tsaiˬ çəˊ mə˧ tə˧ yoˬ piɛn˧] “Ant is on the right side of what?” and

[çəˊ mə˧ tsaiˬ mau˧ mi˧ tə˧ yoˬ piɛn˧] “What is on the right side of kitty?” If the utterance was in response to the first question, the target noun, [mau˧ mi˧] “kitty” was in narrow focus, carrying new information; but if it responded to the second question, the target noun, [mau˧ mi˧] “kitty” was in neutral focus, carrying given information. Table

3.4 indicates the design of precursor questions and answers matching different positions and focus conditions.

Table 3.4 Positions, focus conditions with matched precursor questions and answers.

Position	**Focus**	**Precursor Question**	**Answer**
Initial	Neutral	[mauꜜ miꜜ təꜜ yoꜘ piɛnꜜ çiꜘ çəꜗ məꜜ] “What is on the right side of kitty?”	[mauꜜ miꜜ təꜜ yoꜘ piɛnꜜ çiꜘ maꜗ yiꜛ] “Kitty is on the right side of ant.”
Initial	Narrow	[çəꜗ məꜜ təꜜ yoꜘ piɛnꜜ çiꜘ maꜗ yiꜛ] “Ant is on the right side of what?”	[mauꜜ miꜜ təꜜ yoꜘ piɛnꜜ çiꜘ maꜗ yiꜛ] “Kitty is on the right side of ant.”
Middle	Neutral	[çəꜗ məꜜ tsaiꜘ mauꜜ miꜜ təꜜ yoꜘ piɛnꜜ] “What is on the right side of kitty?”	[maꜗ yiꜛ tsaiꜘ mauꜜ miꜜ təꜜ yoꜘ piɛnꜜ] “Ant is on the right side of kitty.”
Middle	Narrow	[maꜗ yiꜛ tsaiꜘ çəꜗ məꜜ təꜜ yoꜘ piɛnꜜ] “Ant is on the right side of what?”	[maꜗ yiꜛ tsaiꜘ mauꜜ miꜜ təꜜ yoꜘ piɛnꜜ] “Ant is on the right side of kitty.”
Final	Neutral	[çəꜗ məꜜ təꜜ tsuoꜛ piɛnꜜ çiꜘ mauꜜ miꜜ] “Kitty is on the left side of what?”	[maꜗ yiꜛ təꜜ tsuoꜛ piɛnꜜ çiꜘ mauꜜ miꜜ] “Kitty is on the left of ant.”
Final	Narrow	[maꜘ yiꜘ təꜜ tsuoꜛ piɛnꜜ çiꜘ çəꜗ məꜜ] “What is on the left side of kitty?”	[maꜗ yiꜛ təꜜ tsuoꜛ piɛnꜜ çiꜘ mauꜜ miꜜ] “Kitty is on the left of ant.”

Each sentence with the same focus condition and position in the experiment was repeated three times by repeating the precursor questions for three times in random order. All together, 288 sentences (4 (tone)* 4 (tone) * 2 (focus condition) * 3 (position) * 3 (repetition) =288) were produced in randomized order by each subject.

3.2.3 Instrumentation

Microphones produced by Uni-Directional Dynamic Microphone, model TEV TM-728Ⅱ, headphones produced by Grado Prestige Series Headphones, model SR80, and for recording equipment, the MD recorder produced by Sony, model MZ-R4ST, were used for recording in this experiment. Emu software was used for data analysis.

3.2.4 Recording procedure

Recording was conducted at the Acoustic Lab in the Department of Foreign Languages and Literatures, National Chiao Tung University. A microphone was placed about 35cm in front of the subject's mouth. The experimenter and subject were in separate rooms.

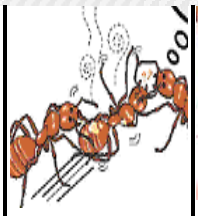
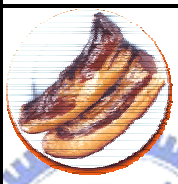
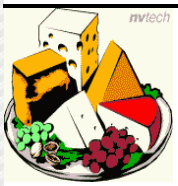
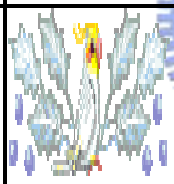
A reading list with precursor questions and answers typed in Chinese scripts and the picture (Table 3.5) was presented to the experimenter. But to subjects, only the picture (Table 3.5) was presented. Subjects were instructed to answer aloud to all precursor questions. The reason for showing subjects only the picture instead of the typed Chinese list with IPA was to ensure the production data acquired from this experiment were more spontaneous and natural, like daily conversation.

The experimenter initiated the recording process by asking the precursor question, and then waited for the subject's answer. Subjects had to determine the focus placement from precursor questions and answer according to the picture. No instructions on where to place the focus were given, because the data of responses were intended to be as spontaneous as possible. There was a 10-minute break every 30 minutes.

During recording, when the experimenter judged that a particular sentence was not produced properly, the precursor question was repeated again, and subjects were asked to repeat the answer. For example, for the question [mau˧ mi˧ tə˧ yo˨ piɛn˧ çɪ˨ çəˊ mə˧] "What is on the right side of kitty?" the correct answer was [mau˧ mi˧ tə˧ yo˨ piɛn˧ çɪ˨ maˊ yiˇ] "Kitty is on the right side of ant." If subjects answered [mau˧ mi˧ tə˧ yo˨ piɛn˧ çɪ˨ la˧ miɛn˨] "Kitty is on the right side of noodles."

or [mau˥ mi˥ tə˧ tsuoˏ piɛn˥ çɪˋ maˊ yiˏ] “Kitty is on the left side of ant”, experimenter repeated the precursor question until the answer was correct.

Table 3.5 The picture presented to experimenter and subjects.

 [mau˥ mi˥] ‘kitty’	 [maˊ yiˏ] ‘ant’	 [la˥ miɛnˋ] ‘noodles’	 [naiˏ yoˊ] ‘butter’
 [yuˋ miˏ] ‘corn’	 [lɪɛnˊ wuˋ] ‘bell fruit’	 [laˋ roˋ] ‘cured meat’	 [nioˊ wa˥] ‘frog’
 [ruˏ luoˋ] ‘cheese’	 [yɪŋ˥ wuˏ] ‘parrot’	 [yiˋ rən˥] ‘beadlike grains’	 [wu˥ meɪˊ] ‘dried plum’
 [nɪmˊ məŋˊ] ‘lemon’	 [yɛnˋ wo˥] “bird’s nest”	 [lɔŋˊ yænˏ] ‘longan’	 [lauˏ yɪŋ˥] ‘eagle’

3.2.5 Data analysis

The speech signals were recorded in MD and digitized onto the hard disk of a Windows2000 computer by Creative Wave Studio with the sampling rate of 22.5kHz. The F0, waveform, and spectrogram for each of the 288 sentences were generated and

then annotated using Emu Software (Figure1). Hand-editing was done to label the segmentation of Romazi level, tone level and focus level. For Romazi level, the utterance was segmented by syllable with romanization tagging transcribed for each syllable. For tone level, the surface tone of each syllable was labeled. For focus level, target syllables were labeled with their focus condition, with “nf” representing narrow focus, and “neut” representing neutral focus.

Data queries of duration, F0, F1, and F2 values were done after segmentation. Durations of target syllables under narrow and neutral focus in different positions were calculated by the query command of Emu. We used the command to query each syllable of target nouns, for example, the first syllable [mauˊ] and the second syllable [miˊ] of the noun [mauˊ miˊ] “kitty.” Results showed the starting point and end point of each syllable of target nouns under narrow focus and neutral focus. Therefore, duration of the target nouns under two conditions was calculated.

F0 values at 5%, 25%, 50%, 75% and 95% interval points in target syllables were queried. Maximum and minimum F0 value were taken to calculate F0 range of target syllables under narrow and neutral focus. F1 and F2 values of three vowels /a/, /i/ and /u/ in target syllables under narrow and neutral focus were queried too. An example of the segmentation is shown in Figure 3.1 below.

The results of this experiment were largely based on statistical analysis of various values taken from the query results of Emu Software. The statistics used in this chapter are the repeated analysis of variance (ANOVA), used to analyze the effect of focus and position on duration and F0 range. The multiple post-hoc Tukey comparison was used to analyze significant main effect. The multivariate analysis of variance (MANOVA) was

used to analyze the effect of focus on vowel quality. In this study, the level of significance was set at 0.05, with a 95% confidence interval.

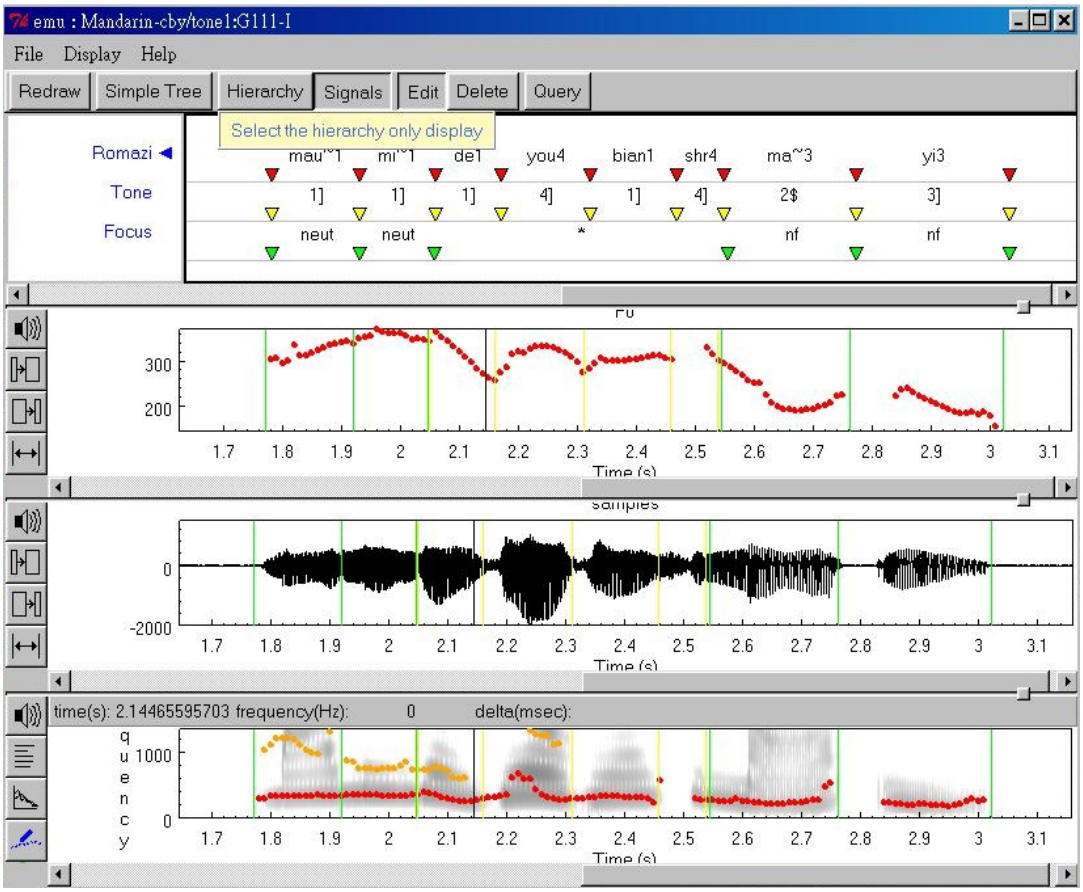


Figure 3.1 Emu Software.

3.3 Results

In this section, the results of recording and statistical analyses upon the three aspects - vowel quality, F0 range and duration - will be reported accordingly.

3.3.1 Vowel quality

Table 3.6 summarizes the results of one-way repeated measure ANOVA (focus)

with focus condition as the independent factor analyzing F1 and F2 separately. A two-dimensional MANOVA (focus) with focus condition, narrow versus neutral, as the independent factor was used to analyze the two-dimensional mean vector of F1 and F2.

As shown in Table 3.6, for subject CBY, no significant differences were found on the effect of focus on F1 or F2 of /i, u, a/ for ANOVA and MANOVA. For subject CKC, the significant effect of focus was found in F2 of /a/ from target syllables, and a significant effect of focus was found on MANOVA analyzing of F1 and F2. For subject RLY, the only significant effect of focus was found in F2 of /u/ from target nouns.

Table 3.6 Statistical test on the effect of focus upon target vowels (N stands for the number of vowels).

Subject	**Vowel**	**Information**	**N.**	**Separate ANOVA**		**MANOVA F1&F2**
1. CBY	/i/	G	18	F=2.33	F=0.60	F=1.32
		N	18	p=0.14$^{n.s.}$	p=0.45$^{n.s.}$	p=0.28$^{n.s.}$
	/u/	G	18	F=1.70	F=1.65	F=1.04
		N	18	p=0.20$^{n.s.}$	p=0.21$^{n.s.}$	p=0.37$^{n.s.}$
	/a/	G	18	F=0.35	F=0.28	F=0.29
		N	18	p=0.56$^{n.s.}$	p=0.60$^{n.s.}$	p=0.75$^{n.s.}$
2. CKC	/i/	G	18	F=0.00	F=1.66	F=0.90
		N	18	p= 0.98$^{n.s.}$	p= 0.21$^{n.s.}$	p= 0.42$^{n.s.}$
	/u/	G	18	F=0.00	F=1.59	F=1.34
		N	18	p= 0.99$^{n.s.}$	p= 0.22$^{n.s.}$	p= 0.28$^{n.s.}$
	/a/	G	18	F=1.42	F=11.63	F=6.51
		N	18	p= 0.24$^{n.s.}$	p= 0.00**	p= 0.00**
3. RLY	/i/	G	18	F=0.31	F=0.00	F=0.16
		N	18	p=0.58$^{n.s.}$	p=0.93$^{n.s.}$	p=0.85$^{n.s.}$
	/u/	G	18	F=0.11	F=4.88	F=2.38
		N	18	p=0.75$^{n.s.}$	p=0.03*	p=0.11$^{n.s.}$
	/a/	G	18	F=0.12	F=0.37	F=0.18
		N	18	p=0.74$^{n.s.}$	p=0.54$^{n.s.}$	p=0.84$^{n.s.}$

* $p < 0.05$

Vowel space (the space formed by the F1 and F2 values of vowel /i/, /u/, /a/) was calculated as following: X_i stands for the F1 value of vowel /i/, Y_i stands for the F2 value

of vowel /i/, X_2 stands for the F1 value of vowel /u/, Y_2 stands for the F2 value of vowel /u/, X_3 stands for the F1 value of vowel /a/, and Y_3 stands for the F2 value of vowel /a/.

$$\frac{\left| (X_2 - X_1)(Y_3 - Y_1) - (X_3 - X_1)(Y_2 - Y_1) \right|}{2}$$

Figure 3.2 – Figure 3.7 below show the vowel space and mean value of F1 and F2 of target vowels /i/, /u/ and /a/ in first and second syllable for three subjects separately. It was found that the vowel space from vowels under narrow focus was larger than its counterpart under neutral focus for subject CBY and subject CKC, but not subject RLY.

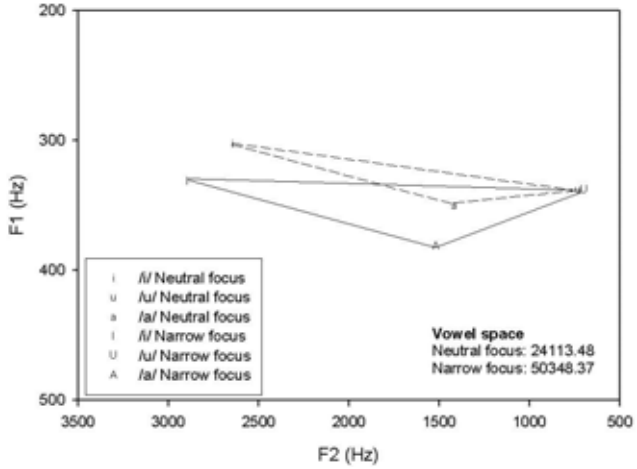


Figure 3.2 Comparison of mean formant frequency of the target vowels as function of focus conditions in first syllable for subject CBY. Capital and small letters correspond to target vowels under narrow and neutral focus. Gray line and short-dashed line encircle the vowel space of target vowels extract from nouns under narrow and neutral focus respectively.

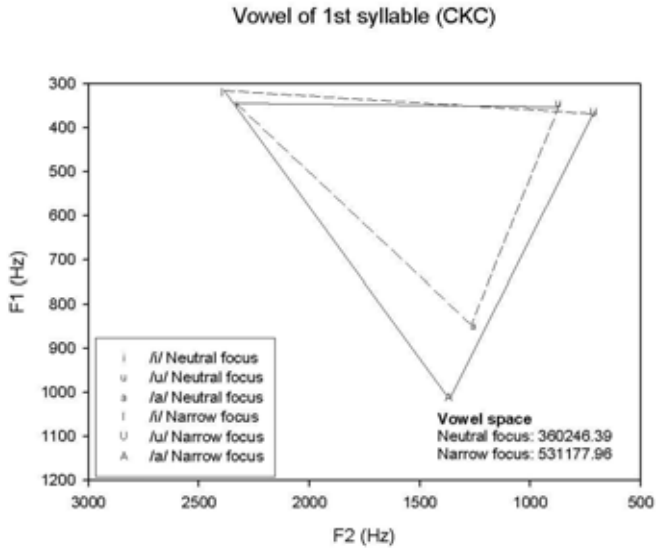


Figure 3.3 Comparison of mean formant frequency of the target vowels as function of focus conditions in first syllable for subject CKC. Capital and small letters correspond to target vowels under narrow and neutral focus. Gray line and short-dashed line encircle the vowel space of target vowels extract from nouns under narrow and neutral focus respectively.

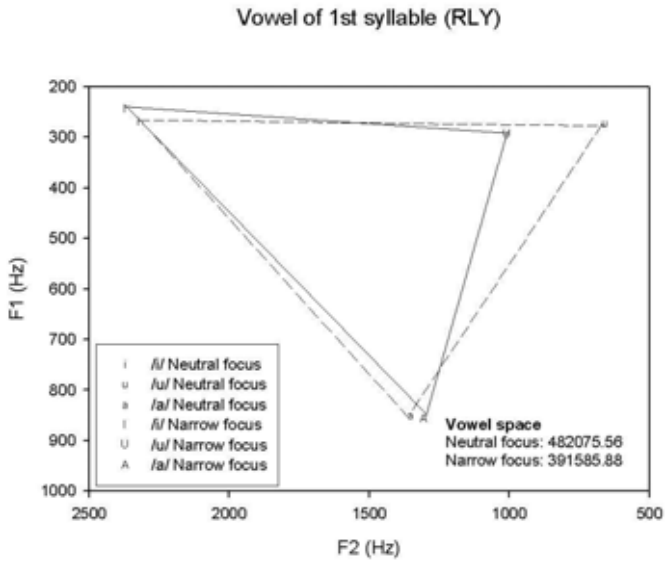


Figure 3.4 Comparison of mean formant frequency of the target vowels as function of focus conditions in first syllable for subject RLY. Capital and small letters correspond to target vowels under narrow and neutral focus. Gray line and short-dashed line encircle the vowel space of target vowels extract from nouns under narrow and neutral focus respectively.

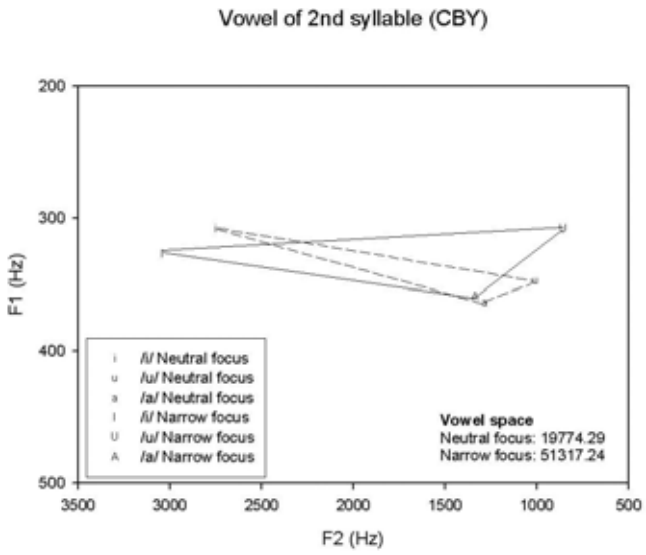


Figure 3.5 Comparison of mean formant frequency of the target vowels as function of focus conditions in second syllable for subject CBY. Capital and small letters correspond to target vowels under narrow and neutral focus. Gray line and short-dashed line encircle the vowel space of target vowels extract from nouns under narrow and neutral focus respectively.

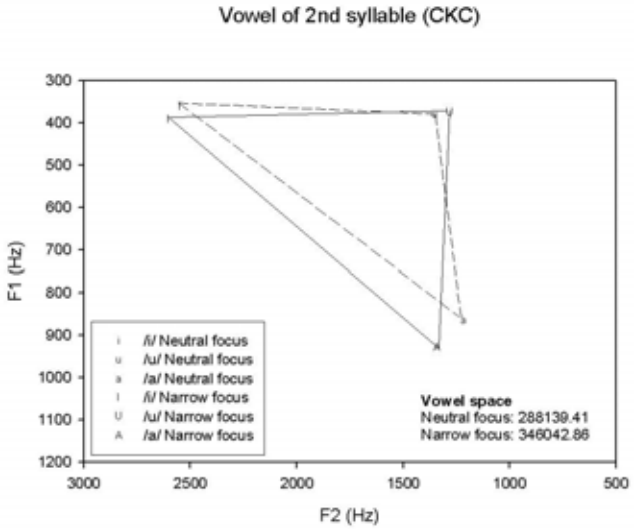


Figure 3.6 Comparison of mean formant frequency of the target vowels as function of focus conditions in second syllable for subject CKC. Capital and small letters correspond to target vowels under narrow and neutral focus. Gray line and short-dashed line encircle the vowel space of target vowels extract from nouns under narrow and neutral focus respectively.

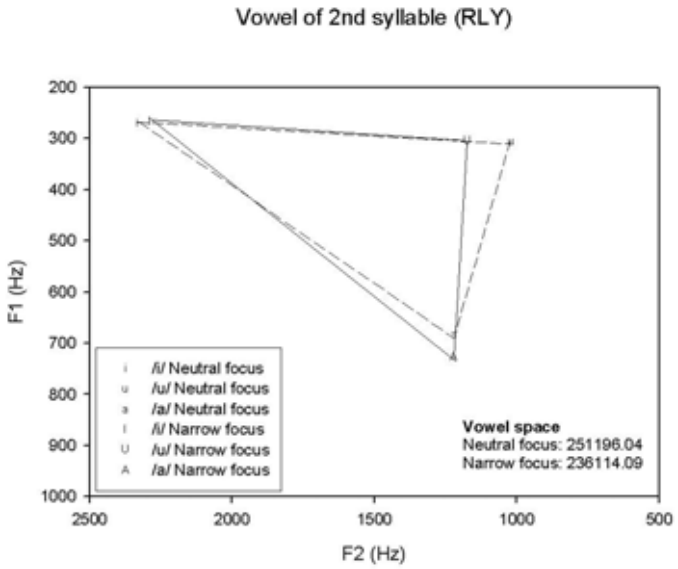


Figure 3.7 Comparison of mean formant frequency of the target vowels as function of focus conditions in second syllable for subject RLY. Capital and small letters correspond to target vowels under narrow and neutral focus. Gray line and short-dashed line encircle the vowel space of target vowels extract from nouns under narrow and neutral focus respectively.

3.3.2 Duration

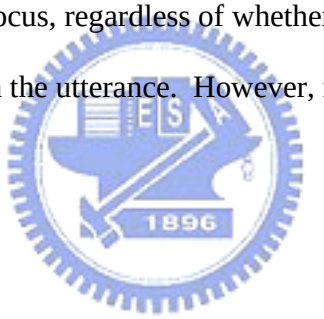
In this section, we will examine the duration of the target syllables under narrow focus and neutral focus, and the duration of target syllables in different positions. We begin by analyzing the variance of the data, and the statistical results will be reported accordingly. Two-way repeated measure ANOVA was performed to investigated if the interaction effect of the two parameters, focus and position, upon duration of target syllables was significantly. If the interaction effect was not significant, the main effect of each parameter was examined by one-way repeated measure ANOVA. The first one-way repeated measure ANOVA was to investigate if the duration of target syllables under narrow focus was significantly longer than the duration of target syllables under neutral focus. The second one-way repeated measure ANOVA was used to examine if the duration of target syllables in different positions of utterances were significantly longer or

shorter than one another.

First, two-way repeated measure ANOVA to examine the duration of target syllables under different focus conditions and positions was performed. Results showed that for subject CBY, the interaction effect of these two independent factors was significant ($F_{(2,190)}$=4.23, p<0.05). Further analyses showed that the duration of target syllables under narrow focus was significantly longer than that under neutral focus in initial position ($F_{(1,95)}$=17.56, p<0.05) and middle position ($F_{(1,95)}$=37.93, p<0.05), but not in final position ($F_{(1,95)}$=2.37, p>0.05). For subject CKC, the interaction effect was not significant ($F_{(2,190)}$=1.06, p>0.05). Therefore, the main effect of each parameter upon the duration of target syllables was examined. Results showed that the durations of target syllables under narrow focus were significantly longer than target syllables under neutral focus ($F_{(1,287)}$=8.38, p<0.05). Also durations were significantly different when the target syllables were in different positions ($F_{(2,382)}$=438.43, p<0.05). Duration of target syllables in middle position was significantly longer than duration of target syllables in initial position (p<0.05). Duration of target syllables in final position was significantly longer than duration of target syllables in middle position (p<0.05). The Duration of target syllables in final position was significant longer than duration of target syllables in initial position (p<0.05). For subject RLY, the interaction effect of these two independent factors, focus conditions and positions, was significant ($F_{(2,190)}$=10.32, p<0.05). Further analyses showed that the duration of target syllables under narrow focus was significantly longer than that under neutral focus in initial position ($F_{(1,95)}$=11.81, p<0.05) and middle position ($F_{(1,95)}$=25.15, p<0.05), but not in final position ($F_{(1,95)}$=1.56, p>0.05).

Therefore, our initial conclusion for the parameter duration is that duration is highly related to the focus condition in spontaneous speech in Taiwanese Mandarin. This is because both subject CBY and RLY's results of two-way repeated ANOVAs were significant, and for subject CKC, the results of one-way repeated ANOVA that analyzed duration were significant. Besides, duration was related to the position in utterances as well studied in previous literatures.

Additionally, figures of duration in first and second syllables of four tones under narrow and neutral focus are attached in the Appendix A. According to the figures, we found the trend that durations of target syllables under narrow focus are longer than their counterparts under neutral focus, regardless of whether they were in the first or second syllable, or their positions in the utterance. However, not all differences were statistically significant.



3.3.3 F0 Range

In this section, we will examine the F0 range of the target syllables under narrow focus and neutral focus, and the F0 range of target syllables in different positions. We begin by analyzing the variance of the data, and the statistical results will be reported accordingly. Two-way repeated measure ANOVA was performed to investigated if the interaction effect of the two parameters, focus and position, upon F0 range of target syllables was significantly. If the interaction effect was not significant, the main effect of each parameter was examined by one-way repeated measure ANOVA. The first one-way repeated measure ANOVA was to investigate if the F0 range of target syllables under narrow focus was significantly larger than the F0 range of target syllables under neutral

focus. The second one-way repeated measure ANOVA was used to examine if the F0 range of target syllables in different positions of utterances were significantly longer or shorter than one another.

First, two-way repeated measure ANOVA to examine the F0 range of target syllable under different focus conditions and positions was performed. Results showed that for all subjects, the interaction effect of these two independent factors was not significant (Subject CBY: $F_{(2,190)}=0.93$, $p>0.05$), (Subject CKC: $F_{(2,190)}=2.17$, $p>0.05$), (Subject RLY: $F_{(2,190)}=1.12$, $p>0.05$). Therefore, the main effect of the two parameters, focus and position, upon F0 range of target syllables was examined by one-way repeated measure individually.

First, the one-way repeated measure ANOVA was performed to analyze F0 range under narrow focus and neutral focus. Results showed that for subject CBY, the F0 range values of target syllables under narrow focus were significantly higher than the F0 range values under neutral focus ($F_{(1,287)}=8.08$, $p<0.05$). For subject CKC, F0 ranges were not significantly different under narrow focus and neutral focus ($F_{(1,287)}=2.75$, $p=0.10$). For subject RLY, F0 ranges of target syllables under narrow focus were significantly larger than F0 range values under neutral focus ($F_{(1,287)}=22.96$, $p<0.05$).

One-way repeated measure ANOVA was performed to analyze F0 range values in different positions. Results showed that for subject CBY, the F0 ranges of target syllables in different positions were significantly different ($F_{(2,382)}=243.50$, $p<0.05$). For subject CKC, F0 ranges were significantly different when target syllables were in different positions ($F_{(2,382)}=49.98$, $p<0.05$). For subject RLY, the F0 ranges of target syllables in different positions were significantly different ($F_{(2,382)}=44.55$, $p<0.05$). The significant

differences of mean F0 ranges for all subjects were further examined through multiple post-hoc comparisons.

Results of post-hoc comparison showed that, for subject CBY, F0 ranges of target syllables in middle position were significantly larger than F0 ranges of target syllables in initial position. ($p<0.05$). F0 ranges of target syllables in final position were significantly larger than F0 ranges of target syllables in middle position ($p<0.05$). F0 ranges of target syllables in middle positions were larger than F0 ranges of target syllables in initial position ($p<0.05$). For subject CKC, F0 ranges of target syllables in middle position were significantly larger than F0 ranges of target syllables in initial position ($p<0.05$). F0 ranges of target syllables in final position were significantly larger than F0 ranges of target syllables in middle position ($p<0.05$). F0 ranges of target syllables in final position were significantly larger than F0 ranges of target syllables in initial position ($p<0.05$). For subject RLY, F0 ranges of target syllables in initial and middle position were not significantly different ($p>0.05$). F0 ranges of target syllables in final position were significantly larger than F0 ranges of target syllables in middle position ($p<0.05$). F0 ranges of target syllables in final position were significant larger than F0 ranges of target syllable in initial position ($p<0.05$). It was found that F0 ranges of target syllables in final positions were significantly larger than F0 ranges of target syllables in middle position; F0 ranges of target syllables in final position were significantly larger than F0 ranges of target syllables in initial position, while the differences of F0 ranges of target syllables between initial and middle positions were not significant for subject RLY.

Thus, our initial conclusion for the F0 range parameter was that F0 range was related to the focus condition and position in spontaneous speech in Taiwanese Mandarin, except

for subject CKC, whose data showed that F0 ranges of target syllables under narrow focus and neutral focus were not significantly different. Besides, bar-chart figures of F0 ranges in first and second syllables of four tones under narrow focus and neutral focus are reported below in Figure 3.8- Figure 3.43. A trend could be observed that the means of F0 ranges of target syllables under narrow focus tended to be larger than those counterparts under neutral focus, though not all of the differences reached statistical significance. Some exceptions are shown in the bar-chart figures: first syllable in Figure3.8, second syllable in Figure 3.9, first syllable in Figure 3.10, first syllable in Figure 3.13, first syllable in Figure 3.14, second syllable in Figure 3.17, first syllable in Figure 3.18, first syllable in Figure 3.24, second syllable in Figure 3.25, first syllable in Figure 3.23, second syllable in Figure 3.27, first and second syllable in Figure 3.28, first and second syllable in Figure 3.32, first syllable in Figure 3.36, first and second syllable in Figure 3.39, first syllable in Figure 3.40, first syllable in Figure 3.41 and second syllable in Figure 3.42.

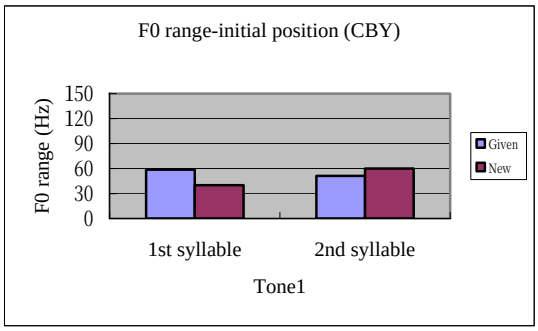


Figure 3.8 F0 ranges in first and second syllables of tone1 (high-level tone) in initial position for subject CBY.

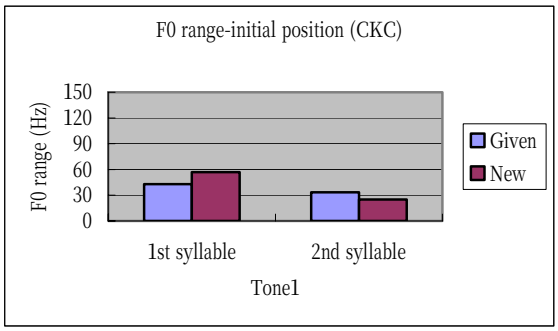


Figure 3.9 F0 ranges in first and second syllables of tone1 (high-level tone) in initial position for subject CBY.

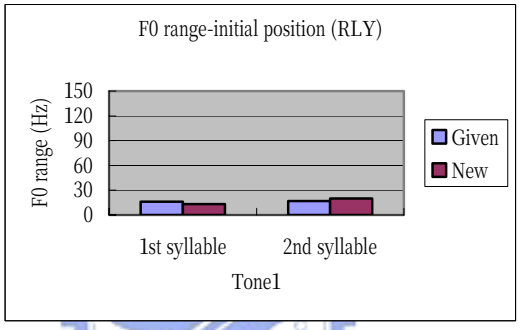


Figure 3.10 F0 ranges in first and second syllables of tone1 (high-level tone) in initial position for subject RLY.

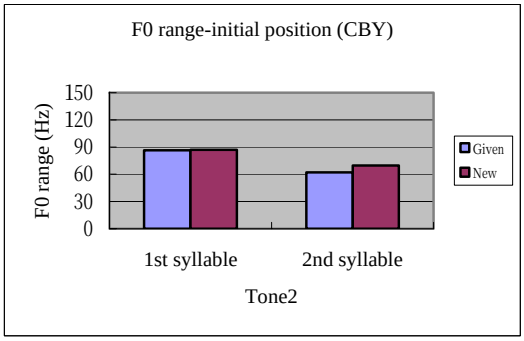


Figure 3.11 F0 ranges in first and second syllables of tone2 (high-rising tone) in initial position for subject CBY.

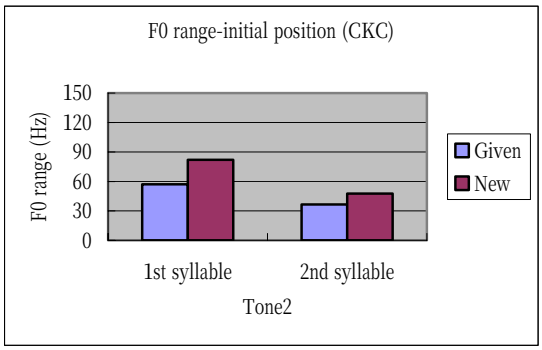


Figure 3.12 F0 ranges in first and second syllables of tone2 (high-rising tone) in initial position for subject CKC.

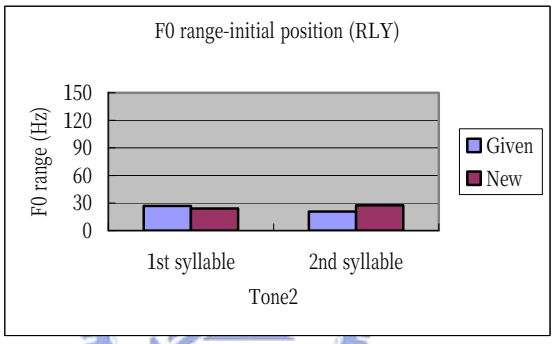


Figure 3.13 F0 ranges in first and second syllables of tone2 (high-rising tone) in initial position for subject RLY.

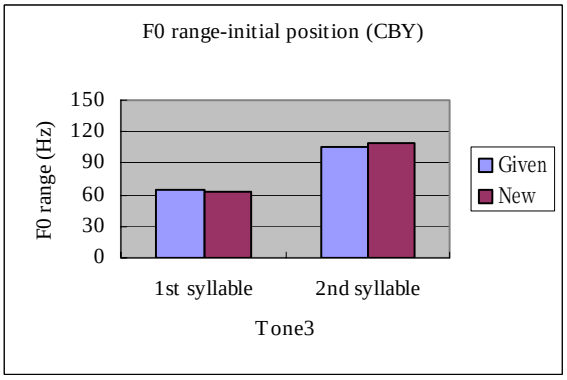


Figure 3.14 F0 ranges in first and second syllables of tone3 (low tone) in initial position for subject CBY.

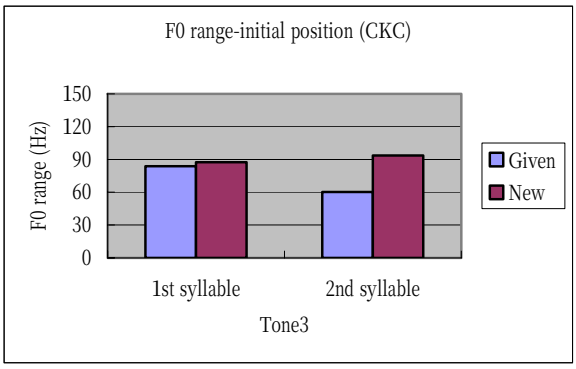


Figure 3.15 F0 ranges in first and second syllables of tone3 (low tone) in initial position for subject CKC.

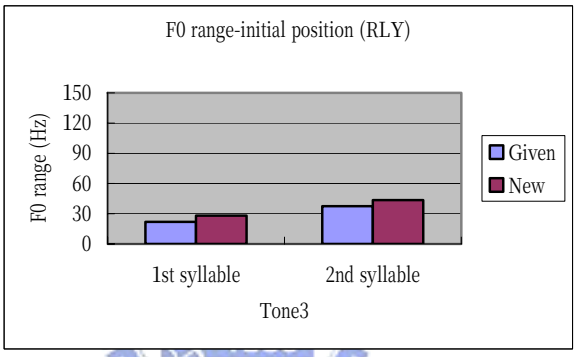


Figure 3.16 F0 ranges in first and second syllables of tone3 (low tone) in initial position for subject RLY.

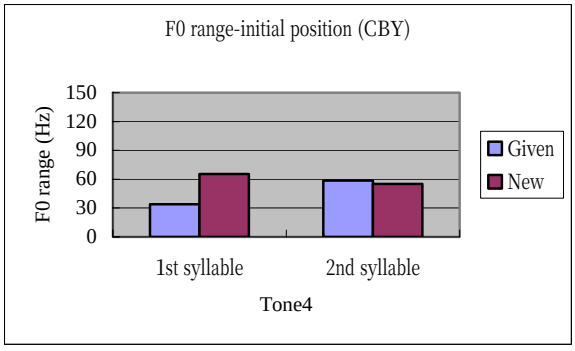


Figure 3.17 F0 ranges in first and second syllables of tone4 (falling tone) in initial position for subject CBY.

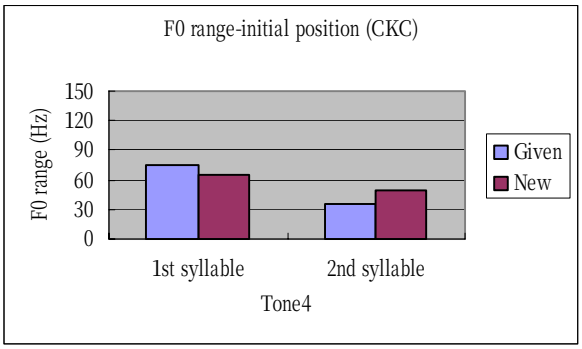


Figure 3.18 F0 ranges in first and second syllables of tone4 (falling tone) in initial position for subject CKC.

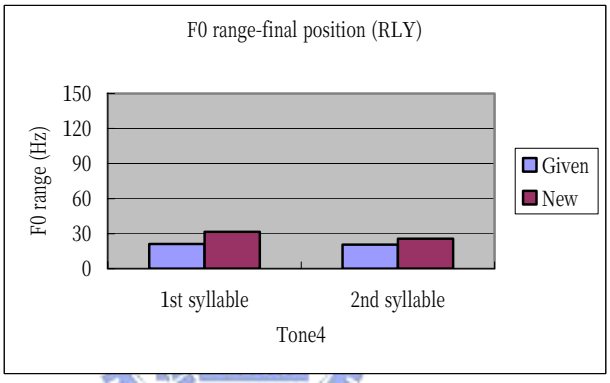


Figure 3.19 F0 ranges in first and second syllables of tone4 (falling tone) in initial position for subject RLY.

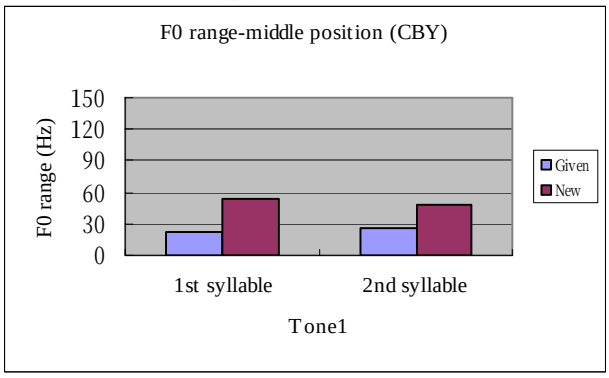


Figure 3.20 F0 ranges in first and second syllables of tone1 (high-level tone) in middle position for subject CBY.

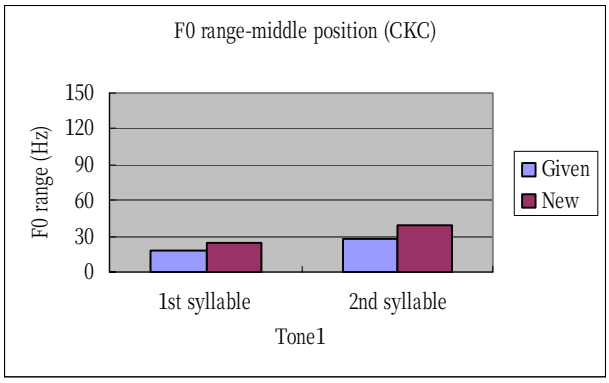


Figure 3.21 F0 ranges in first and second syllables of tone1 (high-level tone) in middle position for subject CKC.

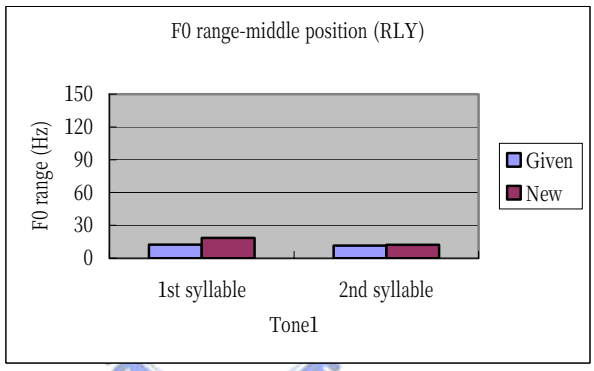


Figure 3.22 F0 ranges in first and second syllables of tone1 (high-level tone) in middle position for subject RLY.

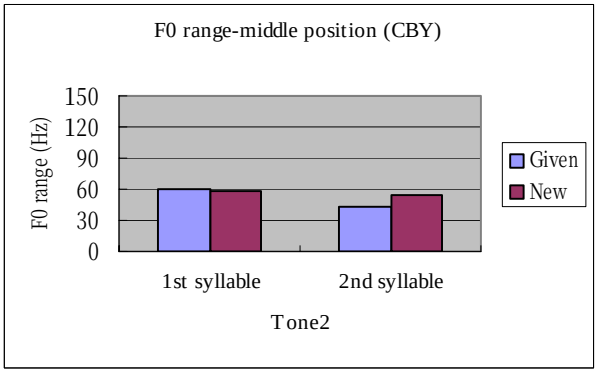


Figure 3.23 F0 ranges in first and second syllables of tone2 (high-rising tone) in middle position for subject CBY.

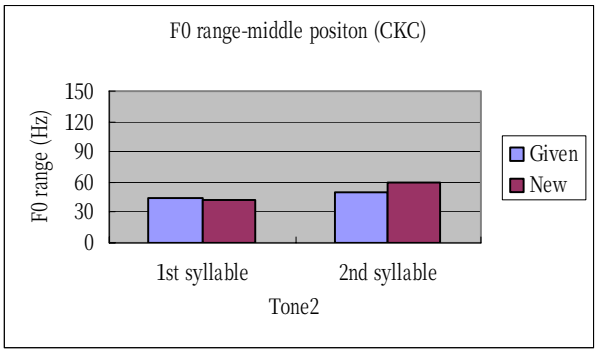


Figure 3.24 F0 ranges in first and second syllables of tone2 (high-rising tone) in middle position for subject CKC.

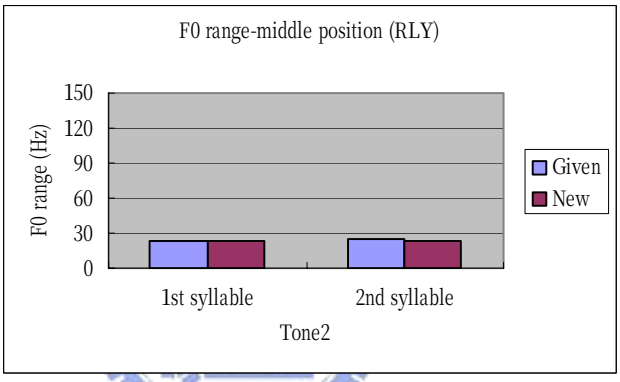


Figure 3.25 F0 ranges in first and second syllables of tone2 (high-rising tone) in middle position for subject RLY.

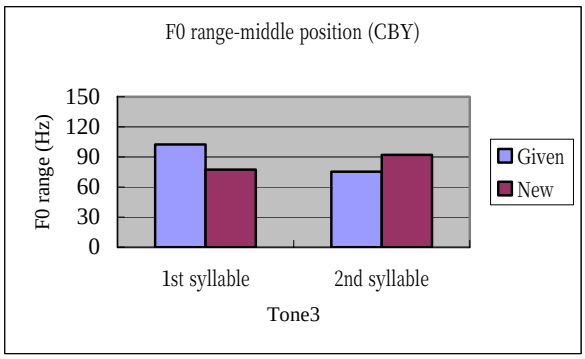


Figure 3.26 F0 ranges in first and second syllables of tone3 (low tone) in middle position for subject CBY.

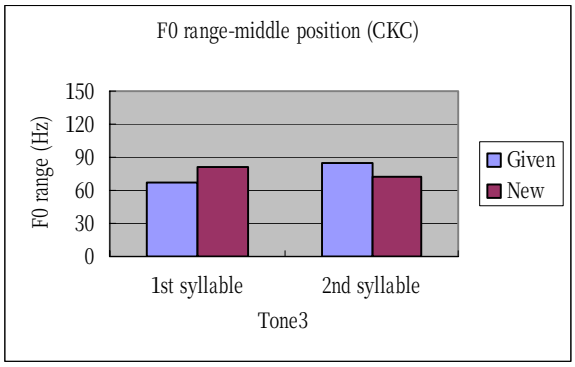


Figure 3.27 F0 ranges in first and second syllables of tone3 (low tone) in middle position for subject CKC.

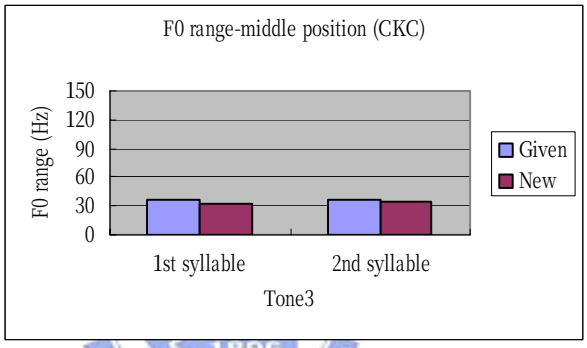


Figure 3.28 F0 ranges in first and second syllables of tone3 (low tone) in middle position for subject RLY.

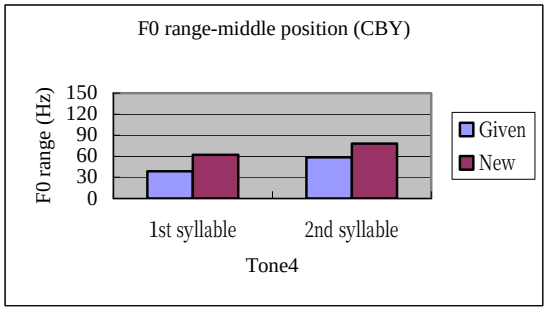


Figure 3.29 F0 ranges in first and second syllables of tone4 (falling tone) in middle position for subject CBY.

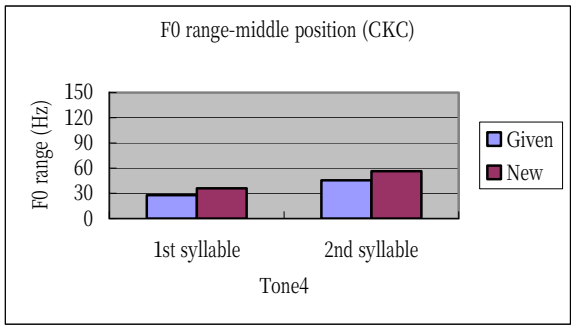


Figure 3.30 F0 ranges in first and second syllables of tone4 (falling tone) in middle position for subject CKC.

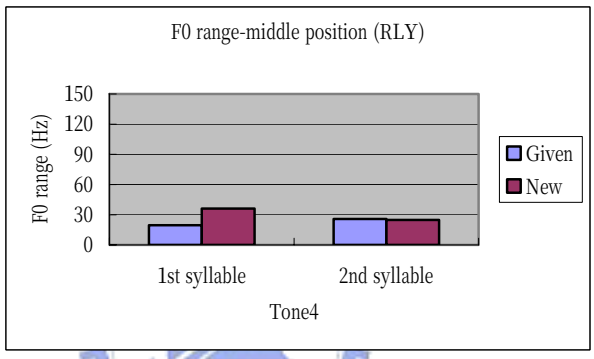


Figure 3.31 F0 ranges in first and second syllables of tone4 (falling tone) in middle position for subject RLY.

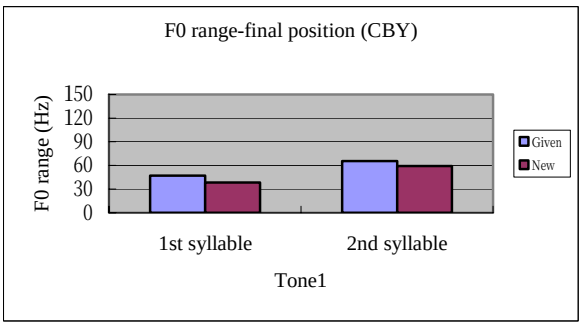


Figure 3.32 F0 ranges in first and second syllables of tone1 (high-level tone) in final position for subject CBY.

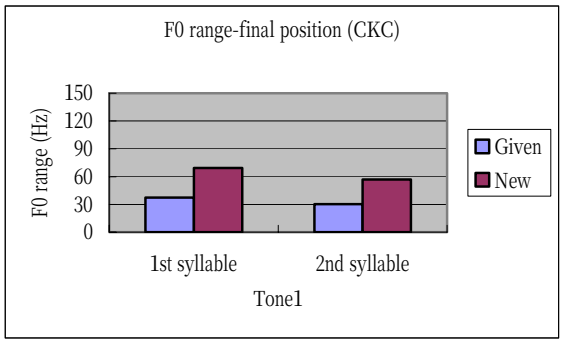


Figure 3.33 F0 ranges in first and second syllables of tone1 (high-level tone) in final position for subject CKC.

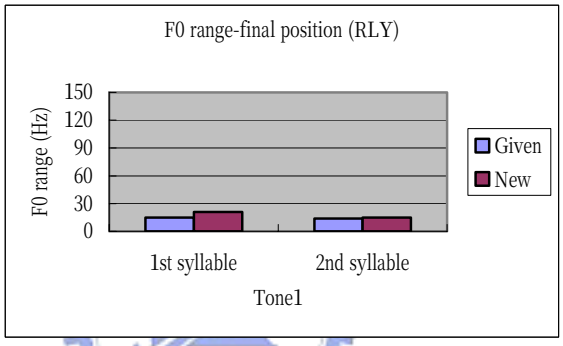


Figure 3.34 F0 ranges in first and second syllables of tone1 (high-level tone) in final position for subject RLY.

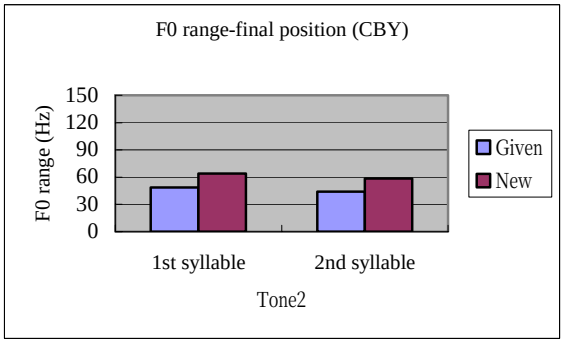


Figure 3.35 F0 ranges in first and second syllables of tone2 (high-rising tone) in final position for subject CBY.

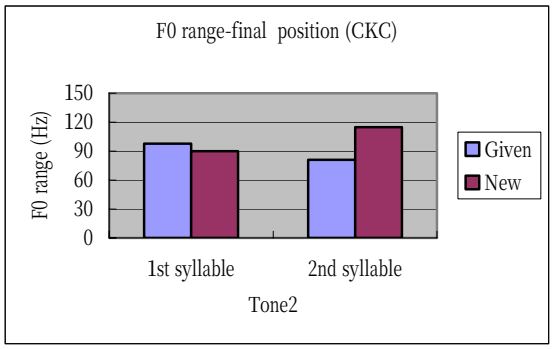


Figure 3.36 F0 ranges in first and second syllables of tone2 (high-rising tone) in final position for subject CKC.

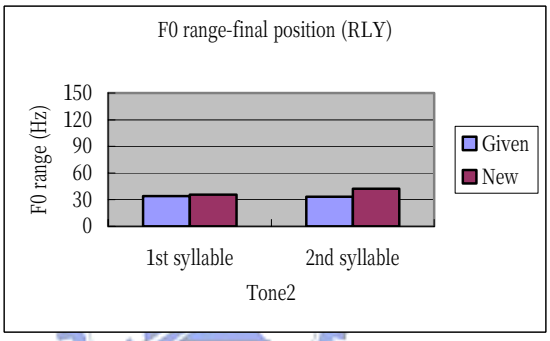


Figure 3.37 F0 ranges in first and second syllables of tone2 (high-rising tone) in final position for subject RLY.

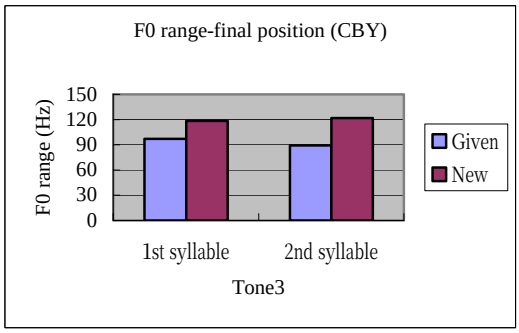


Figure 3.38 F0 ranges in first and second syllables of tone3 (low tone) in final position for subject CBY.

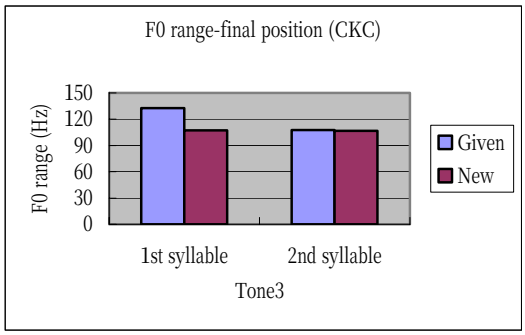


Figure 3.39 F0 ranges in first and second syllables of tone3 (low tone) in final position for subject CKC.

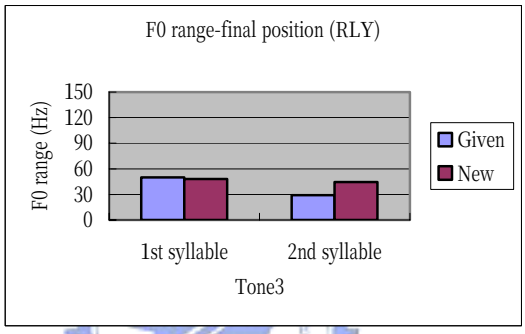


Figure 3.40 F0 ranges in first and second syllables of tone3 (low tone) in final position for subject RLY.

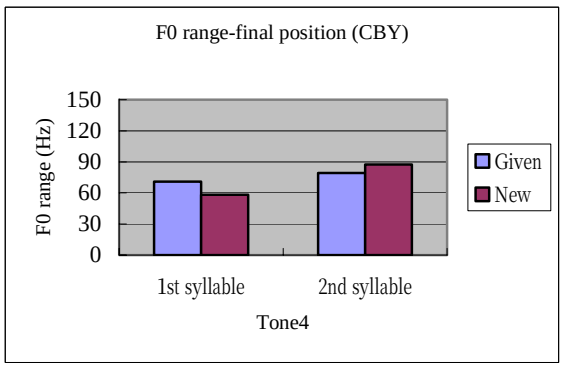


Figure 3.41 F0 ranges in first and second syllables of tone4 (falling tone) in final position for subject CBY.

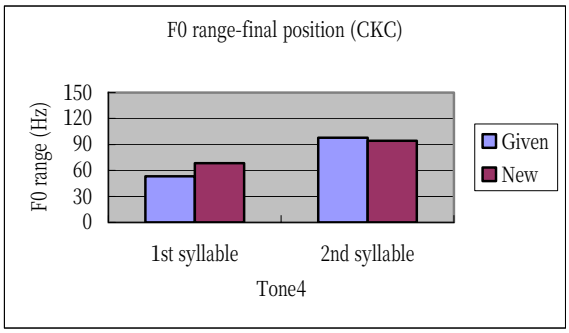


Figure 3.42 F0 ranges in first and second syllables of tone4 (falling tone) in final position for subject CKC.

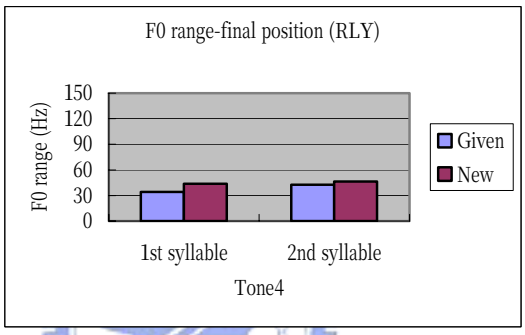


Figure 3.43 F0 ranges in first and second syllables of tone4 (falling tone) in final position for subject RLY.

3.4 Summary

In result section, firstly, F1 and F2 values and space chart of three target vowels have been considered with respect to focus condition. Though the vowel space of target vowels under narrow focus and neutral focus was not significantly different, we found that the vowel space seemed to be enlarged under narrow focus for subject CBY and subject CKC.

Secondly, the parameter of duration of target syllables has been investigated with respect to focus condition and position. Data showed that durations of target syllables under narrow focus was significantly longer than its counterparts under neutral focus in initial and middle position for subject CBY and subject RLY, no matter if they were in

the first or second syllable. For subject CKC, though the interaction effect was not significant, results showed that the duration of target syllables under narrow focus was still longer than that under neutral focus. Therefore, from this study's experiment on spontaneous data, we concluded that duration was highly related to focus conditions; it is one of the most salient parameters in Taiwanese Mandarin.

Finally, we investigated the parameter F0 range of target syllables with respect to focus condition and position. It was found that when target syllables were under narrow focus, F0 range was larger than its counterpart when under neutral focus for subject CBY and subject RLY. However, for subject CKC, F0 range of target syllables under narrow focus and neutral focus were not different. We found that positions of target syllables in the utterance also influenced F0 range. These syllables bore the largest F0 range when target syllables were in the final position. In addition, bar-chart figures show the trend that target syllables under narrow focus had higher mean of F0 range than their counterparts under neutral focus. Therefore, based on the results of the spontaneous data, we suggest that F0 range seems to be a salient acoustic cue in our study, however, not as prominent as the parameter, duration.

Chapter IV

Perception Experiment

4.1 Introduction

From the production results of Chapter III, we found that in spontaneous Taiwanese Mandarin, the vowel spaces of target syllables under narrow focus and neutral focus were not significantly different from each other, though there was a trend for enlargement of vowel space under narrow focus. The mean durations of the same syllables under narrow focus and neutral focus were significantly different from each other. In addition, except for subject CKC, F0 range of target syllables under narrow focus was significantly larger than its counterpart under neutral focus. Since these findings were based solely on production data, a perception experiment was conducted in order to investigate the acoustical cues used by native listeners to distinguish between neutral and narrow focus.

In this chapter, the method and results in perception experiment will be introduced. Finally, a comparison between the production and perception data from this study will be made.

4.2 Method

4.2.1 Subjects

Ten native Mandarin speakers, none of them participated in the production experiment, participated in the perception experiment. They were all born and raised in Mandarin-speaking families in Taipei City or Taipei County. None of them spoke either Taiwanese or Hakka (the two most widely-used dialects in Taiwan). Subjects included five males and five females aged between 20-25 years old. All of them were students in

National Chiao Tung University at the time of this experiment.

4.2.2 Corpus

The stimuli for the perception experiment consisted of 288 sentences produced by subject CBY from the previous production experiment. The reason for choosing subject CBY's data as the stimuli for the perception experiment was because CBY was the subject whose production of F0 range and duration showed significant differences between target syllables under narrow focus and neutral focus. By using subject CBY's production, we can compare the relative perceptual salience between duration and F0 range in spontaneous speech.

The two precursor questions produced by the experimenter to elicit each of the 288 sentences were also used in the perception experiment. For each sentence, there were two precursor questions: one was used to elicit answers with the target syllables under narrow focus, while the other was used to elicit answers with target syllables under neutral focus. Take the sentence [mau˧ mi˧ tə˧ yo˨ piɛn˧ çɪ˨ maˊ yiˋ] "Kitty is on the right side of ant" for example, the target noun was [mau˧ mi˧] "kitty." The sentence could respond to either [mau˧ mi˧ tə˧ yo˨ piɛn˧ çɪ˨ çəˊ mə˧] "What is on the right side of kitty?" or [çəˊ mə˧ tə˧ yo˨ piɛn˧ çɪ˨ maˊ yiˋ] "Ant is on the right side of what?" If the answer was in response to the first question, the target noun [mau˧ mi˧] "kitty" was under neutral focus. However, if the affirmative sentence was in response to the second question, the target noun [mau˧ mi˧] "kitty" was under narrow focus. The pair of focus pattern can be found in Table 2.3.

4.2.3 Stimulus creation and order of presentation

Both the precursor questions uttered by the experimenter and the answers produced by subject CBY were digitized at the sampling rate of 22.5kHz using Creative Wave Studio. A sound file containing the 576 precursor questions and 288 answers was created using Adobe Audition; each set of the two precursor questions and the answer were recorded in ABX order, with a 2 second interval between the first precursor question, the second precursor question, and the answer. A 5 second interval was recorded between each set of two questions and one answer. For each set of questions, there were two questions and one answer. The order in which these question sets were recorded and presented to subjects was randomized. The sound file was 84 minutes in length and was recorded onto a CD by dividing the file into 4 tracks, with 21 minutes for each track. There were 144 questions and 72 answers in each track.

4.2.4 Instrumentation

The CD with the sound files was displayed on a PC and listened to with a set of Grado Prestige Series Headphones, Model SR8. A Media Player program of Windows 2000 was used to play the CD containing the sound files of both precursor questions and answering sentences.

4.2.5 Experimental procedure

The perception experiment was conducted at the Acoustic Lab of the Department of Foreign Languages and Literatures, National Chiao Tung University. Subjects wore

headphones and sat in front of the computer.

Answer sheets were presented to subjects with two precursor questions in the same randomized order as the sound files on the CD to remind subjects of the two questions they heard. After listening to both the two precursor questions and the answer sentence, subjects indicated which one of the two precursor questions was the answer responding to and circled the corresponding choice on the answer sheets. There was a ten-minute break between each one of the four tracks of sound files.

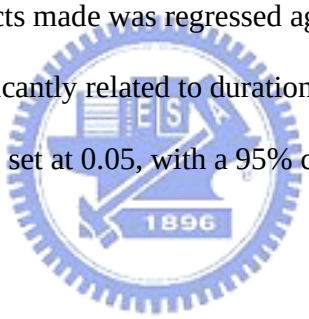
4.2.6 Data analysis

Each choice by every subject was coded. If subjects circled the precursor question that elicited the answering sentence with target syllables under narrow focus and carrying new information, the choice was coded as N; if subjects circled the precursor question used to elicit the answering sentence with target syllables under neutral focus and carrying given information, the answer was coded as G.

The participants' answers were further classified into four categories, G→G, G→N, N→G and N→N, by comparing subjects' choices with the correct answer. In these categories, the first letter represented the correct answer, and the second letter represented the given answers from the subject. For choices of questions eliciting given information on the target syllables that were matched with a answer containing target syllables carrying given information, the choice was counted as one instance of G→G category. For choices of questions containing target syllables carrying given information that were matched with questions eliciting new information on target syllables, the choice was counted as one instance of G→N category. For answers with target syllables carrying

new information that were matched with questions eliciting given information on the target syllable, the choice was counted as one instance of N→G category. For answers carrying new information that were matched with questions eliciting new information, the choice was counted as one instance of N→N category.

ANOVA tests were conducted with categories of answers, i.e. G→G, G→N, N→G, and N→N, as the independent factor, and duration or F0 range values as the dependent factor. Furthermore, results of coded answers were analyzed using Simple Regression. For each set of precursor questions and answers, the number of instances that 10 listeners chose as the precursor question used to elicit target syllables under narrow focus out of all the choices that 10 subjects made was regressed against the duration or F0 range value to exam if N-rating is significantly related to duration or F0 range value. In this study, the level of significance was set at 0.05, with a 95% confidence interval.



4.3 Results

In this section, the results of perception experiments will be reported. First, the effect of multiple-choice categorization on F0 range and duration will be discussed, and then the correlation between the percentage of N choices and the F0 range and duration of the target syllables for each answer sentence.

4.3.1 Effect of multiple-choice categorization on F0 range

According to the definition of coding mentioned in previous section, the answers were coded as G or N according to the choice subjects made. Table 3.1 summarizes the general performance of the 10 subjects with respect to their responses to the 288 pairs of

stimulus sentences.

Table 4.1 shows that the number of G→G category was 77 out of 288, which was 26.74%. The number of G→N category was 67 out of 288, which was 23.26%. The number of N→G category was 80 out of 288, which was 55.56%, and the number of N→N category was 64 out of 288, which was 22.22%.

Table 4.1 Four groups of perceptual answers.

Information status of target syllable	**Category of answers**			
	G→G N=77	**G→N N=67**	**N→G N=80**	**N→N N=64**
Given N=144	26.74%	23.26%		
New N=144			55.56%	22.22%

Results of one-way ANOVA with the four categories of answers as the independent variable were used to analyze mean F0 range of target syllables falling into each of the four categories. The table shows that F0 ranges were not significantly different ($F_{(3, 284)}=1.74=1.74$, $p=0.16$) in the four categories, G→G, G→N, N→G and N→N. Figure 4.1 below shows the distribution of F0 ranges of the four groups:

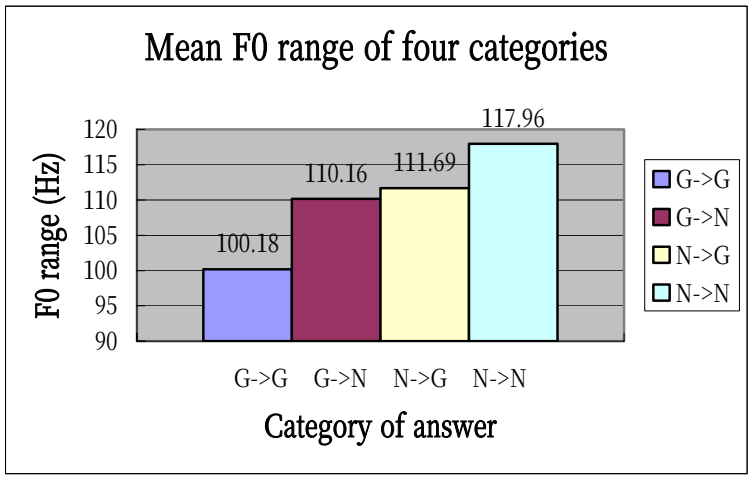


Figure 4.1 Mean F0 range values of the four categories of answers.

4.3.2 Effect of multiple-choice categorization on duration

We also examined the effect of the four categories of perceptual choices on duration. Results of one-way ANOVA with the four categories of answers as the independent variable were used to analyze duration of target syllables falling into each of the four categories. The results show that the durations of the four groups were significantly different from each other ($F_{(3,284)}$=91.78, $p<0.05$). Results of post-hoc Tukey test showed that the mean duration of target syllables of categories G→G and N→G are significantly shorter than categories G→N and N→N (Table 4.2). Figure 4.2 below shows the distribution of durations of the four groups:

Table 4.2 Post-hoc comparison of durations among four groups.

	G→G (290.36)	G→N (465.85)	N→G (317.50)	N→N (496.77)
G→G (290.36)		0.00*	$0.24^{n.s.}$	0.00*
G→N (465.85)	0.00*		0.00*	$0.21^{n.s.}$
N→G (317.50)	$0.24^{n.s.}$	0.00*		0.00*
N→N (496.77)	0.00*	$0.21^{n.s.}$	0.00*	

* $p < 0.05$

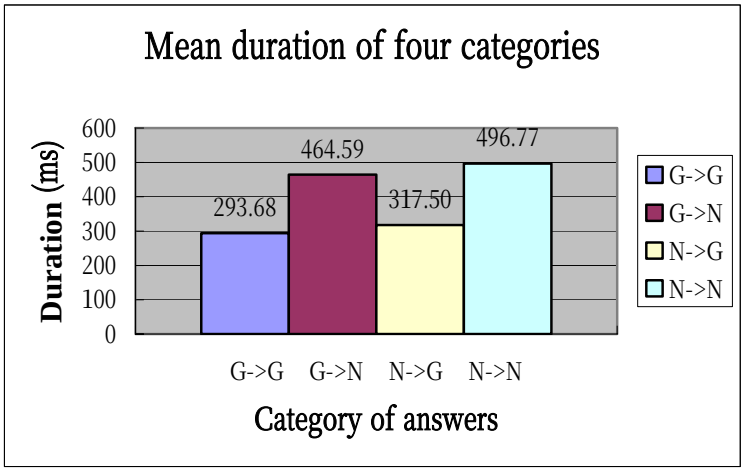


Figure 4.2 Mean duration of four categories of answers.

4.3.3 Correlation between N-rating, F0 range and duration

From all the answers, the N-rating was calculated. N-rating represented the percentage of choices on questions used to elicit target syllables under narrow focus out of the 10 choices that 10 subjects made for each set of precursor questions and answer. For example, in a certain perception question, 4 subjects circled the precursor question used to elicit target syllables under narrow focus, N-rating was 40%, no matter the answer was correct or not.

Table 4.3 shows the statistical results of percentage of choices on N regressed over mean F0 range value and mean duration separately. We found that mean F0 range value of target nouns was significantly related to the number of choices of answers eliciting target syllable under narrow focus ($F_{(1, 286)}$=8.69, $p<0.05$, R^2=0.03), however, R^2 value was only 3%. Mean duration of target syllables was also significantly related to the number of choices of answers eliciting target syllable under narrow focus ($F_{(1,286)}$=228.75, $p<0.05$, R^2=0.44). The longer the duration was, the more the subjects chose the precursor question eliciting target syllable under narrow focus as the answer. Therefore, compared with the statistical result of F0 range value, duration was more related to subjects' decision

Table 4.3 Correlation between N-rating and mean F0 range and duration.

N-rating	**F0 range (Hz)**	**P value**	**R^2**	**Duration (ms)**	**P value**	**R^2**
0.48	109.62	0.00	0.03	384.56	0.00	0.44

For subject CBY in the production experiment, both mean duration and F0 range value of the target syllables under narrow focus were significantly different than target syllables under neutral focus. However, subjects in the perception experiment used duration as the major perceptual cue to distinguish between the stimuli under narrow focus and neutral focus.

4.4 Summary

In this chapter, first, we found that duration was a more salient cue than F0 range in the perception experiment. The difference of F0 range value among the four categories of choices was not significant, but the difference of duration among the four groups was

significantly different. Post-hoc tests showed that the mean duration of target syllable of categories G→G and N→G are significantly shorter than categories G→N and N→N. Therefore, we inferred that subjects realized the two focus conditions when listening to the stimulus sentences because the mean duration of the answers being coded as N was significantly longer than the answers coded as G.

However, in Figure 4.2, the bar-chart figure shows that the difference of duration between G→G and N→G categories, and the difference of duration between G→N and N→N categories were not significant. It is proposed that the high percentage of the category N→G, i.e. 55.56% (see Table 4.1), was because among the perceptual materials which were recorded directly from subject CBY's production data without any manipulation, there were utterances of which the F0 range or duration were not significantly different under narrow focus and neutral focus.

Second, we found that the longer the duration of target syllables were, the more choices of precursor questions with target syllables under narrow focus were made. We concluded that subjects might choose duration as the perceptual cue when listening to the stimulus and choosing the answer. However, we could not explain why subjects chose duration instead F0 range as the cue since in subject CBY's production data, which was used as stimulus in perception experiment, the difference of duration and F0 range under narrow focus and neutral focus was significant.

Therefore, in this chapter, we have used Simple Regressions and ANOVAs to analyze two parameters, duration and F0 range, of target syllables under narrow focus and neutral focus in the perception experiment. The general conclusion is that in perceptual aspect in Taiwanese Mandarin, duration was used to perceive the distinction between neutral and

narrow focus.

Chapter Ⅴ

Conclusion

In this chapter, the results and findings in Chapter Ⅲ and Chapter Ⅳ will be summarized and further discussed in response to the three research questions. In addition, limitations of this study and suggestions on further studies will also be introduced.

5.1 Summary of the results of the two experiments

In the production study, we examined three acoustic parameters - vowel quality, duration and F0 range - with reference to bisyllabic nouns carrying new versus given information. In the perception study, we examined two parameters-duration and F0 range with reference to how native Taiwanese Mandarin speakers perceive bisyllabic nouns carrying new versus given information.

In the production experiment, the vowel space of /i/, /u/ and /a/ of a bisyllabic noun carrying new information was not significantly different from its counterpart carrying given information; however, a trend of vowel space enlargement was observed for syllables carrying new information for subject CBY and subject CKC. The mean duration of target syllables carrying new and given information were significantly different from each other for all subjects, and the interaction effect of focus and position upon duration was significant for subject CBY and subject RLY. For F0 ranges of target syllables carrying new and given information, significant differences were found for all subjects, but the interaction effect of focus and position upon F0 range was not significant for all subjects. It is found that for subject CKC, no significant interaction effect of focus condition and position upon both duration and F0 range of bisyllabic

target nouns was found. To summarize the results, two out of three subjects showed significant results in two-way repeated measure ANOVAs upon duration while none of the subjects showed significant results in two-way repeated measure ANOVAs upon F0 range. Therefore, we conclude that both duration and F0 range may be used as acoustic cues in Taiwanese Mandarin, but duration is more salient than F0 range in this study.

In the perceptual experiment, the effect of multiple-choice categorization on F0 range and duration was examined. We also investigated the correlation between N-rating, F0 range and duration.

We found that duration was more salient than F0 range in perception because the duration of four categories G→G, G→N, N→G and N→N were significantly different, while F0 ranges of the four categories were not significantly different.

Additionally, it was found that, the longer the durations of target syllables were, the more often precursor questions carrying new information would be chosen as the answer. Hence, duration was the most salient perceptual cue in the experiment, while F0 range was not as salient as duration in our study.

5.2 Response to research questions

As mentioned in the first chapter of this paper, there were two research questions in this study. The first one was to investigate the acoustical cues used in spontaneous speech to distinguish bisyllabic nouns carrying new versus given information, and compare these results of spontaneous speech with results of read speech by Jin (1996), Xu (1999), and Hsiung (2002). The second question was to identify the perceptual cues

used by listeners to distinguish new and given information in spontaneous Taiwanese Mandarin.

In spontaneous Taiwanese Mandarin, speakers expand F0 range and lengthen duration, as subject CBY and subject RLY in production to distinguish between new versus given information carried by the same bisyllabic nouns. However, the acoustical parameters used in read speech to distinguish between new versus given information carried by the same nouns were different.

In read Pekinese Mandarin, Jin (1996) and Xu (1999) found both that duration was lengthened and that F0 range was expanded in production, while Hsiung (2002) found that lengthening of duration was the only cue used in Taiwanese Mandarin to distinguish between new and given information. However, the results of spontaneous Taiwanese Mandarin showed that subject CBY and RLY lengthened duration and expanded F0 range for new information. In subject CKC's data, neither the duration was lengthened nor F0 range was expanded when the same syllables carried new information in initial, middle and final position.

A similar pattern was found in vowel space. In Japanese (Maekawa, 1996) and Arabic (de Jong, 2002), the vowel space was enlarged for vowels carrying new information; however, in spontaneous Taiwanese Mandarin, the enlargement of vowel space was not significantly different between vowels carrying new versus given information, in spite of the trend of vowel space enlargement being observed.

When comparing our results with previous studies, it was found that the scale of difference for spontaneous speech was much smaller than with read speech. It might be due to fast speaking rates for spontaneous utterances, since the experiment was conducted

in a speed of normal daily conversation. During fast speaking rates, the acoustical cue may shorten the duration of each component of the utterance, or increase the overlapping of each component, which results in co-articulation and the reduction of total duration. The magnitude of articulation reduced as well (Byrd and Tan, 1996). However, a study that controls the speaking rate is necessary to verify this claim.

To answer the second research question, it was found that duration was the most salient perceptual cue in Taiwanese Mandarin. When presented with production data from subject CBY, which was both lengthened in duration and expanded in F0 range for target syllables carrying new information, listeners were more apt to tune in to the lengthening of duration and to ignore the F0 range expansion while perceiving nouns under narrow focus; that carried new information. As there is a difference in acoustic cues used in production of read versus spontaneous speech, in perception aspect, F0 range was found to be the most salient perceptual parameter used to distinguish between nouns under broad focus and narrow focus in Jin's study (1996), while duration was not as important as F0.

5.3 Conclusion and Suggestions for Further Studies

We conclude that both duration and F0 range seem to be the salient acoustic parameters in spontaneous Taiwanese Mandarin in production experiment; however, duration is found to be more salient than F0 range in our production data. As for perceptual cues, duration is the most important factor in this study.

For future studies, spontaneous data produced by more speakers could be recorded, to provide stronger statistical support of these results. For perceptual experiments, other

sets of perceptual materials can be recorded, one set is that the data is significantly different under narrow focus and neutral focus in only duration, the other set is that the data is significantly different under narrow focus and neutral focus in only F0 range, to verify the perceptual results observed here. Also, speech synthesis can be conducted with subject CBY's production data. F0 range expansion of target nouns can be adjusted to examine to what extent can subjects realize the difference, and so can the parameter, duration.

Reference

Birch, S. & J. Charles Clifton (1995). Focus, accent, and argument structure: effects on language comprehension. *Language and Speech, (38)*, 365-391.

Bolinger, D. (1972). Accent is predictable (if you're a mind reader). *Language, (48)*, 633-644.

Burnell, H. T. (1997). Interaction among f0, duration, and amplitude in the perception of focus. *Journal of Acoustical Society of America, (102)*, 3203-3204.

Byrd and Tan (1996). Saying consonant clusters quickly. *Journal of Phonetics, (24)*, 263-282.

Chao, Yuen Ren (1961). *Mandarin primer: an intensive course in spoken Chinese.* Harvard University Press.

Chao, Yuen Ren (1968). *A grammar of spoken Chinese.* University of California Press.

Cooper, W. E., S. J. Eady, et al. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *Acoustical Society of America,(77)*, 2142-2156.

Cruttenden, A. (1997). *Intonation Second Edition*, Cambridge University Press.

De Jong, K. & B. Zawaydeh (2002). Comparing stress, lexical focus, and segmental focus: patterns of variation in Arabic vowel duration. *Journal of Phonetics (30)*, 53-75.

Eady, S. J. and W. E. Cooper (1986). Speech intonation and focus location in matched statements and questions. *Acoustical Society of America,(80)*, 402-415.

Eady, S. J., W. E. Cooper, et al. (1986). Acoustic characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments. *Language and Speech,(29)*, 233-251.

Fan, K. T. & Chang, Y. J. (2000). *Xian dai han u u fa fen xi [Syntactic analysis of modern Mandarin]*. Hua Dueng University Press.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of Acoustical Society of America,(27),* 765-768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech,(1),* 126-152.

Garding, E. (1987). Speech act and tonal pattern in standard Chinese: constancy and variation. *Phonetica,(44),* 13-29.

Gussenhoven, C. & A. C. M. Rietveld (1992). Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics (20),* 283-303.

He, Y & Jieng, S. (1992). Bei jieng u u diao de sh ian tan suo [Experimental research on intonation in Peking Mandarin]. *U yan jiao xue u ian jiou,(2),* 71-96.

Heldner, M. & E. Strangert (2002). Temporal effects of focus in Swedish. *Journal of Phonetics,(29),* 329-361.

Ho, A. T. (1976). The acoustic variation of Mandarin tones. *Phonetica,(33),* 353-367.

Ho, A. T. (1977). Intonation variation in Mandarin sentence for three expressions: interrogative, exclamatory, and declarative. *Phonetica,(34),* 446-457.

Hsiung, S. C. (2002). *Acoustic characteristics of sentential focus in Mandarin spoken in Taiwan*. MA thesis, National Hsinchu Teachers College.

Jin, S. (1996). *An Acoustic Study of Sentence Stress in Mandarin Chinese.* Doctoral dissertation, Columbus, Ohio, The Ohio State University.

Kreiman, J. (1982). Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics,(10),* 163-175.

Li, Z. Z. (1996). Tueng ueng pain zhang zhueng in ba uo xien xi jian dian [Stress attracts focus of information]. *Uai u ian jiou,(3),* 15-20.

Lo, W. P. (2003). *The psychological reality of Third Tone Sandi in Mandarin Chinese.* MA thesis, Graduate Institute of Linguistics and Cultural Studies, Hsinchu, National Chiao Tung University.

Maekawa, K. (1997). Effects of focus on duration and vowel formant frequency in Japanese. *Computing prosody* (pp. 129-153). Y. Sagisaka, N. Campberll and M. Higuchi. New York, Springer.

Peng, S. H., Chan, M.K.M., Tseng, C., Huang, T., Lee, O. J., and Beckman, M. E. (in press). Towards a Pan-Mandarin prosodic annotation system. To appear in S. A. Jun, ed., *Prosodic Models and Transcription: Toward Prosodic Typology.* Oxford University Press.

Shen, X. S. (1990). Tonal coarticulation in Mandarin. *Journal of Phonetics,(18),* 281-295.

Shen, X. S. (1993). Relative duration as a perceptual cue to stress in Mandarin. *Language and Speech,(36),* 415-433.

Shen, X. S. (1993). The use of Prosody in Disambiguation in Mandarin. *Phonetica,(50),* 261-271.

Shen, J. (1985). Bei jieng hua sheng diao de in u ji u diao [Pitch range and intonation of lexical tones in Peking Mandarin]. *Bei jieng u ien sh ian lu.* Peking University Press.

Shen, J. (1955). Han u u diao gou zao u u diao lei xieng [Structure and type of intonation in Mandarin]. *Fan yan (Dialect),(3),* 221-228.

Shen, J. (1955). Hau u in gao xi tueng de iou sheng xieng han qu bie xieng [Voiced characteristics of pitch system in Mandarin]. *U yan uen zu ieng ueng,(2),* 13-18.

Sun, R. J. (2000). Han u u diao de u qi han kou qi gueng neng [The function of Mandarin intonation]. *Nan tueng sh fan zue uan zue bao,(3),* 69-73.

Suomi, K., J. Toivanen, et al. (2002). Durational and tonal correlates of accent in Finnish. *Journal of Phonetics,(31),* 113-138.

Swerts, M. and E. Krahmer (2002). Prosodic marking of information status in Dutch and Italian: a comparative analysis. *Journal of Phonetics,(30),* 629-654.

Yeh, J. (2001). *Han u u ju uen lu de u fa gueng neng [The pragmatic function of Mandarin sentences].* Hua Dueng Univeristy Press.

Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of Acoustical Society of America, (95),* 2240-2253.

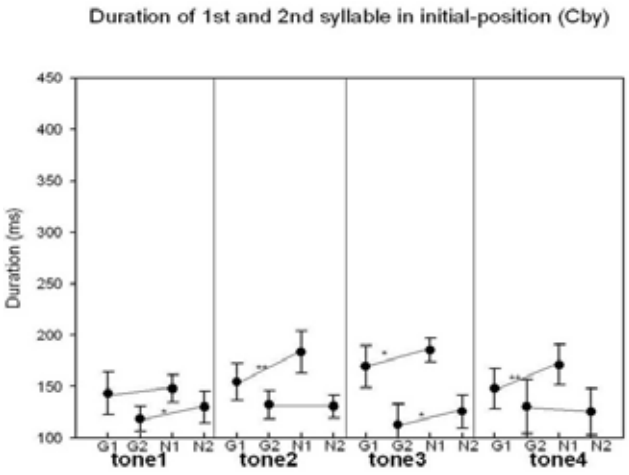
Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics,(25),* 61-83.

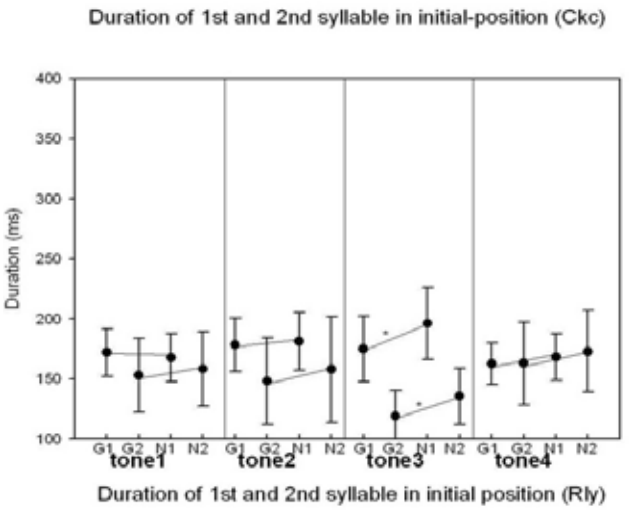
Xu, Y. (1999). Effects of tone on the formation and alignment of f0 contours. *Journal of Phonetics,(27),* 55-106.

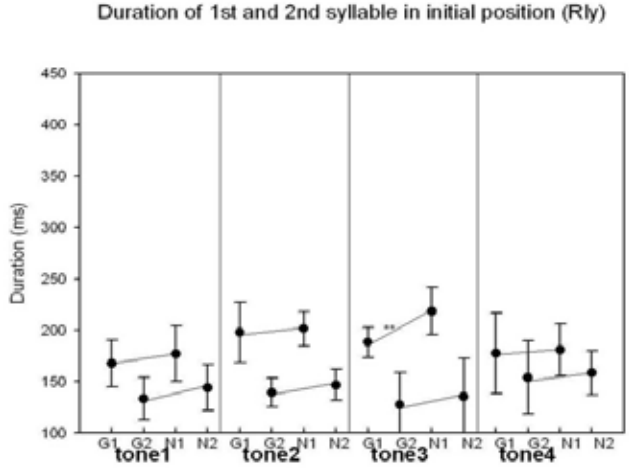
Zubizarreta, M. L. (1998). *Prosody, Focus, and Word Order*, MIT Press.

Appendix A

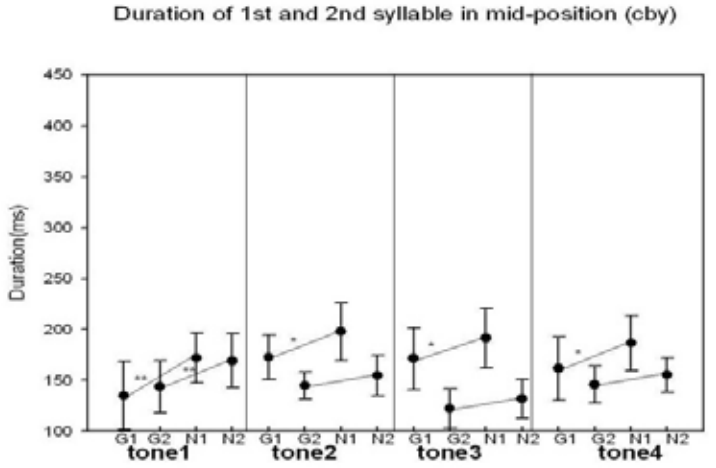
Duration in first and second syllables of four tones under narrow and neutral focus in initial position for subject CBY, CKC and RLY.

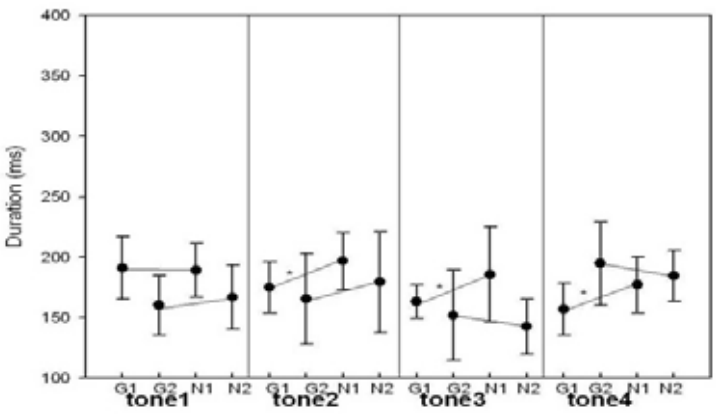


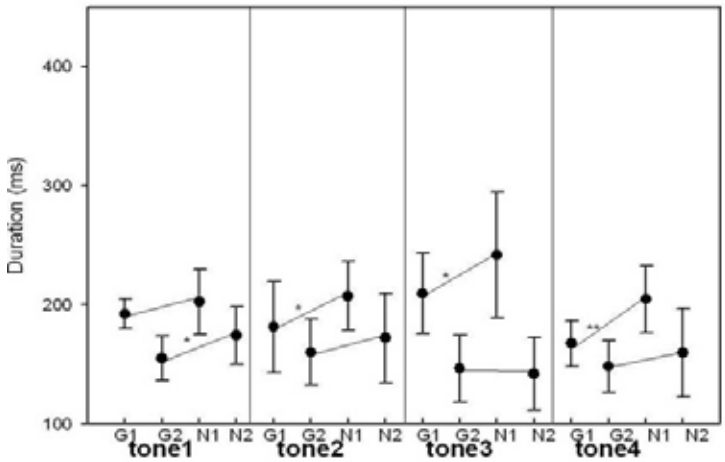




Duration in first and second syllables of four tones under narrow and neutral focus in middle position for subject CBY, CKC and RLY.







Duration in first and second syllable of four tones under narrow and neutral focus in final positions for subject CBY, CKC and RLY.