An Acoustic Study of Sentence Stress
in Mandarin Chinese

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Shunde Jin, B.A., M.A.

* * * * *

The Ohio State University
1996

Dissertation Committee:
Marjorie K.M. Chan
Mary E. Beckman
James H.Y. Tai

Approved by

[Signature]
Adviser
Department of East Asian Languages & Literatures
Copyright by
Shunde Jin
1996
To My Wife and Children
ACKNOWLEDGMENTS

I would like to extend my sincere thanks and appreciation to Marjorie K.M. Chan, my dissertation advisor, for her teaching, guidance and advice throughout this research project. I have benefited tremendously from her teaching and knowledge of the Chinese language. Without her persistent support and encouragement, this dissertation would not have come into existence.

I also consider myself very fortunate in having Mary Beckman as one of the members on my dissertation committee. She first introduced me into the subject of experimental phonetics. And it is from her that I learned to apply statistics to language analysis. Her profound knowledge in these areas has greatly influenced my analysis in this study. I would also like to thank her for letting me use the phonetics laboratory facilities of the Linguistics Department.

Special thanks also go to James Tai for first getting me interested in exploring the semantic properties in the Mandarin prosody. I have learned a great deal from his teaching and research.
I am very grateful to other faculty members in the Department of East Asian Languages and Literatures for broadening my academic background; to my friends and fellow students for making my life a lot more fun and interesting for the past six years.

I am also grateful to my wife, Yanfen, for her never-ending moral support, encouragement and love; to my children, Danshen and Daniel, who have made my life more meaningful and three dimensional.

Lastly, I owe a deep debt of gratitude to my parents. Without their encouragement and support, it would be hardly imaginable that I could ever attain this lifetime goal.
VITA

1981 B.A. in English Linguistics. Department of Foreign Languages, East China Normal University, Shanghai, China


1985–1990 Lecturer, Research Center for Foreign Languages and Literatures, Shanghai International Studies University, Shanghai, China

1990–1996 Graduate Teaching Associate, Department of East Asian Languages and Literatures, The Ohio State University, Columbus, OH.

PUBLICATIONS

1. Monograph/Book


2. Articles


In press. Toward a systematic account of Shanghai tonal phonology. In Jialing Wang and Norval Smith (eds.), *Studies in Chinese Phonology*. 

vi
3. Translations


FIELDS OF STUDY

Major field: East Asian languages and literatures

Studies in Chinese tones and intonation

Studies in Chinese phonology and dialectology

Studies in Mandarin syntax and grammar
## TABLE OF CONTENTS

DEDICATION ................................................................. ii

ACKNOWLEDGMENTS .................................................... iii

VITA ........................................................................ v

Chapter I  Introduction .................................................. 1

1.1 Aim of this study ..................................................... 1
1.2 Data ................................................................... 3
1.3 Target syllables .................................................... 8
1.4 Statistical analysis ................................................. 10
1.5 Outline ................................................................. 15

CHAPTER II  Mandarin stress ......................................... 18

2.1 Tones .................................................................. 18
2.2 Intonation ............................................................ 25
2.3 Stress ................................................................. 32

Chapter III  Focus and stress ......................................... 49

3.1 Background ......................................................... 49
3.2 The results of the perception experiment ............... 68

Chapter IV  The F0 parameter ....................................... 76

4.1 Gross pattern of F0 with respect to stress ............... 76
  4.1.1 F0 and stress ................................................. 77
  4.1.2 Tone and stress ............................................. 84
  4.1.3 Unstressed syllables ................................. 89
4.2 Measurement points ............................................. 95
4.3 The name syllable ............................................... 103
4.4 The syllable mìng ............................................... 112
4.5 The syllable liáo ................................................ 127
4.6 A comparison of stress 3 and stress 4 ................. 132
4.7 Conclusion ......................................................... 136

viii
Chapter V  Duration and intensity .......................... 140

5.1 Duration .................................................. 140
   5.1.1 The name syllable .................................. 141
   5.1.2 The syllable míng ................................. 147
   5.1.3 The syllable liáo ................................. 150
   5.1.4 A comparison of stress 3 and stress 4 ....... 152
   5.1.5 Conclusion ........................................ 157

5.2 Intensity .................................................. 158
   5.2.1 The name syllable .................................. 158
   5.2.2 The syllable míng ................................. 164
   5.2.3 The syllable liáo ................................. 167
   5.2.4 A comparison of stress 3 and stress 4 ....... 171
   5.2.5 Looking across the sentence .................... 175
   5.2.6 Conclusion ........................................ 178

Chapter VI  Perception and stress .......................... 181

6.1 Multiple regression ...................................... 182
6.2 Stepwise regression with four variables .......... 185
6.3 Stepwise regression with twelve variables ....... 197
6.4 Conclusion ................................................ 203

Chapter VII  Conclusion ...................................... 205

Appendices ..................................................... 211

A  Mean value of pitch ranges of the name syllables
   with respect to subject, tone, and stress
   (N = 5) ..................................................... 212

B  Mean values of pitch ranges of the syllable míng
   with respect to subject, stress, and tonal
   context (N = 5) .......................................... 213

C  Mean values (Hz) at point (c) under stress 1 and
   the one-tailed t-test results ........................... 214

D  Mean values (Hz) at point (d) under stress 1 and
   the one-tailed t-test results ........................... 215

E  Mean values of pitch ranges of the syllable liáo
   with respect to subject, stress, and tonal
   context (N = 5) .......................................... 216

F  Mean values of duration for the name syllables
   with respect to subject, tone, and stress
   (N = 5) ..................................................... 217
Mean values of intensity for the name syllables with respect to subject, tone and stress (N = 5) ................................. 218

Mean values of intensity for the syllable míng with respect to subject, tonal context and stress (N = 5) ........................................ 219

Bibliography .............................................................. 220
<table>
<thead>
<tr>
<th>TABLES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Focus patterns</td>
</tr>
<tr>
<td>2.1</td>
<td>Mandarin tones and notations</td>
</tr>
<tr>
<td>3.1</td>
<td>Subjects' responses in the perception experiment</td>
</tr>
<tr>
<td>3.2</td>
<td>95% confidence interval of the proportion of all choices and the corresponding z-values</td>
</tr>
<tr>
<td>4.1</td>
<td>Analysis of variance for the pitch range of the name syllables</td>
</tr>
<tr>
<td>4.2</td>
<td>Mean value (Hz) of pitch ranges for the name syllable with respect to stress (N = 80)</td>
</tr>
<tr>
<td>4.3</td>
<td>Pairwise Tukey comparisons of pitch range values of the name syllables</td>
</tr>
<tr>
<td>4.4</td>
<td>Analysis of variance for the pitch range of the syllable ming</td>
</tr>
<tr>
<td>4.5</td>
<td>Mean value (Hz) of pitch ranges for the syllable ming with respect to stress (N = 80)</td>
</tr>
<tr>
<td>4.6</td>
<td>Analysis of variance for the pitch range values of the syllable liáo</td>
</tr>
<tr>
<td>4.7</td>
<td>Mean values of pitch ranges for the syllable liáo under the four stress conditions</td>
</tr>
<tr>
<td>4.8</td>
<td>Pairwise Tukey comparisons of pitch range values of the syllable liáo</td>
</tr>
</tbody>
</table>
4.9 Comparisons of the mean pitch range values of
the name syllable (N = 20) under stresses 3 and
4 ........................................................................ 132

4.10 Comparisons of the mean pitch range values for
the syllable míng (N = 20) under stresses 3 and
4 ........................................................................ 134

5.1 Analysis of variance for the duration of the
name syllables ....................................................... 141

5.2 Mean values of duration (ms) for the name
syllable with respect to stress (N = 80) ............ 142

5.3 Pairwise Tukey comparisons of the duration of
the name syllables ......................................................... 146

5.4 Analysis of variance for the duration of the
syllable míng .............................................................. 148

5.5 Mean values of duration of syllable míng with
respect to stress (N = 80) ...................................... 148

5.6 Analysis of variance for the duration of the
syllable liáo ............................................................... 151

5.7 Mean values of duration of the syllable liáo
with respect to stress (N = 80) ........................ 152

5.8 Comparisons of the mean duration (ms) for the
name syllable (N = 20) under stresses 3 and 4 ..... 153

5.9 Comparisons of the mean duration (ms) for the
syllable míng (N = 20) under stresses 3 and 4 ..... 155

5.10 Analysis of variance for the intensity of the
name syllables ......................................................... 159

5.11 Mean values of intensity (dB) of the name
syllable with respect to stress (N = 80) ............. 159

5.12 Pairwise Tukey comparisons of the intensity
values of the name syllable ................................. 162

5.13 Analysis of variance for the intensity of the
syllable míng .......................................................... 164

5.14 Analysis of variance for the intensity of the
syllable liáo ............................................................ 167
5.15 Mean values of intensity of the syllable liáo with respect to stress (N = 80) ..................... 168
5.16 Pairwise Tukey comparisons of the intensity values of the syllable liáo ......................... 170
5.17 Comparisons of the mean intensity (db) for the name syllable (N = 20) under stresses 3 and 4 ..... 172
5.18 Comparisons of the mean intensity (db) for the syllable míng (N = 20) under stresses 3 and 4 ..... 174
6.1 Stepwise regression of the four variables of each target syllable against the score "who" ...... 185
6.2 Stepwise regression of the four variables of each target syllable against the score "when" ...... 187
6.3 Stepwise regression of the four variables of each target syllable against the score "do what" ..................................................... 188
6.4 Stepwise regression of the three parameters of each target syllable against the score "what happened" ..................................................... 194
6.5 Stepwise regression of twelve parameters of the three target syllables against the four scores .... 198
C.1 Mean values (N = 5) ............................................. 214
C.2 T-test results ................................................. 214
D.1 Mean values (N = 5) ............................................. 215
D.2 T-test results ................................................. 215
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Grid model</td>
<td>27</td>
</tr>
<tr>
<td>4.1 Stress 1 (T1) for subject 1 (C174)</td>
<td>77</td>
</tr>
<tr>
<td>4.2 Stress 2 (T1) for subject 1 (C150)</td>
<td>78</td>
</tr>
<tr>
<td>4.3 Stress 3 (T1) for subject 1 (C106)</td>
<td>78</td>
</tr>
<tr>
<td>4.4 Stress 4 (T1) for subject 1 (C126)</td>
<td>79</td>
</tr>
<tr>
<td>4.5 Stress 1 (T1) for subject 3 (C374)</td>
<td>79</td>
</tr>
<tr>
<td>4.6 Stress 2 (T1) for subject 3 (C313)</td>
<td>80</td>
</tr>
<tr>
<td>4.7 Stress 3 (T1) for subject 3 (C368)</td>
<td>80</td>
</tr>
<tr>
<td>4.8 Stress 4 (T1) for subject 3 (C330)</td>
<td>81</td>
</tr>
<tr>
<td>4.9 Stress 4 (T1) for subject 1 (C130)</td>
<td>83</td>
</tr>
<tr>
<td>4.10 Stress 1 (T1) for subject 1 (C105)</td>
<td>84</td>
</tr>
<tr>
<td>4.11 Stress 1 (T2) for subject 1 (C108)</td>
<td>85</td>
</tr>
<tr>
<td>4.12 Stress 1 (T1) for subject 2 (C251)</td>
<td>87</td>
</tr>
<tr>
<td>4.13 Stress 1 (T3) for subject 1 (C115)</td>
<td>88</td>
</tr>
<tr>
<td>4.14 Stress 1 (T4) for subject 1 (C136)</td>
<td>88</td>
</tr>
<tr>
<td>4.15 Stress 2 (T4) for subject 3 (C303)</td>
<td>90</td>
</tr>
<tr>
<td>4.16 Stress 3 (T4) for subject 3 (C348)</td>
<td>91</td>
</tr>
<tr>
<td>4.17 Overlaid pitch contours of Figures 4.8 and 4.16</td>
<td>92</td>
</tr>
</tbody>
</table>
4.18 Stress 1 (T1) for subject 3 (C324) .................... 94
4.19 Stress 4 (T1) for subject 3 (C330) .................... 98
4.20 Measurement points (a) and (b) for T1 (C105) ....... 99
4.21 Measurement points (a) and (b) for T2 (C138) ....... 100
4.22 Measurement points (a) and (b) for T3 (C165) ....... 100
4.23 Measurement points (a) and (b) for T4 (C140) ....... 101
4.24 Mean values of pitch ranges of the name syllables with respect to subject, tone, and stress .......................... 107
4.25 Overlaid F0 contour of stress 1 versus stress 4 for subject 4 (C408 and C437) ......................... 109
4.26 Overlaid F0 contour of stress 1 versus stress 4 for subject 3 (C308 and C337) ......................... 110
4.27 Mean values of pitch ranges of the syllable míng with respect to subject, stress, and tonal context .......................... 116
4.28 F0 contour of stress 1 (T3) from subject 2 (C253) ........................................ 117
4.29 Mean F0 values (N = 5) at point (c) under stress 1 ........................................ 118
4.30 Mean F0 values (N = 5) at point (d) under stress 1 ........................................ 119
4.31 Mean values of pitch ranges of the syllable liáo (n = 5) with respect to subject, tonal context and stress ........................................ 130
4.32 Scatter plot of the mean pitch range values of the name syllable (N = 5) with respect to stresses 4 and 3 ........................................ 133
4.33 Scatter plot of the mean pitch range values for the syllable míng (N = 5) with respect to stresses 4 and 3 ........................................ 135
5.1 Mean values of duration (ms) of the name syllable with respect to subject, tone, and stress ........................................ 144

5.2 Mean values of duration of the syllable míng for the four subjects with respect to stress ........... 150

5.3 Scatter plot of duration of the name syllable (N = 5) with respect to stresses 4 and 3 ............ 154

5.4 Scatter plot of duration of the syllable míng (N = 5) with respect to stresses 4 and 3 .......... 156

5.5 Mean values of intensity of the name syllable with respect to subject, tone, and stress
(N = 5) ................................................................. 161

5.6 Mean values of intensity of syllable míng with respect to subject and stress .................. 165

5.7 Mean values of intensity of the syllable míng with respect to subject, tonal context, and stress ......................................................... 166

5.8 Mean values of intensity (db) of the syllable liáo for the four subjects with respect to stress .................................................. 169

5.9 Scatter plot of intensity (db) for the name syllable (N = 5) with respect to stresses 4 and 3 .............................................. 173

5.10 Scatter plot of intensity (db) for the syllable míng (N = 5) with respect to stresses 4 and 3 ...... 174

5.11 Mean values of intensity for the three target syllables with respect to stress ................... 176
Chapter I

Introduction

1.1 Aim of this study

Mandarin Chinese is an official language of both mainland China and Taiwan. According to the 1995 edition of the World Almanac, the language is spoken by a total of 952 million people throughout China and overseas, among whom 836 million are native speakers.¹

The major aim of this study is to investigate various acoustic properties of sentence stress in this language. Although there has been increased recent interest in various prosodic aspects of this language (Connell, Bruce et al. 1983, Coster and Kratochvil 1984, Gårding 1987, Ho 1977, Liao 1994, Lin and Yan 1980, J. Shen 1985, X. Shen 1990a, Shih 1988, and Tseng 1988), most of the previous studies mainly focused on the interaction of tone and intonation. Relatively little attention has been paid to stress, especially sentence stress.

Linguistically, the functions of sentence stress include, among other things, delimiting syntactic boundaries
and providing focus and contrast. This study only deals with the latter function. It is based on the results of two experiments which made use of this function (i.e. focus and contrast).

The stress types we will be investigating can be generally classified into two: contrastive stress and broad focus sentence stress. Contextually, contrastive stress usually occurs in sequences of sentences with parallel constituents that are filled with different morphemes. Functionally, contrastive stress is used to distinguish a particular morpheme from other morphemes that may occur in the same position (Lehiste 1970). Its semantic focus is only on the stressed item. The unfocused part serves as the background which is assumed to be shared by the hearer. Broad focus sentence stress, on the other hand, does not have these restrictions. There is no assumption about the background knowledge on the part of the hearer. In terms of information, its focus is on the whole sentence.

In English, the issue of broad focus and stress has been well studied (Bolinger 1972, Gussenhoven 1983a, Halliday 1967, Jackendoff 1972, Ladd 1980, Schmerling 1976, and Selkirk 1984). The same phenomenon in Mandarin Chinese, however, is poorly understood. Very little study has been done in this area. In this study, we will examine the sentence stress pattern in this language that represents
broad focus and see how it differs from contrastive stress that realizes narrow focus.

The major acoustic parameters for stress, tone and intonation have been generally recognized to be the fundamental frequency (F0), duration and intensity (Lehiste 1970). The main goal of this study is to examine these three acoustic parameters across the sentence with respect to different stress conditions and find out how stress is manifested by these parameters. Four stress conditions across three syntactic positions will be examined. These are described in the following section.

1.2 Data

Data in this study come from two experiments, one testing production and the other testing perception. Four subjects, two males and two females, all in their early thirties participated in the production experiment. They were all born in Beijing and grew up there and had been in this country for less than three years. Analyses in the following chapters code the two male subjects (JM and FR) as subjects 1 and 2 and the two female subjects (QO and LX) as subjects 3 and 4.

Sixteen other subjects, not including the four in the production experiment, participated in the perception experiment. The subjects' ages vary from 24 to 34. Although
these subjects do not originally come from the same place, they all speak the same Northern dialects. Linguistically, they are a homogeneous group. In terms of sex, this is also a balanced group with 8 males and 8 females.

The production test was done in the sound booth in the phonetics laboratory at the Linguistics Department of The Ohio State University. In order to ensure a consistent measurement of intensity, all four subjects were recorded using a head-mounted microphone to keep the microphone and their mouths at a fixed distance and angle so that the input level can be made relatively stable.

The perception test was done in the language laboratory of The Ohio State University. Sitting at separate carousals, subjects listened to the test material played from the control panel through the headphone. Each of them had equal amount of time to make their decisions about a certain choice. Before they listened to the test materials, some examples were given.

The production material consists of four simple six-syllable sentences. The basic pattern is NAME míngnián liáoyàng "NAME + next year + go to the sanitarium" (someone is going to the sanitarium next year). The four sentences vary only in the NAME portion, or more precisely, the surname part of the familiar expression "lǎo + surname" (old + surname). The surnames have identical CV structure and even
identical vowels. They vary principally in tone. The four test sentences are presented in (1.1) below. The tonal diacritics in these sentences will be explained in Chapter Two.

(1.1) Lǎo wū míngnián liáoyǎng.
    Lǎo wū míngnián liáoyǎng.
    Lǎo lǜ míngnián liáoyǎng.
    Lǎo lǜ míngnián liáoyǎng.

Note that each of the surnames and each of the other syllables in the sentence begins with a sonorant sound. This choice of the segmental type was made in an attempt to obtain continuous F0 contours. It can also be observed that the four surnames do not have the same initial consonant. Of course, it would be ideal to find perfect minimal pairs. But in Mandarin lexicon, this is the closest I can get that satisfies both tonal and segmental requirements. In any case, the difference in the sonorant onset will not significantly affect the results of this study.

Each test sentence in (1.1) above, in turn, yields four utterance types which vary in the focus pattern. These focus patterns were elicited by asking the subjects to read a particular sentence using the intonation that he/she felt would answer the question posed by me. The following table summarizes these focus patterns, the corresponding focused
items (underlined) and the eliciting questions. Broad focus is represented by pattern 4.

Table 1.1 Focus patterns

<table>
<thead>
<tr>
<th>Focus</th>
<th>Place of focus</th>
<th>Eliciting questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lǎo X míngnián liáoyāng.</td>
<td>Shéi míngnián liáoyāng? 'Who is going to the sanitarium next year?'</td>
</tr>
<tr>
<td>2</td>
<td>Lǎo X míngnián liáoyāng.</td>
<td>X shénme shíhou liáoyāng? 'When is X going to the sanitarium?'</td>
</tr>
<tr>
<td>3</td>
<td>Lǎo X míngnián liáoyāng.</td>
<td>X míngnián gànshénme? 'What is X doing next year?'</td>
</tr>
<tr>
<td>4</td>
<td>Lǎo X míngnián liáoyāng.</td>
<td>Fāshèngle shénmesì? 'What happened?'</td>
</tr>
</tbody>
</table>

Altogether, we have sixteen different test sentences with four different tones for each surname and four focus patterns for each sentence with a specific tone. In addition, five tokens were used for each individual test sentence. The total of eighty sentences (4 x 4 x 5 = 80) were then randomized and presented to the subjects for recording.

The subjects were told to read the sentence as if they were talking to someone three feet away and at a normal conversational speed. As has been explained above, they
first listened to my question before they started to read a sentence.

The resulting three hundred and twenty sentences (80 x 4 speakers = 320) were then digitized at the sampling rate of 28000 using the CECIL (Computerized Extraction of Components of Intonation in Language) program. The digitized sentences were recorded to another tape in randomized order and serve as the test material for the perception experiment.

For each of the three hundred and twenty test items in the perception experiment, subjects were presented with four choices: (a) shéi 'who', (b) shénme shíhòu 'when', (c) gànshénmé 'do what', and (d) shénmèsì 'what happened', which correspond to the four eliciting questions above in Table 1.1. They were told that after listening to each sentence, they must first judge which of the four questions it was in response to which the sentence was originally read, and circle the corresponding choice on the answer sheet.6

Each choice by one subject counts as one point toward the score for that response for the stimulus. I use score (a) to represent the total points for choice (a) of a particular sentence, score (b) for choice (b), and so on. So for each of the stimuli, there are four scores, (a) through (d). As a result, the value for each score ranges from 0 to 16. The results of the perception experiment are summarized
in Section 3.2. They will be used again in Chapter Six for multiple regression analysis.

1.3 Target syllables

For acoustic analysis, three target syllables are chosen. They are the surname part of the familiar name expression (the name syllable), míng in míngnián and liáo in liáoyâng (see [1.1] above). The name syllable is the only target syllable controlled for tone. For this reason, I have chosen a rising tone for the following syllable míng to avoid possible tone-sandhi effects (see Chapter Two for discussions of tone and tone sandhi). The syllables míng and liáo are chosen because they form part of the word which will be the target for focus.

The reasons why I chose the first rather than the second syllable in míngnián and liáoyâng as the target syllables are as follows: First, the syllable míng shows more clearly whether the inherent tone is overridden by intonation or not. There is no problem if both míng and nián are falling when they are in post-stress position. However, I find that, for some tokens, when míngnián is stressed, the syllable nián is actually falling, with the peak of stress falling on the syllable míng. For these cases, choosing míng as the target syllable can retain the rising information. Otherwise, the information of intonation contour will be lost. Second, the
syllable liáo is chosen because a speaker's voice tends to get creaky by the end of the sentence. This often happens together with weakened signals and a loss of F0 traces. Thus, by not choosing the last syllable as the target, we decrease the chances of not having any F0 information for a certain token.  

This choice of the target might cause a problem because the default lexical stress is claimed to fall on the second syllable in Mandarin Chinese (Chao 1968). However, since we will mainly be concerned with comparing the same syllable under different stress conditions, the effect of the lexical stress, if any, is assumed to be the same.

For each of the target syllables, three acoustic parameters are measured. The measurement of the F0 values is described in Section 4.2. Basically, the point of measurement is the peak or valley of the F0 excursion at a point related to the tone of the target syllable. The point does not necessarily coincide with the syllable boundary. For the duration parameter, however, syllable boundaries are used as the points of measurement. I mainly rely on the waveform for deciding on the location of the boundary. In some doubtful cases, the information from the spectrogram is also used. Lastly, intensity values are taken from the peak of the amplitude within the target syllable. If more than one peak is found, the value at the highest one will be taken.
1.4 Statistical analysis

The results of this study are largely based on statistical analysis of various values taken from the measurement points. The most-used statistics in this study are the analysis of variance (ANOVA), the t-test, and the multiple post-hoc Tukey comparisons. In the following discussion, I will introduce some of the key notions involved in these statistical procedures.

First, like all other statistical analyses, hypothesis testing is central to both t-test and ANOVA. It is mainly used to determine if two or more than two groups of measurement samples have similar means. If the result is significant, not all groups have similar means. On the other hand, an insignificant result indicates that the means being compared are not qualitatively different from each other.

In hypothesis testing, two hypotheses are involved. The one we want to test is called the null hypothesis. It usually takes the form that the mean of a population is equivalent to a claimed value. When two or more than two populations are compared, the null hypothesis will be of the form that the mean values of all the involved populations are equivalent.

When we test the hypothesis, we expect to find that the null hypothesis is not valid; we want to prove an alternative hypothesis. That is, we want to prove that the mean value of
the population is not equal to the claimed value, or, in some cases, it is either smaller or greater than that value. For example, when two populations A and B are compared, the alternative hypothesis can be: (1) A and B are not equal, (2) A is greater than B, or (3) A is smaller than B. The test statistic used to calculate the significance depends on which of the three alternative hypotheses is involved. The result serves as the statistical evidence which can be used according to certain criterion to be explained below either to reject the null hypothesis when the result is significant or, in the case of insignificant values, reject the alternative hypothesis.

In a test by ANOVA, more than two populations can be compared with each other. The null hypothesis is that all populations being compared are not different with respect to their means. The significance of the test depends on the F value, the size of the population, and the number of factors involved, etc. A significant value indicates that at least two of the populations have qualitatively different means. Otherwise, the null hypothesis is valid.

The value to be tested in an ANOVA is called the dependent variable. This variable may be divided into different groups according to factors. These factors are called independent variables. In an ANOVA model, we can test the main effects of the factors or the interaction of these
factors. For example, in the present study, if we want to test the stress effect on the duration of a target syllable, the duration of all the tokens of that target will be set as the dependent variable. Factors to be tested include stress, subject, tone, and possible interactions of these factors. With reference to these factors, individual elements of the dependent variable are then divided into different groups. An insignificant test result means that the variations in the dependent variable does not form a trend that is associated with a certain factor (independent variable). So, that factor has no effect on the variations in the data. For each factor in the ANOVA model, there will be a test result showing the significance of the factor. The same is true for the interaction of the factors.

In a t-test, either the mean value of one population is compared with a set value or two populations are compared with each other. Factors are not involved. The size of the population is usually very small (less than 30). Depending on what the alternative hypothesis is, two kinds of t-test, one-tailed versus two-tailed, can be used. One-tailed t-test is used for the alternative hypothesis of one population having greater or smaller mean values. On the other hand, two-tailed t-test is only appropriate for the alternative hypothesis that the difference between the two population means is not zero, that is, they are not equal.
The significance of the test results for both the ANOVA and the t-test is represented by a value called p-value which ranges from 0 to 1. This value represents the smallest level of significance at which the null hypothesis can be rejected. The level of significance (the α-level) is set by anyone who wants to interpret the result. For linguistic studies, this level is usually set at 0.05. In this study, we will also set our level of significance at this level. Therefore, any p-value smaller than 0.05 will be considered significant enough to reject the null hypothesis.

When more than two groups of data are compared with each other at the same time, I will use the multiple post-hoc Tukey comparisons (i.e., pairwise Tukey comparisons) to perform the task. This statistical procedure assumes that all the data groups have the same standard deviation. The result of the procedure is a confidence interval for each group of data. This provides the basis for the comparison. Overlapping confidence intervals suggest that the data being compared are not significantly different. In this study, confidence intervals are set at 95%. Statistically, it means that, if the same experiment is performed 100 times, the results will fall within the range of the interval for at least 95 times.

In this study, the 95% confidence interval for the multiple post-hoc Tukey comparisons refers to the interval of
the difference between two groups of data being compared. In other words, this interval represents possible range of variations when the mean of the first data group is subtracted from that of the second group. Thus, when the interval includes zero, the data are considered not to be significantly different. I will use two kinds of format for the confidence intervals: $a \pm b$ and $[a, b]$. The former stands for the range from $a - b$ to $a + b$ and the latter from $a$ to $b$.

In Chapters Four and Five, we will examine the three acoustic parameters of F0, syllable duration and intensity with respect to the stress effect. For each of these parameters, I will first carry out the analysis of variance for repeated measures to test the effects of stress and its interactions with the other two variables. Following Clark (1973), I consider "subject" as a random variable and "tone" and "stress" as fixed variables. In other words, we regard the levels of factor "subject" as random samples from a larger population. Hence, the statistical result does not simply apply to the four subjects. It is regarded to be generally true for the larger population from which the subjects are randomly chosen.

When stress and its various interactions are found to be significant, I will perform further analysis on the data. First, the mean value and the standard deviation are
calculated. To save space, I will put the standard deviation immediately after the mean value between parentheses. For ease of reading, I will put most of the lengthy tables of mean values in the Appendices and will use graphs to summarize the data. Second, I will perform the multiple post-hoc Tukey comparisons to determine if the difference between the two is significant.

All the statistical analyses in this study are accomplished with Minitab 8.21 for the Macintosh. The graphic presentations of the mean values are done in DeltaGraph Pro 3.

1.5 Outline

The organization for the rest of the chapters is as follows. In Chapter Two, we introduce previous studies of tone, intonation and stress on Mandarin Chinese. In Chapter Three, we briefly review the literature on focus and present the result of the perception test. Chapter Four is devoted to the analysis of the acoustic parameter F0 and Chapter Five to duration and intensity. In Chapter Six, we will try to correlate subjective perception scores with the objective acoustic values in the multiple regression analysis. In Chapter Seven, we summarize the findings in this study.
Note

1 Broadly speaking, Mandarin Chinese is a cover term for the official language used in mainland China and Taiwan. However, the official names for the language adopted under the two political systems are not quite the same. It is called Pǔtōnghuà (the common language) in mainland China and Guóyǔ (the national language) in Taiwan. Since all my subjects come from mainland China, the term Mandarin or Mandarin Chinese in this study actually refers to Pǔtōnghuà.

2 According to Yuan et al. (1960), an authoritative account of Chinese dialects, Mandarin Chinese can be divided into four sub-dialect groups: Northern, Northwestern, South-western, and Jiang-huai. The standard Mandarin is based on the Northern dialects and Beijing pronunciation, a Northern subvariety of Mandarin Chinese.

3 All twenty subjects were paid for their participation in the study.

4 The vocal cords do not vibrate in producing a voiceless sound. This will show up in the F0 contour as a break.

5 Mandarin syllables are conventionally analyzed as having the structure represented by the following tree diagram (see Cheng 1973).

```
syllable
   /\   
  initial final
     /\      
    medial rime
   /\          
  nucleus vocalic or consonantal ending
```
The main vowel of a syllable is the nucleus. The other elements are all optional. Initials include stops, fricatives, nasals and liquids. There are three possible elements for the medial position, that is, /j/, /w/ and /u/ in the IPA representation. In the pinyin system, they are often represented by y, w and ü if the initial is zero. Otherwise, i and u are used. In this position, where the medial is preceded by an initial, w and ü are in complementary distribution. So only u is used. In this study, both initials and medials are used. For convenience sake, I simply refer to them as consonants without further discussion.

In the actual test, the four choices were given in the order: (a) "what happened", (b) "who", (c) "when", and (d) "do what". I intentionally put choices (a) and (d) at two extremes. There are two reasons for this. First, the two choices differ only from each other by one Chinese character. Keeping them further apart might reduce the probability for subjects to mark the answer they did not intend. Second, the two choices are the closest pair among the four in terms of their presumed places of stress. If they are put too close to each other, the subjects might get confused and might thus lengthen the processing time unnecessarily. In this dissertation, for ease of discussion and presentation, I have switched them back. So, in the analysis below, focus 1 corresponds to score (a), focus 2 to score (b), etc.

F0 traces are partly or totally lost in some tokens even though I have chosen the penultimate syllable as the target.
CHAPTER II

Mandarin stress

In this chapter, we will review previous studies on the topic of stress in Mandarin. As is well known, Mandarin is a tone language. Like other languages, it also has intonation. Since sentence stress is acoustically related to both tone and intonation, I will begin my discussion with a brief survey of previous studies on the topic in this language.

2.1 Tones

Mandarin has four lexical tones, which are frequently referred to as the first tone (T1), second tone (T2), third tone (T3), and fourth tone (T4). Two notations are widely adopted in Chinese linguistics to stand for the four tones.

The first is tone diacritics, used in both pīnyīn (sound spelling) and zhùyīn fúhào (notation for sound representation), two different systems for representing Mandarin sounds. Pīnyīn is used officially in mainland China, and zhùyīn fúhào officially in Taiwan. This system of notation is adapted from Chao's (1930) "tone letters". In
his system, tonal contours are represented graphically. A vertical bar is used to stand for the voice range of a speaker. Tonal contours are drawn on the left side with reference to the vertical bar. For example, the hypothetical tones of "high level", "mid level", "low level", "mid rising" and "high falling" can be represented by the following five tone letters respectively: \( \textup{I} \), \( \textup{I} \), \( \textup{I} \), \( \textup{I} \) and \( \textup{I} \). In the adapted version, however, only the left side of Chao's tone letters is used. The tone marks are thus iconic, graphically representing the actual contours. In pǐnyīn, they are placed over the nuclear vowel (or the nucleus of the syllable) (see Table 2.1 below).¹

The other notation is the numeric representation used in Chao (1930, 1942, 1968).² In this system, the numerals 1 through 5 are used to stand for the pitch range of a speaker with 1 at the lowest extreme and 5 at the highest. A tonal contour is represented by two consecutive numbers which stand for the beginning and end points of the contour. They can be used either in tone-sandhi rules or as a superscript to the syllable over which the target tone is realized. This notation is now becoming more popular in the description of Chinese data. Some dictionaries (for example, Jiānmíng Wúfāngyán Cídìǎn [A Concise Dictionary of the Wu Dialects], Shanghai Cishu Chubanshe, 1988) actually uses this notional system for transcribing tones. The following table
summarizes the four Mandarin tones and the two systems for tonal representation.

Table 2.1 Mandarin tones and notations

<table>
<thead>
<tr>
<th>Tone types</th>
<th>Description</th>
<th>Notations</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>high-level</td>
<td>ma₅⁵</td>
<td>mā</td>
</tr>
<tr>
<td>T2</td>
<td>high-rising</td>
<td>ma₃⁵</td>
<td>má</td>
</tr>
<tr>
<td>T3</td>
<td>low-dipping</td>
<td>ma₂¹⁴</td>
<td>mǎ</td>
</tr>
<tr>
<td>T4</td>
<td>high-falling</td>
<td>ma₅¹</td>
<td>mā</td>
</tr>
</tbody>
</table>

In the above table, the pitch contour for T1 is high-level. It starts high and remains at the high level. T2 is high-rising. It begins at about the middle of the pitch range of a speaker and rises abruptly to the top of the range. T3 in the above representation is the citation form. It starts low and falls to the very bottom of the pitch range and then rises to the mid high level. However, when T3 occurs before a tone other than another T3, it does not have the rising second part. For this reason, it is called half-T3 on those occasions. The pitch contour for half-T3 is thus low dipping. Lastly, T4 is high-falling. It starts at the top of the pitch range and falls abruptly to the low limit.
In addition to the four tones above, Mandarin also has a fifth tone, often referred to as the "neutral tone". The so-called neutral tone is actually no tone. Its surface pitch value is entirely determined by the tonal contour of the preceding syllable or by the sentence intonation. In the Romanization system, this tone is often represented by a dot, a circle or a lack of tone markers. We will come back to the neutral tone in more detail in Section 2.3 below.

Unlike other Chinese dialects in the southern part of the country, Mandarin Chinese does not have a complicated tone-sandhi system. The most frequently occurring tone sandhi is the so-called Half Third-tone Sandhi described by the following rule where only the first half of T3 is realized before T1, T2 and T4 (see Chao 1968):

(2.1) Half-T3 Rule:

\[
214 \longrightarrow 21 / \quad \{ \begin{array}{c} 55 \\ 35 \\ 51 \end{array} \}
\]

Basically, T3 is realized as [21] in the non-final position before another tone that is not T3. The following are some examples for this rule. The second syllable of these examples belongs to T1, T2 and T4 respectively.

(2.2) a. ma^{214} + che^{55} \longrightarrow ma^{21}che^{55} 'horse-drawn cart'
b. ma^{214} + ti^{35} \longrightarrow ma^{21}ti^{35} 'horse's hoof'
c. ma^{214} + li^{51} \longrightarrow ma^{21}li^{51} 'horse power'
When a T3 occurs before another T3, the change is rather dramatic. It now becomes a T2, which can be accounted for by the following rule:

(2.3) T3 Rule:

\[ 214 \quad \rightarrow \quad 35 / \_ \_ \_ 214 \]

The following are some examples:

(2.4) a. \( \text{li}^{214} + \text{jie}^{214} \quad \rightarrow \quad \text{li}^{35}\text{jie}^{214} \) 'understand'

b. \( \text{shui}^{214} + \text{guo}^{214} \quad \rightarrow \quad \text{shui}^{35}\text{guo}^{214} \) 'fruit'

c. \( \text{yu}^{214} + \text{san}^{214} \quad \rightarrow \quad \text{yu}^{35}\text{san}^{214} \) 'umbrella'

The two tone-sandhi rules above all assume that the underlying tone for T3 is /214/. Actually, in connected speech, T3 is rarely realized as [214]. Probably, the only place where the final rising portion of the tone is fully surfaced is the sentence-final position. In this regard, citation tone is only a special case. In my data, in most cases this tone is realized without a final rise even at the end of the sentence. Only when a syllable is stressed does the tone appear as [214], and even then it does not always surface. Thus, it seems to be more appropriate to analyze T3 as being underlyingly /21/ and to derive the other forms from it.\(^3\) In this study, we will regard T3 as a low tone.

It should be noted that T3 sandhi affects only the preceding T3. The following conditioning T3 remains the same unless it is followed by another T3. Thus, in the above examples, the tone of the second syllable remains a T3 no
matter what tones they follow. This is significant for our experiment. It can be recalled from Chapter One that the subject of the test sentence is of the form Lǎo X where Lǎo belongs to T3 and X has all the variations of the four tones. Since the target syllable is X (the name syllable), it only conditions the T3 sandhi when the syllable itself is T3. On the other hand, the tonal contour of the name syllable is not affected by the preceding T3. Besides, since the syllable that follows the name syllable is not T3, it follows that the T3 variation for the target name syllable is fully preserved without being affected by the T3 sandhi process.

Another tone sandhi that is related to the present study is the T2 sandhi. This is related to the changing of a T2 syllable to T1. When a T2 (/35/) syllable is preceded by a T1 (/55/) or T2 (/35/) syllable and followed by a syllable of any other tone except neutral tone, it will become T1 (/55/). This can be accounted for by the following rule.

\[(2.5) \quad \text{T2 Rule:} \quad 35 \longrightarrow 55 / \{ \begin{array}{c} 55 \\ 35 \end{array} \} \longrightarrow \{ \begin{array}{c} 55 \\ 35 \\ 21 \\ 51 \end{array} \} \]

In other words, T2 loses its initial rise when it is flanked by a high pitch. According to Chao (1968), this rule only applies to a three-syllable group. The following are some examples given by him.

\[(2.6) \quad \text{a. } xì^{55} \ yáng^{55} \ shèn^{55} \longrightarrow xì^{55} \ yáng^{55} \ shèn^{55} \quad '(\text{occidental) ginseng}') \]
b. san\textsuperscript{55} nian\textsuperscript{35} ji\textsuperscript{35} \longrightarrow san\textsuperscript{55} nian\textsuperscript{55} ji\textsuperscript{35} '3rd-year grade'

c. han\textsuperscript{35} shu\textsuperscript{35} biao\textsuperscript{214} \longrightarrow han\textsuperscript{35} shu\textsuperscript{55} biao\textsuperscript{214} 'thermometer'

d. hao\textsuperscript{35} ji\textsuperscript{35} zhong\textsuperscript{214} \longrightarrow hao\textsuperscript{35} ji\textsuperscript{35} zhong\textsuperscript{214} 'quite a few kinds'

Chao (1968:28) observes that this kind of tone sandhi only occurs in conversational speed. In more deliberate speech, the original tone is preserved. It is interesting to note that the inherent tones for all the three syllables in the last example belong to T3. The first two tones are actually changed tones after the T3 Rule has been applied simultaneously to them (Cheng 1973). It thus appears that T2 sandhi is a surface rule, the application of which depends on the deliberateness of the speaker. Cheng (1973:48) shows that the rule is not limited to the middle syllable in a three-syllable group as Chao has claimed. It can be applied to larger domains. The following is the example given by him.

(2.7) lǎo 1 Y mái hǎo jiǔ
old Li buy good wine

a. , , , , ,

b. , , , , ,

c. , , , , ,

d. , , , , ,

e. , , , , ,

The above five readings of the same sentence is speed sensitive with (a) being the slowest and (e) the fastest. It
can be seen that T2 sandhi occurs on all three medial syllables at the fastest speed listed by Cheng.

Descriptions of other tone-sandhi rules can be found in the literature, for example, Chao (1968), Liao (1994) and Zhang (1988). They will not be discussed here since the present project is not concerned with tone sandhi in Mandarin Chinese per se.

2.2 Intonation

Intonation in Mandarin Chinese has been receiving more and more attention in recent years. The existing literature on Mandarin intonation deals mostly with how tone interacts with intonation. This is not accidental. Acoustically, both tone and intonation have the same physical correlates: fundamental frequency (F0), intensity and duration (Lehiste 1970). Thus, linguists have been interested in investigating what happens to these physical parameters when tone interacts with sentence intonation.

Unlike English and other non-tone languages, in which the F0 contour is principally shaped by the intonation pattern alone (Beckman and Pierrehumbert 1986, Bolinger 1986, Féry 1992, Finn 1984, Liberman 1975, Pierrehumbert 1980, and Pike 1945), F0 contours in Mandarin also reflects tonal specifications for the component words (Chao 1968). Thus, sentence intonation and the word tonal contours interact with
each other. In some situations, typically when the word is stressed, its contour overrides the intonation. In other cases, the pitch contour of an individual word is overwhelmed by intonation (Liao 1994, Tseng 1981).

Previous models of Mandarin intonation can be basically divided into three groups: pitch range, pitch contour and contour register. Proponents of pitch range theories claim that Mandarin intonation is nothing but a combination of backdrop different pitch range values determined by semantics and/or the grammar of the sentence. Tones are just local pitch perturbations within the given ranges. The contour theories claim that Mandarin intonation is characterized by contrasting contour shapes. These contour shapes provide a backdrop global rise or fall onto which the local word tone contours are superimposed. Lastly, the register theories claim that Mandarin intonation contours are exhibited on different registers according to the grammar and the speaker's attitude.

Gårding's grid model is perhaps the most explicitly stated among all the pitch range theories of Mandarin intonation. In a series of papers (Gårding 1983, 1984, 1985, 1987 and Gårding et al. 1983), Gårding and others actually use the model in an attempt to analyze the Mandarin data. The following figure summarizes the key notions of Gårding's model.
a. Level grid

b. Rising grid

c. Falling grid

Figure 2.1 Grid model
The central notion of this theory is "grid". It consists of two parallel lines which stand for the top and the bottom lines of intonation contour. A grid can be level, rising or falling. In this model, the tones of lexical words are represented by the local pitch contours between the grid lines. The maximum and the minimum values of these contours are called "turning points". Functionally, turning points represent lexical groupings. The direction of the pitch contour is determined by the inherent tone of the lexical items. The grid direction signals speech act type, for example, questions or statements, etc. Semantic information such as focus (to be discussed in the next chapter) is represented by the width of the grid. A grid can be expanded, compressed or even be one line (an extreme example of grid compression).

It can be noted in this model that the top and the bottom lines are assumed to be parallel. So, in both the rising and the falling grids, the amount of rising and falling of the local contours varies with both the bottom and the top lines. The distance between the two lines is equally spaced. When a phrase is focused, the grids will expand. This creates greater distance between the top and the lower turning points. On the other hand, when the grids are compressed, the distance in between becomes smaller. Thus, the whole intonational contour is viewed as the local pitch
variation within grids which may undergo expansion or compression.

A general observation of Gårding's model indicates that tone and intonation are represented in the different parts of the model. It can be seen that pitch ranges are mainly used to carry tonal information, while the global intonation pattern is indicated by the direction of the grid. In this sense, we can regard her model as a hybrid of pitch range and contour approaches.

Shih (1988:93) found that pitch range variation with respect to prominence is not the same for the high and the low targets. While a high target can become much higher, a low target will remain at the same level or slightly lower. Shih's observation seems to suggest a model in which the bottom line is fixed and the top line is moveable.

Chao (1968) basically takes the contour approach in his description of Mandarin intonation. To describe the relationship between word tones and sentence intonations, he compares them to small ripples riding on big waves. When the two are combined, "the result is an algebraic sum of the two kinds of waves. Where two pluses concur, the result will be more plus; when a plus meets a minus, the algebraic addition will be an arithmetical subtraction" (Chao 1968:39). Thus, depending on the direction of the global intonation, a local tone contour will be intensified if it glides in the same
direction as the global intonation. On the other hand, when the pitch directions of tone and intonation differ, the two will counterbalance each other.

Shen (1990a) is another important study on Mandarin intonation in recent years. To her, intonation in this language is also manifested by pitch contours, however, not necessarily on the same pitch level. Following DeFrancis (1976), she claims that intonation contours in Mandarin are exhibited on two separate registers, an upper one for questions and a lower one for statements. Tones are local pitch variations on these two separate levels. Her conclusion is based on the results of an experiment that takes into account six different sentence structures: statements, unmarked questions, particle questions, A-not-A questions, alternative questions and Wh-questions.5

In another study, He and Jin (1992) even claim a system in which intonation contours vary on three different registers: high, low and extra low. Their results are based on an experiment that examines the subjects' different intonational patterns in reading the target sentences with various "moods", for example, declarative, interrogative, imperative and exclamatory. As compared with Shen (1990a), they are more concerned with the intonational patterns which reflect a speaker's attitude. For example, they only consider two kinds of interrogative sentences which are
differentiated by the speaker's assumptions. In the first kind, the speaker assumes that something is not true. He is more doubtful than he is willing to believe. For example, in chōu yīhè yán jǐukuāiduō qián? (It costs more than nine dollars to smoke a packet of cigarettes?), the speaker assumes that what he is asking is not true. The question has the overtone that he does not believe such a thing. The intonation contour for this kind of question is found on the highest register. The other type of questions is related to the positive assumption from the speaker. When the question is asked, the speaker is almost sure what the answer might be. Here, the question is used as a means of confirmation. The expectation is that the answer would be positive. For example, in Shéiyà? Xiāolǐ? (Who is it? Xiao Li?), the question Xiāolǐ? is asked simply for confirmation. The speaker assumes that the answer would be yes. The intonation contour for this kind of interrogative question is found on the extra low register. The other sentences (declarative, imperative, exclamatory) all have intonation contours on the low register.

In this study, I will assume the pitch range model of intonation. It can be seen above that neither Shen (1990a) nor He and Jin (1992) are specifically concerned with stress. They only examined sentence patterns with different grammatical meanings. To a certain extent, He and Jin's
study addresses the semantic aspect of the intonation from the point of view of the speaker's attitude. However, their point of interest is mainly limited to concerns of expressiveness of intonation rather than on the information and focus structure part. On the other hand, the main object of the present study is the part of the sentence stress that is related to information and focus. Previous studies which we will examine in the next section all show that, when a syllable is stressed (or focused), its pitch range is expanded. A pitch range expansion model offers a ready tool for the description of the phenomenon. On the other hand, the target sentences to be examined in the present study all belong to the same grammatical category of statement and is rather neutral with respect to the expressive part of the speaker's attitude. Thus, the conclusions reached by Shen (1990a) and He and Jin (1992) will not affect us because, if register is really a parameter of intonation, all our target sentences will be assigned the same parameter and the results will not be affected.

2.3 Stress

The study of stress in Mandarin Chinese, according to Li (1981), began with Yuen Ren Chao's Guóyǔ Luómǎzī Yánjiū [The Study of Romanization in Mandarin Chinese] (1922) when he, for the first time, put forward the idea of qǐngshēng
'neutral tone' in Chinese linguistics. In 1932, following the British tradition, he further defined stress as corresponding to "the tonic stress" in a tone group (see Shen 1993b:416). By 1968, Chao's remarks about stress in Mandarin Chinese became more specific:

Stress in Chinese is primarily an enlargement in pitch range and time duration and only secondarily in loudness. Thus, when a 3rd Tone is stressed it is dipped lower, and when a 4th Tone is stressed it starts higher and falls lower.

(Chao 1968:35)

Stress in Mandarin, according to Chao (1968), can be classified into three categories: normal, contrasting and weak. Normal stress is a stress that is neither contrastive nor weak. Contrasting stress is what we call contrastive stress. It has wider pitch range and longer duration, usually with associated increase in loudness. For the weak stress (i.e. neutral tone), "the tone range is flattened to practically zero and the duration is relatively short"(35).

In the literature, Mandarin stress is in most cases discussed in reference to the weak stress on the lexical level. For disyllabic words or phrases, two stress patterns, iambic (weak-strong) versus trochaic (strong-weak), can be recognized. Neutral tones only occur in the weak position of the trochaic rhythm. The following are two examples:
(2.8) a. iambic  
   weak-strong  
   châoguō  'frying pan'

   b. trochaic  
   strong-weak  
   xuēshēng  'student'

In the above example, only sheng surfaces as a neutral tone. The morpheme chāo 'fry' retains the inherent tone even though it is in the weak position.

The pitch level of the neutral tone is determined by the inherent tone of the preceding syllable. Basically, it is either the result of a low-tone insertion or the rightward spreading of the tonal contour of the preceding syllable. In this case, the behavior of the neutral tone in Mandarin strikes a strong resemblance to the tone-sandhi behavior of the Wu dialects (Chan 1991b, Chan and Ren 1989, Kennedy 1953, Sherard 1972, and Wright 1983) and the neutral tone in Fuzhou (Chan 1985, Hung 1989). For example, in the following examples, the pitch values of the neutral tone after syllables of the four different tones vary according to the type of tones they follow (from Chao 1968:36):

(2.9) a. ta₅₅de²  'his'

   b. huang₃₅de³  'yellow'

   c. ni₂¹de⁴  'your'

   d. da₅¹de¹  'big'

In the above examples, the tone of the morpheme de is neutral. It can be seen that the exact pitch value of this tone is not fixed. It is determined by the tone of the preceding syllable. When the preceding syllable ends with a
high tone, it surfaces as a low tone. On the other hand, when the preceding tone is low, the inherent tonal contour of the preceding syllable is carried over to the following neutral tone. In terms of the duration of the neutral tone syllables, Chao (1968:35) also observes that they are relatively short.

The description of the neutral tone phenomenon by Chao (1968) has been confirmed acoustically to be generally true by Lin and Yan (1980). Their results are based on 29 pairs of disyllabic words with both iambic and trochaic rhythms. In some cases, a difference of meaning is involved depending on what rhythm is used. They find that when a syllable is pronounced with a neutral tone (second syllable in the trochaic rhythm), its duration is only about half of the preceding stressed syllable. Intensity of the neutral-tone syllables is also found to be reduced in most cases.

Their results for the behavior of the F0 excursion of the neutral tone are summarized by using the numeric notation for the tones. Unfortunately, the exact procedures of how they arrive at these tonal values are not explicitly given. Presumably, these are the results of rough estimation from a comparison of the F0 excursions between two related rhythmic patterns since sixteen pairs of F0 excursions of such words are given at the end of their paper. Their description of
the acoustic data, presented in the following table, generally conforms to the one given by Chao (1968).

Table 2.2 Tonal contours of trochaic words

<table>
<thead>
<tr>
<th>Two subjects</th>
<th>1st syllable (strong)</th>
<th>2nd syllable (weak)</th>
</tr>
</thead>
<tbody>
<tr>
<td>male</td>
<td>T1 55</td>
<td>Tx 41</td>
</tr>
<tr>
<td></td>
<td>T2 35</td>
<td>Tx 51</td>
</tr>
<tr>
<td></td>
<td>male</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>T3 322</td>
<td>Tx 44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>T4 51</td>
<td>Tx 21</td>
</tr>
<tr>
<td>female</td>
<td>T1 55</td>
<td>Tx 41</td>
</tr>
<tr>
<td></td>
<td>T2 35</td>
<td>Tx 51</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>T3 312</td>
<td>Tx 33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>T4 51</td>
<td>Tx 21</td>
</tr>
</tbody>
</table>

In the above table, Tx refers to an inherent tone of any type. Although there are minute differences between the two speakers, the general tendency is clear. It can be observed that the characteristic feature of the second syllable in
such words is the loss of the inherent tones since all tones in this position have the same tonal value when they follow a syllable with a specific tone. Basically, the surface pitch contour of the weak second syllable is determined by the tonal contour of the first syllable. Roughly, it is an extension of the tonal contour of the preceding syllable in citation form plus final fall in most cases. In contrast to the neutral tone of the second syllable, the tone of the stressed first syllable is fully preserved.

Neutral tones not only occur on the lexical level, they can also be found on the sentential level. In the latter case, they are typically associated with grammatical morphemes, and with morphemes, the meanings of which are largely predictable from the discourse environment (Hung 1989). The same phenomenon is found in other Chinese dialects (Hung 1989, Jin 1986). Since grammatical morphemes are rarely stressed (Tao Lin 1962), one would naturally expect that neutral tones at the sentence level are weakly stressed. On the other hand, semantically prominent words, that is, focused words (to be discussed in the next chapter), are usually stressed.

Wu (1982b) discusses a case where T3 sandhi is affected by sentence stress. The result is that the intonation of the whole sentence is affected. The following three sentences are taken from Wu (1982b:442):
(2.10) a. Wǒ xiǎng qǐng nǐ gěi wǒ mǎi liǎng bǎ hǎo
    I want ask you for me buy two CL good
yǔsān, bú yào dēngwài pǐn.
umbrella not want substandard product
'I want to ask you to buy two good umbrellas for me. Don't buy the substandard ones.'

b. Wǒ xiǎng qǐng nǐ gěi wǒ mǎi liǎng bǎ hǎo
    I want ask you for me buy two CL good
yǔsān, yídīng yào mǎi lái.
umbrella make sure must buy DIRECTION
'I want to ask you to buy two good umbrellas for me. Make sure you buy them.'

c. Wǒ xiǎng qǐng nǐ gěi wǒ mǎi liǎng bǎ hǎo
    I want ask you for me buy two CL good
yǔsān, bìé gěi tā mǎi.
umbrella do not for him buy
'I want to ask you to buy two good umbrellas for me. Don't buy them for him.'

Although all the above three sentences express the same request, their focus is different. As a result, stress falls on different parts of the sentence. In (2.10a), hǎo 'good', in (2.10b), mǎi 'buy', and in (2.10c), the second wǒ 'me' are stressed. Note that all the syllables in the first part of these sentences are T3. In the neutral reading where none of the words are stressed, that part of the sentence will have the following tone-sandhi domain (Shih 1986, Hung 1989, Mok 1993).
(2.11) [Wǒ xiǎng] [qǐng nǐ] [gěi wǒ] [mǎi liǎng bǎ]
I want ask you for me buy two CL
[hǎo yǔsǎn]
good umbrella

Given this domain assignment, the words, mǎi 'buy' and hǎo 'good' will both surface as a low T3, while the second wǒ 'me' can be either realized as T3 or T2. However, when it is stressed as in the sentence environment of (2.10c), it will surface as T2.

It has been generally observed in Chinese linguistics that lexical tones, when put in a sentence environment, will be modified by the intonation of the sentence. When the local tone glides in the same direction as the sentence intonation, its F0 value will increase. On the other hand, when the local tone and the global sentence intonation are at odds, the values of the two will be counterbalanced (Chao 1968, Ho 1976a, 1977, Liao 1994, Shen 1990a, Tseng 1981). When syllables are stressed, their tonal values are fully realized, while weakly stressed syllables are usually overridden by sentence intonation. Thus, in the post-stress position, a syllable will typically lose its inherent tone, become tonally neutralized, and conform to the general falling F0 contour of the sentence (Liao 1994, Tseng 1981).

As has been noted in the preceding section, this is the typical position taken by the contour theory. An alternative account in the pitch range theory is that when a syllable is
stressed, its pitch range will expand. On the other hand, the range for a weakly stressed syllable will be reduced. In the post-stress position, this range will be further reduced to zero. So, the inherent tonal contour of the syllable is not realized.

As far as the data in (2.10) are concerned, it is noted by Wu (1982b) that, when hǎo 'good' and wǒ 'me' are stressed, the immediately following syllable becomes neutral-toned. This is obviously related to their post-focus status. The F0 traces of gěi wǒ mǎi liǎng bā hǎo yǔsān 'buy two good umbrellas for me' given by Wu clearly show that the inherent tones of the two syllables yǔ 'rain' and mǎi 'buy' are lost. The F0 excursions for these two syllables basically conform to the general sentence intonation.

However, when mǎi 'buy' is stressed as in the case of (2.10b), the following syllable liǎng 'two' fails to become neutral-toned. The corresponding pitch track given by Wu shows that, for this particular syllable, the F0 peak falls on liǎng 'two' rather than on mǎi 'buy'. Thus, it seems to be the case that liǎng 'two' is not considered to be the neutral tone because its tonal contour is not falling after the stressed syllable.

When the fundamental frequency contours of the three sentences are compared, it can be found that when mǎi 'buy' is stressed in (2.10b), a noticeable peak that coincides with
the syllable yǔ 'rain' can still be found. However, when wǒ 'me' is stressed in (2.10c), the F0 exhibits a general fall after the stressed syllable without any obvious peaks. This seems to suggest that, when a word toward the beginning of an intonational phrase is stressed, the resulting falling intonation will override all the lexical tones of the following words. This implies that stressing the sentence initial word will reduce the rise of all the following words to nothing.

This phenomenon is also observed in Gårding (1985, 1987). In her model, syllables immediately following a focused word exhibit a sharp fall in F0 values.

In the present study, we will also examine the F0 excursion of syllables that immediately follows a stressed word. Our result shows that, in some cases, the local tonal contour of a lexical word is not fully overridden by the falling intonation. Noticeable rising peaks can still be observed. This suggests that a pitch-range reduction model, rather than a contour model, is appropriate to describe the Mandarin data.

So far, we have considered the F0 parameter with respect to stress. The results of previous studies show that duration and intensity are also involved in sentence stress as well as in lexical stress.
Shih (1988) reports that, in addition to the pitch range expansion, duration and intensity are both involved in stress.

It is apparent from the figures that prominence is reflected by expanding pitch range: high targets become much higher, while low targets remain at the same level or are slightly lower. Aside from the increased pitch range, more prominent forms also have longer duration and higher intensity.

(Shih 1988:93)

In addition to this generalization, she also observes that, although the stressed syllable is always longer in duration and higher in intensity, the higher pitch sometimes falls on the adjacent but less prominent syllable. This shows that the pitch contour peaks of focus may not be coterminous with the stressed syllable.⁹

Liao's (1994) description basically agrees with the other studies; that is, both pitch range expansion and duration are involved in focus. Generally speaking,

[In] an intermediate phrase, the pitch curve of the stronger syllable is usually longer or spans a larger pitch range than that of any other syllable.

(Liao 1994:263)

Shen (1993b) is another recent work on Mandarin sentence stress. In her study, subjects are presented with three sets of pre-recorded and manipulated perception data. They are asked to mark the place of stress as the utterance is originally spoken. In the first set, all three acoustic
parameters -- F0, duration and intensity -- are present. In the second set, the information of the original F0 is removed by replacing it with a constant 135 Hz through a filter. So the subjects only have access to intensity and duration. In the third set, intensity and F0 are removed. As a result, the only acoustic parameter available to the subjects is duration. In addition to these variable controls, segmental information with respect to the test sentences in all the three sets is also masked by using the low-pass filter at a cut-off frequency of 400 Hz. Thus, the subjects cannot possibly use the manipulated segmental information as additional cue in determining the place of stress.

The result shows that there is no significant difference in the subjects' response to the three sets of data. This indicates that the F0 and intensity in sets one and two are redundant to the subjects when they judge a certain syllable as being stressed. Thus, she concludes that F0 is not crucial for identifying stress. Of the other two parameters, duration is more important than intensity as far as perception of stress is concerned.

In an earlier study, Tseng (1988) examined the disyllabic stress pattern in Mandarin and had reached a similar conclusion. She finds that "the main difference between emphatic and non-emphatic appears to be in the domain
of syllable duration rather than a wider F0 range or more energy information" (610).

The importance of this durational aspect of stress is further supported by Chan (1992) from a different perspective. In her study, she examines the correspondence of Mandarin words as they are read and sung in six Mandarin TV and movie songs. Her findings indicate that, while pitch contours of words can sometimes be overridden by the tune of music if the two are not compatible, the length of those words is largely preserved. It is shown that full-toned grammatical morphemes usually occur as suffixes in songs. Her data also show that neutral-tone syllables are almost never sung on a note that is longer than the full-toned stressed syllable in the same word or phrase. This is similar to the fact that these morphemes do not receive phrasal stress in an utterance. Based on this evidence, she suggests that duration plays a very important part in Mandarin stress. At least in songs, duration is a crucial acoustic correlate of stress.

Although the general conclusion in the literature seems to favor the view that intensity is not related to stress in Mandarin, there is at least one paper in which the two co-authors imply that intensity is of primary importance in stress. In their conclusion, He and Jin (1992:93) write:
The strong stress in Beijing dialect not only is characterized by an increase of intensity, but under usual circumstances, it is also characterized by the expansion of pitch range and the lengthening of duration.

In this study, we will use multiple regression models to examine the relative importance of the three acoustic parameters for sentence stress. Our result shows that F0 is more important than duration in indicating sentence stress.

To summarize the discussion of stress in Mandarin Chinese, neutral-tone syllables most often occur immediately after a stressed syllable. This is certainly the case on the lexical level. On the sentential level, although no definitive conclusion can be made, there is evidence in the literature that the loss of inherent tone which is induced by weak stress is found after sentence stress, especially when the stress falls on the beginning of an intonational phrase. In terms of acoustic parameters, F0, duration, and intensity are all involved in lexical as well as sentence stress. However, there is still controversy as to the relative importance of the three involved parameters.

In the following chapters, the word "stress" refers to sentence stress, not lexical stress. We will examine the effects of this stress on the three acoustic parameters across different sentence positions. It will be shown that these parameters do not behave in the same way in these positions. We will also examine the relationship between the
different acoustic parameters as they are affected by the various stress patterns on the one hand and the subjects' perception to these different stress conditions on the other. On the basis of these information, we will reach our conclusion about the acoustic nature of sentence stress in Mandarin Chinese for our data.
Notes

1 The phonetic symbols used in this dissertation are pinyin since they have wider acceptance.

2 Chao (1968) mainly uses this notation for describing pitch values of tones. His language examples are coded in a different Romanization system devised by him, in which tones are incorporated into the syllables by using different letters. For example, two syllables with the same segmental phonemes but with different tones are spelt in two different ways.

3 This position is actually adopted by early structuralist grammarians such as Hartman (1944) and Hockett (1947). Yip (1980) also uses this form.

4 I am indebted to Professor Mary Beckman for pointing that out to me.

5 Here, unmarked questions refer to the kind of questions that do not differ grammatically from statements. For example, "He is here?" is a question that can be recognized orthographically by the use of a question mark. However, as far as word order and choice of words are concerned it is not different from a statement. Particle questions refer to questions containing the sentence-final question marker ma. In Mandarin, this particle is typically used to form a yes-no question. A-not-A questions are rather unique with the Chinese language. In Mandarin as well as other Chinese dialects, a verb or an adjective can be used together with the negative form of that verb or adjective to form a yes-no question. For example, in the sentence ta qu 'he is going', qu 'go' is a verb. The negative form of that verb is bu qu 'not go' as in ta bu qu 'he is not going'. These two forms can be combined to form a question: ta qu bu qu? Alternative questions refer to questions using haishi 'or'. Lastly, Wh-questions refers to questions asking why, what and when, etc.

6 This three-degree stress system has been widely accepted among Chinese linguists (Yin 1982:168).
In the case of [mǎi liǎng bǎ] 'buy two [umbrellas]' and [hǎo yǔsăn] 'good umbrella', the second and the third syllables are semantically closer, and so, tone sandhi between these two syllables occurs first. This operation will change the second syllable to T2, thus bleeding the tone-sandhi environment for the first syllable. In the case of gěi wǒ 'for me', tone sandhi can occur in fast speech as long as the tone of the following syllable mǎi 'buy' does not change to T2.

This fact can be accounted for by tonal coarticulation (see Shih 1988).

This observation is further confirmed in my data and affected the way I measured pitch contours of the focused words.
Chapter III

Focus and stress

In this chapter, we will first introduce the concepts of focus and stress by briefly reviewing the related literature. Then, in Section 3.2 below, the results of the perception experiment will be discussed with respect to the two concepts. Consequently, the focus and the stress patterns for this study will be established for subsequent discussions in the following chapters.

3.1 Background

Generally speaking, focus refers to the property of semantic prominence which is determined by pragmatic and discourse contexts, while stress belongs to the phonological component of the grammar and is responsible for the phonetic manifestation of focus. From the speaker's point of view, focus is first determined, and then, stress is assigned accordingly. From the standpoint of the hearer, the process is just the opposite. He is first given the place of stress,
and then, focus is determined on the basis of that information.

Phonetically, a focused word or phrase is usually stressed. This contrasts with the other parts of the sentence that are not stressed. Similarly, on the semantic level, the same distinction can be found. For example, Jackendoff (1972) uses "focus" and "presupposition" to characterize the same contrast in meaning. "Focus" refers to that part of the information assumed by the speaker that the hearer does not share. "Presupposition", on the other hand, refers to the part of the information which the speaker assumes to be shared by him and the hearer.¹

The relationship between focus and non-focus (or presupposition in Jackendoff's term) on the one hand, and stress and non-stress on the other seems to be straightforward at first glance. The focused words or phrases are stressed and the non-focused ones are not. For example, in the following exchange, the focused word and stress coincide.

(3.1)  a. Did Maxwell kill the judge with a HAMMER?
       b. No, he killed him with a ZAPGUN.

In these two sentences, the speaker of (a) assumes that the speaker of (b) is aware of the fact that Maxwell killed the judge with something. So the focus of his question falls on the word "hammer". Consequently, that word is stressed. In response to this question, the speaker of (b) stresses the
word "zapgun", indicating a semantic prominence by bringing it into focus. In both cases, the focused word is stressed, and conversely, the stressed word is focused.

However, this one-to-one relationship between focus and stress does not work in all cases. Specifically, when the whole sentence is focused — that is, when the speaker assumes that the information of the sentence is not shared by the hearer — it is not the case that all the words in the sentence are stressed. The usual case is that only the last word is stressed. This kind of focus is referred to by Ladd (1980) as "broad focus" and the related stress as "normal stress". For example, in answering the question "What happened?", we can say "John left" by placing nuclear sentence stress on "left", following Chomsky and Halle's (1968) Nuclear Stress Rule (NSR). If we put stress on "John" as in "Jóhn left", then the sentence would sound as if it is used to answer the question "Who left?'.

In the literature, the discussion of focus and stress generally falls under two approaches: syntactic and semantic (see Ladd 1980). In the syntactic approach, the relationship between focus and stress is viewed as something that can be accounted for by syntax. However, the semantic approach proponents, Bolinger being the major figure, claim that the place where stress goes is purely determined by where the information focus is and has nothing to do with syntax.
Bresnan (1971) is a good example of the syntactic approach to stress. She attempts to reconcile Chomsky and Halle's NSR to the following types of sentence pairs originally pointed out by Newman (1946).

(3.2) a. George has plans to leave.

b. George has plans to leave.

(3.3) a. Helen left directions for George to follow.

b. Helen left directions for George to follow.

The (a) pattern in each of these pairs can be spoken out of the blue without indicating any special presupposition about the unstressed phrase "to leave" or "to follow". That is, the focus domain is the whole sentence. On the face of it, the place of stress seems not to be determined by syntax since (a) and (b) sentences have similar surface structures, so that according to the NSR, stress should always fall on the last word under broad focus. Bresnan points out, however, that the (a) sentences are actually syntactically different from the (b) ones. In the underlying structure, the words "plan" and "directions" are the direct object of "leave" and "follow" in (a), but not in (b). Thus, she suggests that the NSR be applied cyclically to underlying S's and NP's.

Bolinger (1972), by contrast, argues that sentence accents (or stress) are determined by information focus rather than by syntax. No systematic relationship can be
found between the structure of a sentence and its accents. A speaker can choose any part of the sentence to be the point of his information focus and this will determine the place of accents for that sentence. For example, in "Maxwell killed the judge with a hammer", a speaker can choose to stress any NP's or the verb depending on where his information focus is. This has nothing to do with the structure of the sentence. As to the examples discussed by Bresnan, he argues that they are also related to the inherent meaning rather than to the structure. Other counter-examples to the NSR are given to illustrate the point. The following are some of them.

(3.4) a. The end of the chapter is reserved for various problems to computerize.
   b. The end of the chapter is reserved for various problems to solve.

(3.5) a. I have a point to emphasize.
   b. I have a point to make.

In these examples, (a) and (b) have the same underlying structure, but they obviously have different stress patterns. Bolinger argues that the difference lies in semantics rather than in syntax. For the (a) sentences, the phrases "computerize problems" and "emphasize a point" are less frequent or predictable than "solve problems" and "make a point" in the corresponding sentences in (b). When a speaker decides to choose one particular phrase over another, he already has decided where his information focus will be and
that will be the place of his accent. In other words, stress falls on the point of information focus.

Other examples are given to illustrate the same point. In examples like "bóoks to write", "wórk to do", "clóthes to wear", "fóod to eat", "léssons to learn", "gróceries to get", the verbs are highly predictable. On the other hand, stress is more likely to fall on less predictable verbs. Hence, it is more natural for primary stress to fall on verbs in those cases, for example, "pássages to mémorize", "It's time to léave", etc.

Similar discussions can be found concerning stress in Mandarin Chinese. It was pointed out in the last chapter that neutral tone syllables are weakly stressed. Although in many cases the occurrence of the neutral tone is limited to certain syntactic structures (for example, various grammatical markers cannot be stressed), neutral tones do occur in an NP VP structure similar to the English one. If the meaning of the verb is highly predictable from the meaning associated with the NP, then that VP can be weakly stressed and becomes a neutral tone.

This phenomenon is first discussed in Meng (1982). He finds that the verb kàn 'read' in mǎi běn shū kan 'buy a book to read' is often realized as neutral tone. Other examples include the following.
(3.6) zuò yīshāng chuan
make clothes wear
'make clothes to wear'

(3.7) jiè liǎng chē qí
borrow CL bicycle ride
'borrow a bicycle to ride'

(3.8) qī hú chá he
brew CL tea drink
'brew a pot of tea to drink'

Syntactically, all the examples above involve the structure VOV. The second verb can be optionally read with a neutral tone. Meng (1982:58) analyzes these cases in terms of syntax. He points out that in order for the second verb to become neutral-toned in the above sentences, the object of the first verb must be able to serve semantically as the object of the second verb.

This point is picked up and further developed by Hung (1989). He points out that the second verb in the three examples above does not add much information to the whole sentence. The meaning of the whole sentence can be largely retained and the sentence is still acceptable to a native speaker even if the second verb is dropped altogether. This shows that the meaning of the second verb is highly predictable from the meaning of the noun it modifies. On the other hand, if the word contributes its own meaning to the whole sentence and deleting it would either yield a different interpretation of the sentence or make it ungrammatical, then neutral tone cannot be used. The following are two examples:
(3.9) a. shū méi rén mǎi
    book not-have person buy
    'There's no-one to buy the book.'

   b. shū méi rén *(mai)
    book not-have person buy

(3.10) a. qián ràng tā huā
    money let him spend
    'Let him spend the money.'

   b. qián ràng tā *(hua)
    money let him spend

In the above examples, the verbs mǎi 'buy' and huā 'spend' cannot be deleted because the information they carry contributes to the meaning of the whole sentence. In those cases, the meaning of the verb is much less predictable from the inherent meaning of the noun it modifies. Thus, they cannot be so weakly stressed as to be produced with a neutral tone.

However, Bolinger's semantic approach to focus and stress also has its problems. These are discussed in Schmerling (1976). Although she agrees with Bolinger (1972) on the point that there is no systematic relationship between the structure of a sentence and its stress contour, she finds Bolinger's claim that accented words are points of information focus problematic. The examples she gives are related to the following sentences:

(3.11) Trūman díed.

(3.12) Jōhnson díed.
These two sentences were actually taken from her real life experience. She observed that (3.11) was used in a situation in which President Truman's death was everybody's guess whenever his name was mentioned. At the time, Truman was hospitalized. Because of his advanced age, it was reasonable to assume that he would not survive the crisis. In this context, the word "die" has little semantic weight. Yet, when the word finally came that the President died, she was told the news by her mother who used the stress pattern indicated in (3.11). In contrast, she was told about President Johnson's death when she least expected it. The news came as a surprise to everybody. Yet, her husband, who broke the news to her used the stress pattern indicated in (3.12). According to Bolinger's claim, the verb "died" in (3.11) should not be stressed since it cannot be the information focus. On the other hand, the same verb in (3.12) should be stressed because this verb carries the information focus. However, Bolinger's prediction is just the opposite of what is shown in the above two examples.

In order to account for the difference of the two sentences, she relies on different principles using notions from grammar, logic and discourse. Sentence (3.11) is a topic-comment structure, in which case both the topic and the comment receive the same degree of stress. Then, one of her principles makes the necessary adjustment which results in
the last syllable receiving heavier stress. On the other hand, (3.12) is not a topic-comment structure, and is thus analyzed as subject and predicate. In this case, the verb receives lower stress than its argument. This explains why the subject of this sentence receives heavier stress.

However, Schmerling's analysis is also not free from problems. In her book, she admits in note 7, Chapter Two, for example, that her analysis cannot account for the stress pattern exhibited in the following simple sentence.

(3.13) John is a wondrous man.

She is aware that this sentence is also a counter-example to Bolinger's analysis. Since the meaning of "man" is fully predictable from the word "John", it should not be stressed. Ladd (1980) points out that Schmerling could not account for this simple case because she denies the existence of normal stress, which can be easily handled by the syntactic approach.

The relationship between focus and stress is discussed in Hailiday (1967), Chomsky (1970), Jackendoff (1972), Ladd (1980), Selkirk (1984) and Gussenhoven (1983a, 1983b). Although there is no consensus among them about the existence of normal stress, these papers all agree that while most accent placements leave focus narrow (i.e., focus falls on one particular word), there is a particular one which will
define an unspecified domain of focus called "broad focus" by Ladd (1980).

Basically, both Chomsky (1970) and Jackendoff (1972) try to maintain the position that when the focus domain is broad, the position in which the sentence stress is going to fall can be predicted by the configuration of the constituent structure. Building on Chomsky (1970), Jackendoff (1972) uses the primitive marker F to designate the focus domain. The marker is attached to the syntactic trees and is, therefore, supposed to be given. In other words, it is determined by the meaning. The focus domain can be defined as the lexical items the marker F dominates. The place of stress for a particular focus domain is basically determined by the NSR; that is, primary stress falls on the word dominated by the rightmost constituent. The following is an example in Chomsky (1970) that is discussed in Jackendoff (1972).

(3.14) was it an ex-convict with a red shirt that he was warned to look out for?

In this sentence, the word "shirt" is stressed. However, it can have multiple focus domains in that the stressed word is the rightmost constituent for different higher nodes, that is, nodes dominating "shirt", "a red shirt", "with a red shirt", "an ex-convict with a red shirt". Corresponding to these four different domain interpretations, the following
four sentences are possible answers to the question above in (3.14).

(3.15) a. No, it was an ex-convict with a red tie that he was warned to look out for.

   b. No, it was an ex-convict with a carnation that he was warned to look out for.

   c. No, it was an ex-convict wearing dungarees that he was warned to look out for.

   d. No, it was an automobile salesman that he was warned to look out for.

However, when the focus domain is the whole sentence, the stress does not fall on the rightmost constituent in (3.14). Jackendoff explains that the particular syntactic structure of this sentence is "it was NP that S". The emphatic element is NP, not S. Thus, the primary stress has to fall on a word that is within NP, but not within S.

Although Jackendoff is able to explain why stress does not fall on the last word of sentence (3.14) for the broad focus reading, it is still difficult to account for the stress pattern for the following sentence noted by Halliday (1967:208).

(3.16) John painted the shed yesterday.

This sentence can be used to answer the following two questions:

(3.17) a. What did John paint yesterday?

   b. What happened yesterday?
The second question signifies broad focus. However, when "yesterday" is stressed as in (3.18) below, it only has narrow focus reading, not broad focus.

(3.18) John painted the shed yesteray.

(3.19) a. When did John paint the shed?
b. "What happened?"

According to Chomsky and Jackendoff, there is no reason why "yesterday" cannot be stressed for a broad focus reading since the word is the rightmost constituent of the whole sentence.

Halliday himself does not explicitly explain why (3.19b) is not an acceptable interpretation of the focus for (3.18). However, when commenting on the difference between marked (narrow) and unmarked (broad) focus, Halliday (1967:209) observes that "marked focus may be focus on a reference or other closed system item, whether final or not". On the same page, he states explicitly that the word "yesterday" is a reference item. Thus, his explanation of (3.19b) seems to be that "yesterday" cannot bear stress for a broad focus reading because it belongs to a particular class of words.

Ladd (1980) approaches the problem from a slightly different perspective by using a vaguely defined notion of "accentability". In his system, accentability is relative, rather than absolute, within the focus domain. Generally speaking, content words are higher on the accentability
hierarchy than function words. Similarly, nouns are higher on this hierarchy than verbs. For a particular focus domain, accent goes on the most acceptable syllable. In the case of (3.14), "shed" has the highest degree of acceptability. Thus, it ends up being accented in a broad focus interpretation.

Selkirk (1984) also discusses the problem. Her system is more complicated. Her focus structure represents a level that mediates between intonational structure and the focus related intonational meaning. The relationship between the focus structure and the intonational meaning is accounted for by the Focus Interpretation Principle, which basically says that a syntactic node marked by the focus (F) feature is equivalent to the new information of the sentence. On the other hand, she uses the following two rules to connect the intonational structure and the focus structure.

(3.20) a. Basic Focus Rule

A constituent to which a pitch accent is assigned is a focus.

b. Phrasal Focus Rule

A constituent may be a focus if (i) or (ii) (or both) is true:

(i) The constituent that is its head is a focus.

(ii) A constituent contained within it that is an argument of the head is a focus.

Rule (a) only applies to a word or smaller unit. Pitch accent refers to abstract pitch levels that determine
intonational contours of a sentence. Roughly speaking, we may say that the primary sentence stress falls on the word that bears the pitch accent.

When a word is assigned the F feature by the Basic Focus Rule, Phrasal Focus Rule can be used to determine if it is possible to assign the same features to the other nodes. This possibility is related to the argument-head relationship. Basically, when a V or N is focused, its mother node VP or NP can also be focused (3.20b[i]). When an object NP is focused, the dominating VP, which is the head of the argument NP contained in the VP, can also be focused (3.20b[ii]). However, subject focus does not imply VP focus since the subject NP is not contained in VP. In the case of (3.16), when the NP "shed" is focused, it is also possible that the dominating VP "painted the shed" is focused. The same process cannot be applied to "yesterday" if it is focused, since it is not an argument of the verb. Thus, when "yesterday" is stressed, it can only have narrow focus. Similarly, when the subject "John" is stressed, its F feature cannot be inherited by the VP node. In that case, NP is the highest focus node implying narrow focus on "John". In contrast, when "shed" is stressed, the F feature can percolate upward to the VP node, resulting in broad focus.

Gussenhoven (1983a, 1983b) also treats the problem by using the notions of predicate, argument and condition.
Generally speaking, predicate (P) refers to verbs, argument (A) refers to subjects and objects, and finally, conditions (C) refers to adverbials. In his framework, accent goes to the focused element. This is true for the condition and argument. However, when a predicate is in the same focus domain with an argument, the predicate is not accented. The following are some examples. The focused elements are underlined.

(3.21) Our dóg's disappeared. (AP)

(3.22) Jáne's had an accident in Londón. (APC)

(3.23) Jóhn beats Máry. (APA)

(3.24) He béats her. (APA)

In the above four sentences, the only place where a predicate gets accented is when all the arguments are unfocused as in (3.24). For the other three sentences, as the predicate belongs to the same domain of focus as the argument, it is not accented.

The above four sentences show that, in Gussenhoven's model, it is possible to assign multiple accents in one focus domain. For example, in (3.22) both the argument and the condition are accented. The same is true for the two arguments in (3.23). Gussenhoven explains (1983a:17) that when this happens, the nucleus (the primary sentence stress) goes to the last accented word. For him, broad focus, which he calls news reading, involves focus on the whole sentence.
Thus, it is possible to assign more than one accent to one sentence. If the sentence does not have an object or an adverbial, the argument of the predicate (subject) will be accented. On the other hand, if a condition or another argument (object) are found after the predicate, then these elements will bear primary stress.

So far, we have presented a brief survey of the different models of focus representation. It can be seen that these models have somewhat different predictions regarding the relationship between focus and stress. The motivation for these different models seems to be related to different focus predictions with respect to a particular stress. The most discussed problem is how to account for broad focus. No study, so far as I know, has denied the existence of broad focus. Opinions differ only in how this focus pattern is represented by stress.

In the present study, we will also examine the type of sentence stress that realizes broad focus in Mandarin Chinese. In the literature, stress placement that represents unspecified focus interpretation has been only indirectly referred to. For example, Tao Lin (1962), in attempting to characterize neutral tones that are related to grammatical structures, also mentions the existence of the intonational neutral tone that has nothing to do with grammatical structures. This type of neutral tone can appear anywhere in
the sentence. These neutral tone syllables have inherent
tones and can be optionally stressed without affecting the
meaning of the sentence. The following is one example given
by Lin (304), in which the target syllable is shì 'be'.

(3.25) tā shì xuéshèng
       (he be student) 'He is a student'

   a. neutral tone              neutral reading
   b. stressed                 (I am right)
   c. strongly stressed        (Don't think that he is
                                 not a student)

According to Lin, the three different degrees of stress on
the word shì do not change the lexical meaning of the
sentence. They merely express the speaker's tone of speech.
However, Lin does not explain why a particular reading has
been associated with a specific degree of stress.

In the context of the present study, it is obvious that
the morpheme shì 'be' in (3.25c) is contrastively stressed.
The speaker is trying to distinguish it from bù shì 'is not'.
The difference between this reading and that of (3.25b) is
not qualitative. It merely expresses two different degrees
of stress. In both cases, the word shì 'be' is narrowly
focused with the tone of the word fully preserved. On the
other hand, the scope of focus of (3.25a) seems to extend
over the whole sentence. For example, it could well be the
case that the sentence is spoken in response to the question
"What did you say?".
Gårding's (1985, 1987) studies are also related to focus. Her data involve five focus types: broad focus statement, broad focus yes-no question and after focus, focus left, focus right statements.² What is interesting about her data is that the intonational contours for broad focus statement and focus right statement (that is the rightmost word is focused) are almost the same. This suggests that, in the two focus patterns, the same syllable is stressed. However, since she is mainly concerned with the constant features true for all speakers, the similarity and difference between the two patterns are not much discussed.

In the present study, we also find that broad focus is closely related to the stress condition for narrow focus on the last word of the test sentence. However, our results show that the acoustic effect of stress for broad focus is not exactly the same as the one that realizes narrow focus on the last word. We find that, for broad focus, the pitch range of the name syllable is also affected. Perceptually, the two stress conditions are also very close. This is shown by the fact that the two are the most easily confused. However, subjects are generally able to differentiate the two, suggesting that certain perceptual differences still exist between them.
In the the next section, we will present results of the perception experiment. Acoustic analysis of the target syllables will be given in the next two chapters.

3.2 The results of the perception experiment

In this section, we give an overall description of the subjects' score in the perception experiment. The basic pattern of focus for this study will then be established. The acoustic features related to these patterns will be discussed in the next two chapters. In Chapter Six, we will come back again to the perception experiment and see what correlation there is between acoustic parameters and the perception scores.

First, we will present a general description of the perception score. As has been described in Chapter One, the perception score is related to the subjects' perception of different focus domains. They first listened to the test sentence, as represented by the format below.

(3.26) X míngnián liáoyǎng.
    X next year go to sanitarium
    'X is going to the sanitarium next year.'

On the basis of the stress pattern, they were asked to choose the question that prompted the sentence when it was originally read. Four questions were listed for them to choose from. These questions are presented below in (3.27).
(3.27) a. Who is going to the sanitarium next year?
b. When is X going to the sanitarium next year?
c. What is X going to do next year?
d. What happened?

Of these four sentences, three show narrow focus. First, in (3.27a) the focus falls on the subject. Second, in (3.27b) it falls on the adverbial of time, and third, in (3.27c) the narrow focus is on the verb. In the following discussion, I will code these foci as focus 1, focus 2 and focus 3 respectively. Lastly, question (3.27d) elicits a broad focus statement. We will use focus 4 to code the related focus pattern (for a summary of the place of focus and the eliciting questions, see Table 1.1).

Table 3.1 below summarizes the general performance of the 16 subjects with respect to their responses to the 320 stimulus sentences that vary in focus. The numbers in the first four rows stand for the total number of responses from all subjects with respect to a stimulus sentence with the particular focus pattern shown at the top of the column. The four possible choices which represent the four focus patterns are listed at the left of the row. The correct responses in each case are highlighted.
Table 3.1 Subjects' responses in the perception experiment

<table>
<thead>
<tr>
<th>Description of responses</th>
<th>Focus 1</th>
<th>Focus 2</th>
<th>Focus 3</th>
<th>Focus 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>who</td>
<td>1091</td>
<td>53</td>
<td>50</td>
<td>216</td>
</tr>
<tr>
<td>when</td>
<td>40</td>
<td>1138</td>
<td>44</td>
<td>64</td>
</tr>
<tr>
<td>do what</td>
<td>42</td>
<td>19</td>
<td>718</td>
<td>375</td>
</tr>
<tr>
<td>what happened</td>
<td>107</td>
<td>70</td>
<td>468</td>
<td>625</td>
</tr>
<tr>
<td>Total number of responses</td>
<td>1280</td>
<td>1280</td>
<td>1280</td>
<td>1280</td>
</tr>
<tr>
<td>Total number of wrong responses</td>
<td>189</td>
<td>142</td>
<td>562</td>
<td>655</td>
</tr>
<tr>
<td>Percentage of correct responses</td>
<td>85.23%</td>
<td>88.91%</td>
<td>56.10%</td>
<td>48.83%</td>
</tr>
</tbody>
</table>

The above table shows that subjects did very well for focus 1 and 2. For focus 3 and 4, however, the percentage is rather low. Assuming that the percentage of the correct responses is directly related to the degree of difficulty in identifying the correct focus pattern, we can see that these two focus patterns are obviously more challenging than the other two.

The data also suggest that focus 3 and focus 4 are the most confusing categories. It can be seen that many focus 3 sentences have been misidentified as focus 4 ones and vice versa. In contrast, the number of wrong choices for the two other focus patterns is relatively low.
In terms of the distribution of scores, focus 4 has the most evenly spread responses, as well as the greatest number of wrong responses (655 out of 1280 choices). Comparing the three types of mistakes (that is the choices for "who", "when" and "do what" for focus 4), we can see that subjects had the least confusion with "when". This shows that focus 2 is the least confusable with focus 4. Similarly, a quick look at the subjects' performance for focus 2 will show that this is the easiest category to identify. On the other hand, the most mistakes subjects made in identifying a focus 4 pattern is in the "do what" category, which corresponds to focus 3. Conversely, "what happened" was the most common mistakes for focus 3 sentences.

After focus 3, focus 1 is the next category that tends to get confused with focus 4. When a focus 4 pattern stimulus is presented, "who" receives 216 points, which is much higher than the corresponding score of 64 for "when". Conversely, when a focus 1 pattern stimulus is presented, "what happened" receives 107 points, which is also higher than the corresponding points of 70 when a focus 2 pattern stimulus is presented.

Generally speaking, narrow focus patterns are less likely to be confused with each other. If confusion arises at all, it is between a narrow focus pattern and the broad focus. This is shown in our data by the fact that the
largest number of wrong choices in each of the first three columns always falls under "what happened". Also, focus 4 has the highest error rate in comparison with the other three focus patterns.

Basically, the results of the perception experiment show that three focus patterns can be easily identified, that is, stress 1, 2 and 3/4. The distinction between focus 3 and focus 4 is the most difficult to make. This conforms to the acoustic data presented in Gårding (1985). It suggests that broad sentence focus in Mandarin Chinese is realized by stressing the last content word of the sentence. For focus 3, the same word is stressed. This has made the distinction between the two focus patterns very difficult.

Given the high ratio of the wrong responses for focus 3 and focus 4, one might suspect that the answers subjects gave are largely based on guessing rather than on any possible acoustic differences between the stimuli. In order to evaluate this possibility, a 95% confidence interval for each of the choices is calculated. The z-values for testing the proportion of all the possible choices at the random level of 25% are also given in the parentheses. Both sets of data are summarized in the following table.
Table 3.2 95% confidence interval of the proportion of all choices and the corresponding z-values

<table>
<thead>
<tr>
<th>Focus</th>
<th>Focus 1</th>
<th>Focus 2</th>
<th>Focus 3</th>
<th>Focus 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who</td>
<td>85.2% ± 1.9% (49.76)</td>
<td>4.1% ± 1.1% (-17.35)</td>
<td>3.9% ± 1.1% (-17.43)</td>
<td>16.9% ± 2.1% (-6.71)</td>
</tr>
<tr>
<td>When</td>
<td>3.1% ± 1% (-18.07)</td>
<td>88.9% ± 1.7% (52.80)</td>
<td>3.4% ± 1% (-17.85)</td>
<td>5% ± 1.2% (-16.52)</td>
</tr>
<tr>
<td>Do what</td>
<td>3.3% ± 1% (-17.93)</td>
<td>1.5% ± 0.7% (-19.43)</td>
<td>56.1% ± 2.7% (25.69)</td>
<td>29.3% ± 2.5% (3.55)</td>
</tr>
<tr>
<td>What happened</td>
<td>8.4% ± 1.6% (-13.72)</td>
<td>5.5% ± 1.2% (-16.14)</td>
<td>36.6% ± 2.7% (9.55)</td>
<td>48.8% ± 2.7% (19.69)</td>
</tr>
</tbody>
</table>

At the 95% confidence level, a z-value in the above table that is either greater than 1.6449 or smaller than its negative value of -1.6449 means that the associated proportion is significantly different from the 25% chance proportion. The data suggest that all the proportions are not by chance since none of the intervals include 25% and, similarly, all the z-values fall outside the range of insignificance. However, it can be observed that, for focus 3 and focus 4, the z-values for the wrong choices (highlighted) are significant in the wrong direction. This again shows that the two focus patterns are two confusing categories. Some similar acoustic qualities obviously influenced the subjects in picking the wrong answers. Nevertheless, we cannot say that the two categories are exactly the same because the two confidence intervals for the score of "do what" and "what happened" for both focus 3 and
focus 4 do not overlap. This shows that subjects do not have the same confidence level with respect to the distinction of focus 3 and focus 4. The implication is that, to a certain extent, they are able to tell the difference between the two.

Based on these data, we can say that the focus patterns produced by the four subjects in the production experiment correspond to the four focus patterns perceived by the other sixteen subjects in the perception experiment. What is noteworthy is the subjects' different performance in judging the various focus patterns. Thus, a closer acoustic examination of the stimulus data is necessary. In the next two chapters, we will examine the three acoustic parameters, that is, F0, syllable duration and intensity, with respect to various focus related stress patterns. For ease of discussion, we will code the stress pattern that realizes focus 1 as stress 1, focus 2 as stress 2, focus 3 as stress 3 and focus 4 as stress 4. Thus, we will investigate four stress conditions, each of which represents a different focus pattern, including broad focus.
Other terms used in the literature, although in a somewhat different sense, include "new (information)" versus "old/given (information)" (Chafe 1970, 1976, Halliday 1967, and Selkirk 1984), "variable" versus "background" (Gussethoven 1983a, 1983b).

The term she uses is "focus free" instead of broad focus. However, they mean the same thing.

The formula used for calculating the 95% confidence interval for percentage is given below. Here $\hat{p}$ stands for the proportion of a certain number of choices for an answer over the total number of 1280 (see Woods et al. 1986).

$$\hat{p} \pm \left[ 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} + \frac{1}{2n} \right]$$

$Z$-values are derived by the following formula in which $\pi$ represents the hypothesized value of 25% (see Devore and Peck 1993):

$$z = \frac{\hat{p} - \pi}{\sqrt{\pi(1-\pi)/n}}$$
Chapter IV

The F0 parameter

In this chapter, we will investigate the acoustic parameter F0 and see how it interacts with stress and produces various pitch contours for sentences having different stress patterns. We will first present a general picture of the F0 behavior of stress in Section 4.1; then in Section 4.2, we will briefly analyze the measurement points of F0 with respect to the pitch range data of the three target syllables. The three target syllables are analyzed separately in Sections 4.3, 4.4 and 4.5. In Section 4.6, we compare stress 3 and stress 4 patterns. Finally, in Section 4.7, we summarize the findings in this chapter.

4.1 Gross pattern of F0 with respect to stress

In this section, we will examine the general behavior of F0 for different stress patterns. Basically, we will address the following three questions: 1. What happens to the F0 contour when a syllable is stressed? 2. How does tone influence the F0 excursion with respect to stress? 3. What
is the behavior of the F0 contour in pre- and post-stress positions? We will address these three questions in the following three subsections.

4.1.1 F0 and stress

The general pattern of F0 with respect to stress shows that when a syllable is stressed, its pitch range will be greatly expanded. This observation conforms to the results of previous studies (Liao 1994, Gårding 1985, J. Shen 1985, Shen 1990a, Tseng 1981, etc.). In the following figures, F0 excursions for Lào Wù mínghnián liáoýâng with respect to the four stress conditions for subject 1 and subject 3 are given.

![Figure 4.1 Stress 1 (T1) for subject 1 (C174)](image-url)
Figure 4.2 Stress 2 (T1) for subject 1 (C150)

Figure 4.3 Stress 3 (T1) for subject 1 (C106)
Figure 4.4  Stress 4 (T1) for subject 1 (C126)

Figure 4.5  Stress 1 (T1) for subject 3 (C374)
Figure 4.6 Stress 2 (T1) for subject 3 (C313)

Figure 4.7 Stress 3 (T1) for subject 3 (C368)
Some generalizations can be made from the patterns shown above. First, when a syllable is stressed, its pitch range is greatly expanded. It can be observed that, for all the stress patterns, the pitch range on the stressed word is obviously greater than the cases when the word is not stressed. Second, the F0 excursion for stress 3 strikes a strong resemblance to that of stress 4. Figures 4.3 and 4.4, for example, look similar with respect to the location of F0 peaks. For both figures, three peaks can be identified. And in both cases, the peak for the target syllable liáo is a bit lower than the previous syllable míng. Although it can be
observed that the peak value for stress 4 is higher than that for stress 3, the overall pattern for the two stresses are the same. Likewise, for subject 3, the two F0 excursions in Figures 4.7 and 4.8 are similar.

However, if we compare across speakers, the patterns of the two stresses for the two speakers are not identical. For subject 3, there is no obvious peak for míngnián in pre-stress position. The reason for this lack of tonal contrast is closely related to T2 sandhi discussed in Chapter Two. Both tones in the word míngnián 'next year' are T2 (/35/) preceded by a T1 syllable, Wù (/55/). Hence, in fast speech, the rising T2 is replaced by the level T1. Although subject 1 does not show this T2 sandhi in Figure 4.4, there is at least one token for the same tone and stress produced by the same speaker in which T2 sandhi is obvious. This is shown in Figure 4.9 below. The duration for this particular token is 873 ms, while the one for Figure 4.4 is 1087 ms. This contrast clearly shows that speed is an important factor in determining T2 sandhi. Another difference between the two tokens is that the pitch contour in Figure 4.4 has a rising tail as opposed to a falling contour of Figure 4.9. Again, this can be accounted for by the fact that the token represented in Figure 4.4 has a longer duration. In both cases, the last syllable is T3. When syllable duration becomes longer, T3 tends to surface as [214].
Another observation of the data is related to the position effect of stress. When the target syllable is at the beginning of a sentence, it has the highest peak. However, as the stress moves progressively back toward the end of the sentence, the target syllable's peak becomes obviously lower. In some cases, it is even lower than the F0 peaks of unstressed syllables before it. This is related to the phenomenon called "downstep", which is a lowering of the top of the pitch range. Similar phenomenon is also found in other studies of the same language (Liao 1994).
4.1.2 Tone and stress

In this section, we will discuss how the interaction of tone and stress is manifested through the variations of the F0 values. Since the name syllable is the only syllable controlled for tones, we mainly focus on this syllable.

Our data show that the basic patterns of F0 contour for the four tones can be divided into two categories, rising and falling, if the phonetic shapes of the tones are taken into consideration. T1 and T2 are rising, while T3 and T4 are falling. In many cases, T1 and T2 can be hardly distinguished by the F0 shapes alone. Two examples are given below in Figures 4.10 and 4.11.

![Figure 4.10 Stress 1 (T1) for subject 1 (C105)](image-url)
It was noted in Chapter Two that the tonal excursion for T1 is high-level and that for T2 is high-rising. The fact that these two tones have similar F0 contour in this study is related to the preceding T3 which is basically a low tone. From the low level of T3 to the high level of T1, the tonal contour undergoes a transition in which the beginning part of the high tone for the second syllable is not reached (i.e., a case of undershoot).

The phenomenon is called "tonal coarticulation", which Shih (1988) argues to exist in the interaction of tone and intonation in Mandarin Chinese. The same phenomenon is also discussed in Shen (1990c), and in Shen and Lin (1991). It is
basically the undershooting of a low or high target tone when a series of such tones run together. The motivation is ease of articulation.

In Chapter Two we also noted that T2 sandhi is a surface tone sandhi rule affected by speech tempo. In this sandhi process, the low target of the inherent T2 is not realized. The tonal excursion glides from the preceding high tone of T1 or T2 to the high target of the following T2. This results in a level tone similar to T1. Since T2 sandhi is optional and is controlled by speech tempo, we can regard it as a particular case of tonal coarticulation.

In our experiment, when the name syllable is T1, the toneme sequence of the whole name phrase is /21 55/ (or low-low high-high). In this context, it is very natural to undershoot the first high tone, resulting in a sequence of [21 35]. This makes the tonal contour of the name syllable of T1 and T2 look similar when it is combined with the preceding low tone (Figures 4.10 and 4.11).

This similarity between the two tones does not exist for some tokens, though. I have found that, in some cases, subjects try to maintain the high level tone of T1. When this happens, a glottal stop is inserted between the name syllable and the preceding low tone. This creates a noticeable break in the F0 excursion between T1 (Wū) and the preceding T3 (Lǎo). Figure 4.12 below is such an example.
As for the other two tones, T3 and T4, it is easier to tell the difference. Although both of them are falling, T3 actually falls lower than T4. Apart from this, another very obvious cue that can be used to tell the difference between the two tones is the pattern of F0 excursion of the following syllable. While a T2 (i.e. the syllable míng) following a stressed T3 may still maintain a considerable rising contour, the same tone following a stressed T4 hardly shows any rise at all. This can be illustrated by the contrast between Figure 4.13 in which the name syllable is the T3 syllable Lù and Figure 4.14 in which the name syllable is the T4 syllable Lù.
Figure 4.13 Stress 1 (T3) for subject 1 (C115)

Figure 4.14 Stress 1 (T4) for subject 1 (C136)
The sharp contrast between the two different behaviors of the syllable mǐng above strongly suggests that T3 is basically a low tone while T4 is simply a fall from a high tone. When T4 is stressed, it does not quite reach a very low target at its end. On the other hand, when T3 is stressed, its low target must be realized.

If we assume that the two post-stress syllables also have a top line, then the contrast between the two can be explained naturally. In the case of Figure 4.13, the tonal target for the stressed syllable Lǚ is lower than the top line of the following mǐng. In the case of Figure 4.14, the low target of the stressed syllable Lǜ is roughly at the same level of the top line of the following syllable. Thus, the difference between the two is basically a contrast of pitch ranges with the former being wider than the latter.

4.1.3 Unstressed syllables

In this section, we will examine the pitch range behavior of the unstressed syllables. According to the positions relative to a stressed syllable, they can be generally divided into pre- and post-stress categories. A rough examination of the data shows that the pitch range of the unstressed syllables in pre-stress positions is not so much affected by the following stressed syllable, especially when they are at the sentence-initial position. This can be
illustrated by Figure 4.15, where two peaks can be identified. The second major peak corresponds to the stressed syllable míng, while the first one is the pre-stress T4 syllable, Lù. It can be observed that the pitch range for this pre-stress syllable is still quite large. On the other hand, when the non-initial word míngnián occurs in the pre-stress position, the pitch contour appears much flatter. An example is given in Figure 4.16 below. The downward pointing arrows in both figures delimit the rising portion for the syllable míng.

![Graph showing pitch contour for different syllables](image)

Figure 4.15 Stress 2 (T4) for subject 3 (C303)
Figure 4.16 Stress 3 (T4) for subject 3 (C348)

In Figure 4.16 above, the rising contour of the T2 syllable mìng can be hardly seen. However, the falling portion of T4 for the name syllable is very conspicuous. If the name syllable is T1, as in the case of Figure 4.8, not much fall in F0 can be observed. As a matter of fact, if we compare Figures 4.8 and 4.16, we will find that the F0 excursion connecting the name syllable and the stressed syllable liáo is shifted slightly upward for the T1 name syllable. For ease of comparison, I have overlayed the two patterns in the following figure. The F0 range of Figure 4.8 has been readjusted to the same range of Figure 4.16.
In the above figure, the F0 excursion between a T1 ($\tilde{W\bar{u}}$) and the stressed syllable $li\acute{a}o$ is higher because $W\bar{u}$ ends with a high tone (T1). On the other hand, a T4 syllable ($L\bar{u}$) ends with a low tone. The F0 excursion connecting that syllable with the stressed one appears to be lower.

The fact that tonal coarticulation affects the syllable $m\tilde{\text{i}}ng$ more that it does the name syllable is related to the position of the two syllables. The name syllable belongs to the sentence-initial word, while $m\tilde{\text{i}}ngni\acute{a}n$, in which the syllable $m\tilde{\text{i}}ng$ is located, is flanked on both side by other words. This environment subjects the tonal contour of $m\tilde{\text{i}}ng$
to greater influence from the overall stress pattern (intonation) and from the tones of the neighboring syllables. For the name syllable, there is no preceding high tone that can undershoot the low target of T3 that immediately precedes the name syllable. As a result, the pitch range of the syllable is less affected.

In contrast to the pre-stress position, when a syllable occurs in a post-stress position, its pitch range is drastically reduced. However, our data show that syllables in this position may still retain their inherent tonal identity. An obvious example is Figure 4.13 discussed in the preceding section: the tonal contour for the syllable míng is completely preserved after a stressed T3 syllable.

In other cases, when a rising contour is not obvious in the post-stress position, we can still identify a flat plateau following the falling contour of the preceding stressed syllable, suggesting that there is still the high tone target at the end of the syllable míng even if the reduced pitch range, coupled with undershoot of the mid target at the syllable's onset, effectively eliminates the local difference between low and high. Figure 4.14 presented earlier is such an example.

When a syllable does not immediately follow a stressed word, its rising tonal contour can still be recognized in some cases. This can be illustrated by the following figure.
In this example, the boundaries for the syllable liáo are indicated by the two downward-pointing arrows. The rising contour for that syllable is rather obvious.

All these data show that the post-stress syllables can still retain their tonal identity even though their pitch range is reduced considerably. This seems to suggest that the post-stress neutral tone is caused by a reduction of the tonal ranges rather than by the loss of their inherent tones to the general intonation contour. In the rest of this chapter, we will present statistical evidence for the general F0 excursion patterns discussed in this section.
4.2 Measurement points

In this section, I will briefly describe the points of measurement for the F0 data with respect to the three target syllables. It can be recalled from Chapter One that the three target syllables are: the name syllable, the syllable míng in míngnián and the syllable liáo in liáoyāng. Since there is suggestion in the preceding sections that stress affects pitch range, we want key points that allow us to assess differences in pitch range around the target syllable - preferably the F0 minima and maxima that can be attributed to local low and high tone targets.

The first word in our data is "lǎo + name". Here, the target name syllable is controlled for tone. It can be observed that, of the four tones in Mandarin, T1 and T2 do not have a low-tone target. T1 is high-level, while T2 is high rising. However, since the preceding lǎo is T3, a low tone, we will assume that the pitch level realized by that tone represents the bottom level for our target name syllable. Consequently, we will use the lowest F0 point of lǎo as our reference point for the bottom line of the pitch range and the peak in the T1 or T2 of our name syllable as the reference point for the top of the pitch range. When the name syllable itself is T3, we will also use the preceding tone as the reference point. It was pointed out in Chapter Two that, when two T3 syllables run together, the first T3
becomes a T2. Thus, the pitch range can still be calculated since we have a sequence of a high tone followed by a low tone. Finally, T4 is high falling. The pitch range can be calculated from high and low targets in the syllable itself. The pitch range for the second and the third target syllables can be calculated in the same way. Since both syllables are rising, they have pitch range values if a local rise of F0 contour can be identified. Otherwise, their pitch range is zero.

The F0 contour of the three target syllables constitutes part of the intonational contour for the whole sentence, which includes both the tonal and the (sentence) stress information that we are looking for. Fortunately, our data are controlled for stress. Thus, by comparing different values of the same syllable with respect to different stress patterns, we can estimate the effect of stress on these syllables. For this purpose, it is important that we are consistent in measuring the same point of interest for different stress patterns.

When a syllable is stressed, the points of interest are obviously the peak and the valley (the lowest point) of the F0 excursion for that syllable. When the syllable is not stressed, the F0 values at the two extremum points become closer and closer together, and in some extreme cases, the rising and falling relationship is reversed as a result of
undershoot and pitch range reduction. In those cases, we still need to look for traces left over by the original tone. If an F0 peak and valley can still be identified, then those points are still used. However, when the tone is more drastically influenced by the stress pattern, its original contour can only be identified by a flat plateau in the falling contour. In those cases, the inflectional points, that is, the points at which the F0 trend changes direction, become crucial since they reflect inherent tonal contours to a certain extent. Only when the inherent tone of the syllable is totally obscured, that is, when it is completely overridden by intonation, do we rely on other criteria, such as syllable boundary, to choose the appropriate measurement points. In other words, syllable boundary should not be used when inflectional points are present. For example, in the following figure, the syllable boundaries are indicated by the two vertical markers. If we simply measure the F0 value at these markers, the information of the rising contour will be lost. This is of course counter-intuitive and contrary to the rise evident over the last half of the syllable.
Figure 4.19 Stress 4 (T1) for subject 3 (C330)

It is clear from the above figure that, if we want to include tonal information for consideration, we need to get the F0 values at the lowest point of the syllable and the peak of the F0 excursion to the right of that point since liáo is a T2 (rising) syllable.

In this study, six measurement points are selected, two for each target syllable. In the following discussion, I will describe those points one by one.

The first two points are related to the first target syllable, that is, the name syllable. Since this syllable varies in tonal specification across the corpus, specific measurement points depend on what tone is involved. For T1
and T2, point (a) is at the turning point between the low target of the tone of the preceding syllable and the transition to the high tone of the target syllable. Basically, it is the lowest point within the name phrase. Point (b) is then the peak following point (a). These points are shown below in Figures 4.20 for T1 and 4.21 for T2 respectively. For T3 and T4, point (a) is the peak at the beginning of the downward glide. In the case of T3, it is the high target at the end of the preceding T2 (sandhi tone for lǎo). Point (b) is then the lowest point between the name syllable and the syllable mǐng. These are illustrated in Figures 4.22 and 4.23 below.

Figure 4.20 Measurement points (a) and (b) for T1 (C105)
Figure 4.21 Measurement points (a) and (b) for T2 (C138)

Figure 4.22 Measurement points (a) and (b) for T3 (C165)
Figure 4.23 Measurement points (a) and (b) for T4 (C140)

The third point of measurement – point (c) for the target syllable míng – is also determined by the tone of the immediately preceding name syllable. When the name syllable is T3 or T4, point (c) coincides with point (b). It is the lowest F0 point between the two syllables (see Figures 4.22 and 4.23 above). When the preceding name syllable is either T1 or T2, the situation is more complicated. The inherent tonal contour for T1 and T2 can be numerically represented as /55/ and /35/ (see Table 2.1). Since míng is also a T2 syllable, we have a toneme sequence of /55 35/ or /35 35/. In this situation, it is often the case that there is a tonal
transition between the two syllables from 5 to 3. This can be seen clearly in Figure 4.20 where the rising tone of míng is preceded by a falling transitional glide. In those cases, the lowest point of the F0 excursion between the name syllable and the syllable míng will be taken as the measurement point. In other cases where míng is not rising, if an inflectional point toward the beginning part of the syllable can be detected, then that point will be taken as (c). Otherwise, if the tonal contour is falling without inflection, point (c) coincides with the syllable boundary.

The fourth point of measurement – point (d) – is the peak point of the F0 excursion for the syllable míng if the syllable exhibits a rising contour (see Figures 4.22 and 4.23). In other cases where the rising portion of the syllable is not obvious due to a particular stress pattern, I will again use the inflectional points as a marker for point (d). If there is no inflection, the syllable boundary will then be taken as the point. It should be noted that inflectional points, like F0 peaks and valleys, do not necessarily coincide with syllable boundaries. An obvious case is the name syllable. In most cases (T1, T2 and T4), point (a) actually belongs to the preceding syllable. For the syllable míng, the peak point sometimes is associated with the following syllable nián. In those cases, that point is still taken as (d).
The principle involved in setting up measurement points — (e) and (f) — for the syllable liáo is basically the same as the one for the syllable míng. If the syllable has a rising contour, then the lowest point and the peak will be used for those two points. Otherwise, either inflectional points or syllable boundary will be used as the marker for the points. However, inflectional points have higher precedence.

4.3 The name syllable

In this section, we will examine the pitch range data for the name syllable. I will begin by defining the pitch range of the name syllable in my data as the distance in hertz between measurement points (a) and (b). The value is derived by taking the absolute value of the difference between the two points. Note that point (a) for T1 and T2 is located at the valley of the F0 excursion, while point (b) is at the peak. On the other hand, for T3 and T4, the relation between these two points is exactly the opposite. So, if we just simply take the difference between the two points as the pitch range data, we will end up getting negative values.

We first begin our analysis by calculating the pitch range value for every token. Then, we use this value as the dependent variable in the ANOVA model in which the three independent variables, subject, tone, and stress, are tested
for main effects and various interactions. The result of the ANOVA is presented in the following table.

Table 4.1 Analysis of variance for the pitch range of the name syllables

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>102973.3</td>
<td>34324.4</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>18845.6</td>
<td>6281.9</td>
<td>2.37</td>
<td>0.139</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>75323.6</td>
<td>25107.9</td>
<td>19.89</td>
<td>0.000</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>23504.6</td>
<td>2611.6</td>
<td>8.48</td>
<td>0.000</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>11363.3</td>
<td>1262.5</td>
<td>4.03</td>
<td>0.002</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>10921.2</td>
<td>1213.5</td>
<td>3.87</td>
<td>0.003</td>
</tr>
<tr>
<td>subject<em>tone</em>stress</td>
<td>27</td>
<td>8458.8</td>
<td>313.3</td>
<td>3.06</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>26190.1</td>
<td>102.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>277980.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated

The above table shows that the pitch range of the name syllables is related to stress (F [3, 256] = 19.89, p = 0.000). The factor "stress" is involved in three of the four interactions (denoted by the star "**"), all found to be statistically significant.

To further evaluate how stress affects the pitch range of the target syllable, the mean values for the name syllables under the different stress patterns are calculated and presented below in the following table. The number in parentheses designates standard deviation.
Table 4.2 Mean value (Hz) of pitch ranges for the name syllable with respect to stress \((N = 80)\)

<table>
<thead>
<tr>
<th>Stress 1</th>
<th>Stress 2</th>
<th>Stress 3</th>
<th>Stress 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>86.23 (33.54)</td>
<td>51.35 (20.68)</td>
<td>46.81 (20.07)</td>
<td>57.57 (24.69)</td>
</tr>
</tbody>
</table>

The table shows that when the name syllable is stressed (stress 1) its mean pitch range is greater than when it is not stressed. In order to show that the value listed under stress 1 is greater than those under the other stresses (that is when another syllable in the sentence is being stressed), multiple post-hoc Tukey comparisons are performed. The results show that the pitch range values under stress 1 are significantly greater than those under the other stresses. The 95% confidence intervals of the difference between the mean target values under stress 1 and those under the other stresses do not include zero (34.88±10.28 Hz for the difference of stress 1 minus stress 2, 39.42±10.28 Hz for the difference of stress 1 minus stress 3, and 28.66±10.28 Hz for the difference of stress 1 minus stress 4). In other words, the mean values under the two stresses being compared are sufficiently separated such that the minimum value under stress 1 is still greater than the maximum values under the other stresses. On the other hand, when stress 2 is compared with stresses 3 and 4, the differences between the compared
means are very small. In both cases, their confidence intervals include zero (4.54±10.28 Hz for stress 2 and stress 3 and -6.22±10.28 Hz for stress 2 and stress 4), suggesting that those means are not significantly different. However, when stress 3 and stress 4 are compared, we find that their means are marginally different (-10.76±10.28 Hz). This indicates that the pitch range of the name syllable under stress 3 is slightly smaller than that under stress 4.

Now we analyze the stress effect for each subject. Figure 4.24 below summarizes the mean values of pitch ranges for the name syllables with respect to subject, tone, and stress (see Appendix A for actual values). It can be observed that subject 2 has the smallest pitch ranges overall. However, the general tendency that the greatest pitch range occurs at stress 1 is still true for all subjects and tones. Another apparent tendency in the data is that the pitch ranges under stress 4, in most cases, appear to be higher than those under stresses 2 and 3. For subject 1 and T2, that value is almost the same as the corresponding value under stress 1.

In order to show that the values under stress 1 is significantly greater than those under the other stress conditions, again we perform the multiple post-hoc Tukey comparisons for each subject with respect to stress. The results are summarized in Table 4.3 below.
Figure 4.24  Mean values of pitch ranges of the name syllables with respect to subject, tone, and stress
Table 4.3  Pairwise Tukey comparisons of pitch range values of the name syllables

<table>
<thead>
<tr>
<th>Stress Diff.</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1 - s2</td>
<td>27.84, 60.25</td>
<td>12.51, 24.85</td>
<td>29.31, 58.66</td>
<td>15.83, 49.77</td>
</tr>
<tr>
<td>s1 - s3</td>
<td>35.33, 67.75</td>
<td>11.40, 23.73</td>
<td>37.56, 66.92</td>
<td>19.35, 53.29</td>
</tr>
<tr>
<td>s1 - s4</td>
<td>16.95, 49.43</td>
<td>5.59, 17.92</td>
<td>32.98, 62.34</td>
<td>5.08, 39.02</td>
</tr>
</tbody>
</table>

The above table shows the 95% confidence intervals of the difference between the mean values of the two stresses being compared. It suggests that the values under stress 1 are significantly greater than those under the other stresses for all subjects. This is shown by the fact that none of the confidence intervals for the comparisons include zero. This confirms our initial observation that the pitch range of the name syllable is significantly expanded when the syllable itself is stressed.

Figure 4.24 shows that, when the name syllable is not under stress 1, it has the greatest pitch range values under stress 4. This is also indicated in Table 4.3 above. It can be observed that, when stress 1 is compared with another stress, the smallest value among the lower bounds is usually found between stress 1 and stress 4. This is especially true for subject 4. For her, the value is only 5.08 hertz. This implies that the mean target values under stress 1 can be as
little as 5.08 hertz greater than the corresponding mean values under stress 4. Indeed, for this subject, some tokens do not show any difference at all between the two values. The following is an example in which F0 traces of the two stress patterns are overlaid in the same figure. The pitch ranges of the name syllable (T2) under stress 1 and stress 4 do not show much difference.

![Graph of F0 contours for stress 1 and stress 4](image)

Figure 4.25 Overlaid F0 contour of stress 1 versus stress 4 for subject 4 (C408 and C437)

Although the above figure shows that, for stress 4, the peak of the name syllable is higher than the corresponding peak under stress 1, it can be seen that the rise for stress
1 starts at a slightly lower level. If we take this into consideration, the pitch range of the name syllable for the two stress conditions are not much different.

In contrast, subject 3 shows a different tendency. For her, the pitch range for the name syllable is obviously greater under stress 1. The following is an example. The tone of the name syllable is also T2.

![Figure 4.26 Overlaid F0 contour of stress 1 versus stress 4 for subject 3 (C308 and C337)](image)

Two observations can be made by comparing the two figures above. First, in Figure 4.26 the pitch range for the name syllable under stress 4 is quite different from that
under stress 1. For subject 3, the pitch range is obviously smaller under stress 4 than under stress 1. The other difference between the two subjects is the dipping before the syllable liáo. It can be seen that this dipping for subject 4 is quite deep. However, since both contours represent the same stress 4 pattern, obviously the difference is not important. What is significant is the different F0 excursion patterns between the two stresses. In this respect, the two subjects have the same behavior. For stress 1, in both cases, the peak of the stressed name syllable stand out conspicuously against the neighboring syllables. However, for stress 4, the same name syllable does not form a sharp contrast against the other syllables. This suggests that the pitch range expansion of the name syllable for stress 4 is not qualitative, and that, the F0 behavior for stress 1 is not simply an expansion of the pitch range for the target syllable. The reduction of the pitch range after that syllable is also important. Our data show that this reduction is basically the dropping of the top line. Thus, we may say that both pitch range expansion and the dropping of the top line of the following syllables that forms a contrast for the stressed syllable are important for identifying this stress pattern. Note that the stress 4 pattern of subject 4 also exhibits a range expansion for the name syllable. However, the expansion is not accompanied by
the subsequent dropping of the top line of the following syllables. Thus, it does not make a contrast between the name and the following materials, and so, the syllable is not stressed in the same way as in stress 1.

To summarize, when the name syllable is stressed, its pitch range will expand. The general tendency is that this target syllable shows the greatest value for the pitch range expansion when it is stressed. However, there is also a noticeable tendency for some subjects, especially subject 4, to expand their pitch range of the target syllable under stress 4. The difference between the two stress conditions is that, for stress 1, the pitch range expansion is immediately followed by a general fall of F0 excursion. However, for stress 4, the overall F0 excursion is rather flat. So the pitch range expansion of the name syllable does not stand out in sharp contrast against the following syllables.

4.4 The syllable míng

In this section, we present the analysis of stress effects on the pitch range of the syllable míng. First, we define the pitch range of this target syllable as the difference between F0 values at points (c) and (d). Since míng is T2, a rising tone, the F0 value at point (c) should be smaller than that at point (d). Thus, when we subtract
the former from the latter, a positive value indicates a rising contour. This value will be taken as the pitch range. On the other hand, a negative value signifies a falling F0 excursion. When a rising tone becomes falling, it does not have any pitch range at all. Its F0 excursion is basically determined by the sentence intonation. Thus, when the pitch range values for every token are calculated from points (c) and (d), all negative values are changed to zero. The following table summarizes the analysis of variance for the pitch range data. The result shows that the main effect of "stress" is significant (F [3, 256] = 21.90, p = 0.000).

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>18780.9</td>
<td>6260.3</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>40357.8</td>
<td>13452.6</td>
<td>14.00</td>
<td>0.001</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>68506.8</td>
<td>22835.6</td>
<td>21.90</td>
<td>0.000</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>8647.3</td>
<td>960.8</td>
<td>4.69</td>
<td>0.001</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>9384.5</td>
<td>1042.7</td>
<td>5.09</td>
<td>0.000</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>7486.5</td>
<td>831.8</td>
<td>4.06</td>
<td>0.002</td>
</tr>
<tr>
<td>subject<em>tone</em>stress</td>
<td>27</td>
<td>5525.9</td>
<td>204.7</td>
<td>3.25</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>16117.7</td>
<td>63.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>174807.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated
The mean pitch range values are summarized in the following table. It can be observed that, when the syllable míng itself is stressed (stress 2), it has the largest pitch range value. When the preceding name syllable is stressed (stress 1), on the other hand, the value is drastically reduced.

<table>
<thead>
<tr>
<th>Stress 1</th>
<th>Stress 2</th>
<th>Stress 3</th>
<th>Stress 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.56 (19.05)</td>
<td>49.14 (19.25)</td>
<td>18.04 (16.00)</td>
<td>18.15 (18.88)</td>
</tr>
</tbody>
</table>

Table 4.5 Mean value (Hz) of pitch ranges for the syllable míng with respect to stress (N = 80)

In order to further evaluate the stress effect on the syllable míng, multiple post-hoc Tukey comparisons are performed. The results show that the mean value of the syllable míng under stress 2 is significantly greater than those under stress 1, stress 3 and stress 4. None of the comparisons include zero in their 95% confidence intervals (37.58±7.44 Hz for the difference of stress 2 minus stress 1, 31.1±7.44 Hz for the difference of stress 2 minus stress 3, and 30.99±7.44 Hz for the difference of stress 2 minus stress 4). On the other hand, when the syllable is not stressed,
the mean pitch range values do not differ under various stress conditions.

Now, we consider the interaction of subject and stress against the tonal background of the name syllable. The analysis of variance indicates that there exists significant interactions among the three independent variables (F [27, 256] = 3.25, p = 0.000). The Figure 4.27 below summarizes the mean pitch range values over five tokens for the syllable míng with respect to subject, stress, and tonal contexts (actual values are presented in Appendix B).

Basically, the figure shows that the syllable míng has the greatest pitch range when it is stressed. This is a typical pattern for T1, T2 and T4 contexts. However, for T3 context, the increase of the target value is not so dramatic. An obvious reason for this discrepancy is the low tone of the preceding syllable. We observed in Section 4.1 that, when the stressed name syllable is T3, there will be a very conspicuous rise in the F0 excursion for the following syllable míng.

It can also be observed from the figure that, in tonal context T3, the pitch range values under stress 1 are well above zero for all subjects. This means that when the preceding name syllable is stressed, the F0 excursion of the following míng is still rising as has already been observed above.
Figure 4.27 Mean values of pitch ranges of the syllable "ming" with respect to subject, stress, and tonal context.
In our experiment, subject 2 generally has the smallest pitch range overall. Even for him, conspicuous rising F0 excursions for the target syllable mǐng can still be detected when the syllable is in T3 context under stress 1. An example is given below in Figure 4.28. Points (c) and (d) are also given. It is apparent that the F0 value at point (d) is considerably greater than the corresponding value at point (c). This shows that in tonal context T3 the post-stress syllable mǐng still has a pitch range.

---

Figure 4.28 F0 contour of stress 1 (T3) from subject 2 (C253)
The mean pitch range values of the syllable míng shown in Figure 4.27 above are reflections of the difference between F0 values at points (c) and (d). However, so far we do not know if this is the result of an increase of the F0 value at point (d) relative to other tonal contexts or of a decrease of the F0 value at point (c). To further analyze the data, we need to compare the absolute F0 values in T3 context with those in other tonal contexts under stress 1. In the following figure, we first present the mean F0 values at point (c) under stress 1 with respect to the four tonal contexts.

![Graph showing F0 values for different subjects and tonal contexts](image)

Figure 4.29 Mean F0 values (N = 5) at point (c) under stress 1
In the above figure, two different patterns, subjects 1 and 3 on the one hand, and subjects 2 and 4 on the other, can be identified. For the first two subjects, the target points in the context of T3 are much lower than those in the context of T4, whereas for the other two subjects they are not. However, all four subjects show much lower target values in tonal context T3 as compared with either T1 or T2. (see Appendix C for the mean values and the one-tailed t-test results).

Now we present the mean values at point (d) under stress 1 with respect to the tonal contexts in the following figure.

Figure 4.30 Mean F0 values (N = 5) at point (d) under stress 1
In this figure, we can see that the values in the context of T3 are in all cases higher than those of T4. One-tailed t-tests yield significant values for every subject (see Appendix D). On the other hand, only about half of the t-tests turn out to be significant when T3 is compared with T1 and T2 contexts.

To summarize the observations in the above two figures, we can say that point (c) is generally lower in T3 context relative to T1 and T2 and point (d) in the same context is generally higher relative to T4. Overall values at point (d) are relatively more stable than those at point (c).

Similar observations have been made in Shih (1988) and Xu (1995). They both observe that a post-stress T2 syllable starts at a much lower F0 point when the preceding stressed syllable is T3. Shih (1988:94) explains the phenomenon as being related to the tonal environment. When the stressed tone is H or L, the following tone that has the opposite value (L vs. H or H vs. L) will bring the pitch level back to normal immediately. A T3 is a low tone. When it is stressed, the following T2 (a rising tone) starts out lower because its tonal value at the beginning is the same as the preceding syllable. Xu (1995) offers two other possible explanations for the same phenomenon. The first is that it is a process of dissimilation. This is actually an exception to his general observation that dissimilation only affects a
preceding tone in the context of a following stressed syllable. His second explanation is that it is related to the final rise of T3 in citation form. However, it is not clear how such a relationship could hold, since it has been generally observed (Liao 1994, Tseng 1981, etc.) that a T3 is realized as nothing but a low tone target in non-final position, and even in final position, in many cases, the rising portion may still not materialize.

The explanation that we are going to offer is related to the reduction of the top lines after a stressed syllable. It was noted in the preceding section that a stressed name syllable is characterized by the pitch range expansion of the stressed syllable and the dropping of the top line of the following syllables. Given our definition of the pitch range for the syllable míng, when the name syllable is stressed, point (d) represents the top line value after focus, while point (c) is the value at the bottom line if it is lower than point (d).

The mean F0 values at points (c) and (d) in Figures 4.29 and 4.30 show that the greater pitch range values for the post-stress syllable míng in the T3 context are largely due to the lowering of the F0 values at point (c). This suggests that when a T3 name syllable is stressed, its F0 contour dips lower since point (c) in this particular tonal context is at the same location as point (b). It is apparent that we can
attribute the low F0 starting point for the post-stress T2 syllable to the preceding stressed low tone.

In the same stress and tonal context, our data show that the syllable míng ends on a higher F0 level. This is especially true when tonal context T3 is compared with T4. The phenomenon can be easily explained in terms of different tonal targets of the preceding stressed syllable. When T4, a high-falling tone, is stressed, the top line of the following syllable is reduced more than if T3, a low-dipping tone, is stressed. Apparently for these two different tonal targets, the post-stress top line of the syllable míng is not scaled to the same height. An obvious explanation is that a stressed high tone reduces the top line of the following syllable more than a stressed low tone.

The fact that the post-stress syllable míng has different F0 excursions is also related to the beginning point of the syllable. Since this point is the same as the end point of the preceding name syllable, the contour shape of the whole syllable is obviously affected by the F0 level at which the preceding name syllable ends. When the name syllable is T3, the post-stress syllable míng will start on a much lower F0 level. Since this syllable is inherently rising and there is a top line above the starting point, it is natural for F0 excursion to go up. By contrast, in a T4 context, the starting point of the syllable míng is found on
a higher level. In addition, the top line is relatively low because of the preceding stressed high tone. The resulting F0 excursion depends on the difference between the starting point and the top line. In most cases, since the difference is not much, a flat plateau is a typical pattern to be found.

The F0 excursions of the syllable míng in T1 and T2 contexts are very different. T1 is high-leveling, while T2 is high-rising. Both tones end on a higher F0 level than either T3 or T4. When they are stressed, the top line of the following syllable is effectively reduced. However, the starting point of the following syllable míng is not on a lower level. In fact, it is even higher than the reduced top line. Thus, the post-stress syllable míng in the T1 and T2 contexts starts on a higher level and will not glide upward from that point on. Typically, it does not reach a lower F0 level where it is supposed to start and glides directly toward the reduced top line (due to tonal coarticulation), resulting in a falling excursion. In some cases, however, when it does reach a level lower than the reduced top line, a rising contour can be observed. For example, in Figures 4.10 and 4.11, the post-stress syllable míng still exhibits a clear rising contour.

In the pre-stress position, the F0 excursion of the syllable míng is influenced by the tone of the preceding name syllable in the same way except that the top line is not
reduced. When the preceding name syllable is T1 or T2, the following syllable míng starts on a higher level. Thus, the resulting F0 excursion for the syllable is rather flat. However, when the preceding syllable is T3 or T4, the syllable míng will start on a slightly lower level because of the lower end point of the preceding name syllable.

This difference is related to the T2 sandhi discussed in Chapter Two (see Section 2.1). This sandhi process changes a rising T2 to a leveling T1 when the target T2 follows another T2 or a T1 in a non-final position. However, when the preceding tone is T3 or T4, sandhi does not take place. T1 or T2 ends in a high tone. When they precede a T2 syllable, the starting point of that syllable will be on a high F0 level. T3 or T4, on the other hand, has a low-tone ending. When they occur before the target T2, the F0 excursion of that syllable will start on a lower level. Thus, the target tone will not become a T1 (55). For example, in Figure 4.17, the tone of the name syllable determines the shape of the following syllable míng. In the T4 context (Lù), the syllable starts on a lower level. This results in a rising F0 excursion for the syllable míng. By contrast, in the T1 context (Wū), the target syllable starts off on a higher F0 level, thus creating a level F0 excursion for the syllable míng.
The occurrence of the T2 sandhi presupposes two other conditions. First, neither the preceding nor the target tone must be stressed. Second, the speech tempo must be fast. The first condition is crucial since when the preceding syllable is stressed, the top line of the target syllable will be reduced to a lower level. The resulting tone cannot be T1 (55). If the target tone itself is stressed, because of pitch range expansion, its inherent rising tone will surface. The second condition is related to the ease of articulation since it is more likely for tonal coarticulation to occur in fast and casual speech than in slow and deliberate speech. When coarticulation does occur, the target low tone is not reached. The tonal contour of the syllable glides directly from the beginning high point toward the top line of the syllable. This creates a level sandhi tone.

To summarize the discussion in this section, the F0 parameter is generally related to stress for the syllable míng. Like the name syllable, when míng is stressed, its pitch range will expand greatly. When the syllable is not stressed, the direction of the F0 excursion is largely determined by the tones of the name syllable. We have found that, in the post-stress position, the F0 excursion for the syllable míng is typically falling in the T1 and T2 contexts, flat in the T4 context and rising in the T3 context. We have
attributed the difference to the top line reduction of the syllable mǐng after the stressed name syllable and the F0 height at the end point of that syllable. Our data show that a stressed high tone reduces the top line of the following syllable more than a stressed low tone. In a T3 context, the top line of the post-stress mǐng is less affected. On the other hand, in other tonal contexts, the top line of the syllable is somewhat lower. The rising post-stress contour is accounted for by the lower starting point and the higher top line of the syllable. The flat contour of the target syllable after a stressed T4 is accounted for by a low starting point and a much reduced top line. Lastly, when the target syllable starts off on a higher F0 level with a much reduced top line, the result is a falling contour. This is a typical situation when the syllable follows a stressed T1 or T2 name syllable. In the pre-focus position, the tonal contexts of the T2 sandhi can also be explained by the tone at the end of the preceding syllable and the top line of the target syllable. The sandhi does not apply if the preceding syllable ends with a low tone. Because in this particular tonal context, the following T2 syllable has to start on a lower F0 level, thus effectively preserving its inherent rising quality. Similarly, the sandhi will not take place if a stress has occurred since this will change the top line of the target syllable.
4.5 The syllable liáo

In this section, we will examine the stress effect on the pitch range of the syllable liáo. Again, we define the pitch range of this syllable as the difference in hertz between points (e) and (f). Like the syllable míng, this syllable is also T2. Therefore, when the F0 value at point (e) is subtracted from the corresponding value at point (f), a positive value indicates a rising contour and the presence of a pitch range. If the result is negative, we will reset the value to zero since a falling intonation in this context indicates no pitch range for this syllable. The following table summarizes the result of the analysis of variance for the data.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>4716.60</td>
<td>1572.20</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>68.05</td>
<td>22.68</td>
<td>0.23</td>
<td>0.871</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>21678.59</td>
<td>7226.20</td>
<td>10.05</td>
<td>0.003</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>874.65</td>
<td>97.18</td>
<td>1.76</td>
<td>0.122</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>6469.18</td>
<td>718.80</td>
<td>13.05</td>
<td>0.000</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>402.69</td>
<td>44.74</td>
<td>0.81</td>
<td>0.609</td>
</tr>
<tr>
<td>subject<em>tone</em>stress</td>
<td>27</td>
<td>1487.16</td>
<td>55.08</td>
<td>1.60</td>
<td>0.034</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>8792.74</td>
<td>34.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>44489.64</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated
The result shows a significant stress effect (F [3, 256] = 10.05, p = 0.003) on the pitch range values of our target syllable. To further analyze our data, we summarize the mean values with respect to stress in the following table.

Table 4.7 Mean values of pitch ranges for the syllable liáo under the four stress conditions

<table>
<thead>
<tr>
<th>Stress 1</th>
<th>Stress 2</th>
<th>Stress 3</th>
<th>Stress 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.02 (5.98)</td>
<td>2.82 (4.68)</td>
<td>23.08 (10.85)</td>
<td>15.93 (10.91)</td>
</tr>
</tbody>
</table>

It was noted in Chapter Three that narrow focus on the word liáoyāng is closely related to broad focus on the whole sentence. This is exactly what is shown in the above table. It can be seen that the mean pitch range values at stresses 1 and 2 on the one hand and stresses 3 and 4 on the other form two separate groups. This implies that when the name syllable (stress 1) or the syllable míng (stress 2) is stressed, the pitch range values for the syllable liáo are similar in that they are both relatively low as compared with the corresponding values when the syllable itself is stressed.

In order to further evaluate the data, we again perform the multiple post-hoc Tukey comparisons. The results show
that the mean values under stress 3 are significantly greater than those under the other three stress conditions. None of the 95% confidence intervals for the difference between the paired means include zero (18.06±3.45 Hz for stress 3 minus stress 1, 20.26±3.45 Hz for stress 3 minus stress 2 and 7.15±3.45 Hz for stress 3 minus stress 4). Thus, our data show statistically that, when the syllable liáo is stressed, it has a greater pitch range.

Now we examine the stress effect on the pitch ranges of the syllable liáo with respect to each subject. The mean values are summarized in Figure 4.31 below (see Appendix E for actual values). It can be observed from the figure that, when the syllable liáo is under stress 3, it generally has higher mean values. This is especially true when the values under stress 3 are compared with those under stresses 1 and 2. Although the mean values under stress 4 are generally lower than those under stress 3, subject 4 shows a different pattern. For her, the values under the two stress conditions do not seem to differ very much. Only in one tonal context is the mean value under stress 4 clearly smaller than the corresponding value under stress 3. In all other cases, the values under the two stress conditions are almost the same. This observation is further confirmed by the results of the multiple post-hoc Tukey comparisons summarized in Table 4.8 below.
Figure 4.31 Mean values of pitch ranges of the syllable liáo (n = 5) with respect to subject, tonal context and stress.
Table 4.8 Pairwise Tukey comparisons of pitch range values of the syllable liáo

<table>
<thead>
<tr>
<th>Stress Diff.</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>s3 - s1</td>
<td>24.81, 35.13</td>
<td>0.22, 7.58</td>
<td>14.87, 22.58</td>
<td>12.52, 26.77</td>
</tr>
<tr>
<td>s3 - s2</td>
<td>24.73, 35.05</td>
<td>2.03, 9.39</td>
<td>18.45, 26.16</td>
<td>16.02, 30.28</td>
</tr>
<tr>
<td>s3 - s4</td>
<td>8.57, 18.88</td>
<td>2.62, 9.98</td>
<td>3.55, 11.26</td>
<td><strong>-5.98, 8.28</strong></td>
</tr>
</tbody>
</table>

The above table shows that only one confidence interval includes zero. The general pattern is that the syllable has a greater pitch range under stress 3. For subject 4, our data show that she tends to place the same degree of stress on the syllable liáo under stress 3 and stress 4 since the pairwise comparison of the mean values under the two stress conditions does not produce significant result. After further comparisons, we find that, for this subject, the target mean values under stress 4 are still significantly greater than those under stress 1 and stress 2 (18.49±7.13 Hz for the difference of stress 4 minus stress 1 and 22±7.13 Hz for the difference of stress 4 minus stress 2).

To summarize, when the syllable liáo is under stress 3, its pitch range will experience a significant expansion. For most subjects, this expansion is significantly greater than the corresponding expansion of the same syllable under stress 1, stress 2 and stress 3.
4.6 A comparison of stress 3 and stress 4

We have shown in the previous section that, while syllable 例 exhibits higher pitch range values in most cases under stress 3 as opposed to stress 4, for subject 4, this is not the case. In this section, we will re-examine the name syllable and the syllable 明 in order to compare their values under these two stress patterns. In the following table, the mean values of the pitch ranges of the name syllable with respect to subject and the two stress conditions are presented with the two-tailed t-test results.

<table>
<thead>
<tr>
<th></th>
<th>stress 3</th>
<th>stress 4</th>
<th>2-tailed t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>subj 1</td>
<td>44.5 (20.00)</td>
<td>62.8 (26.6)</td>
<td>-2.47/0.019/35*</td>
</tr>
<tr>
<td>subj 2</td>
<td>25.11 (7.20)</td>
<td>30.92 (6.47)</td>
<td>-2.63/0.011/37</td>
</tr>
<tr>
<td>subj 3</td>
<td>53.2 (14.00)</td>
<td>57.8 (18.4)</td>
<td>-0.88/0.38/35</td>
</tr>
<tr>
<td>subj 4</td>
<td>64.5 (12.3)</td>
<td>78.7 (14.1)</td>
<td>-3.40/0.0016/37</td>
</tr>
</tbody>
</table>

*t-value/p-value/degree of freedom

The above table shows that three subjects out of four have significant greater values under stress 4. Thus, it appears that the pitch range of the name syllable undergoes
greater expansion under broad focus than it does when the syllable liáo is narrowly focused.

This result is not surprising. Note that, under broad focus, background information about the sentence is not presupposed on the part of the speaker. So, it is natural to explain this larger pitch range expansion as a result of heavier semantic load for the name syllable.

In the following scatter plot, average values of five tokens are plotted for stress 4 again stress 3. It can be observed that most of the data points fall on the side of the stress 4. Although in a few cases there is no difference in the target values under the two stress patterns, the general tendency is that higher pitch ranges correlate with stress 4.

Figure 4.32 Scatter plot of the mean pitch range values of the name syllable (N = 5) with respect to stresses 4 and 3
Now, we examine the syllable míng. The mean pitch range values are first compared under the two stress patterns and the results are presented in the following table.

Table 4.10 Comparisons of the mean pitch range values for the syllable míng (N = 20) under stresses 3 and 4

<table>
<thead>
<tr>
<th></th>
<th>stress 3</th>
<th>stress 4</th>
<th>2-tailed t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>subj 1</td>
<td>24.5 (12.8)</td>
<td>20.5 (19.2)</td>
<td>0.75/0.45/33*</td>
</tr>
<tr>
<td>subj 2</td>
<td>9.66 (5.55)</td>
<td>10.80 (6.68)</td>
<td>-0.59/0.56/36</td>
</tr>
<tr>
<td>subj 3</td>
<td>14.0 (17.7)</td>
<td>13.4 (20.1)</td>
<td>0.10/0.92/37</td>
</tr>
<tr>
<td>subj 4</td>
<td>24.0 (19.6)</td>
<td>27.8 (21.9)</td>
<td>-0.58/0.56/37</td>
</tr>
</tbody>
</table>

*t-value/p-value/degree of freedom

It can be observed in the above table that the mean values of the pitch ranges under the two stress conditions are not significantly different. In Figure 4.33 below, we present a scatter plot of the data. The plot shows that the mean target values are basically distributed rather evenly on both sides of the equal value line. The four data points that are separated from the rest of the data are largely due to T3 and T4 tonal contexts. In Figure 4.27, it can be observed that the pitch range values under stress 3 and stress 4 are typically separated from each other. T3 and T4
tend to have larger values. However, because of large standard deviations, they are not shown to be significantly different by the two-tailed t-test. Generally speaking, we can say that the pitch range value of the syllable míng is not affected by the stress patterns 3 and 4. The variations in the data are largely due to the different tonal contexts.

![Graph showing scatter plot of mean pitch range values for the syllable míng (N = 5) with respect to stresses 4 and 3.]

Figure 4.33 Scatter plot of the mean pitch range values for the syllable míng (N = 5) with respect to stresses 4 and 3

In Section 4.5, we have already shown that stress 3 tends to produce larger pitch range values for the syllable liáo (except for subject 3). This difference can also be accounted for by the semantic load of the syllable. Under stress 3, the syllable liáo is narrowly focused. It
obviously carries heavier semantic weight. So, it is within our expectation that the syllable tends to have greater pitch range values.

To summarize, the acoustic distinction between stress 3 and stress 4 mainly shows up in the name syllable and the syllable liáo. Stress 4 generally induces larger pitch ranges for the name syllable. However, for the syllable liáo, stress 3 tends to produce greater values of pitch ranges. As for the syllable míng, the effect of tonal contexts overshadows the difference between the two stress patterns. To explain the different acoustic behaviors of the two stress patterns, we suggest that they are related to the semantic load of the syllables involved. The name syllable carries a heavier semantic load under stress 4 than it does under stress 3. On the other hand, the syllable liáo under stress 3 is narrowly focused. In terms of information focus, it obviously is more important than the case in which the whole sentence is under broad focus. This can account for the tendency in our data that the syllable liáo tends to have larger pitch range values when it is narrowly focused.

4.7 Conclusion

In this chapter, we have considered the pitch range variations of the three target syllables with respect to the four stress conditions. The general conclusion is that, when
a syllable is stressed, its pitch range will expand dramatically. The extent of this expansion is related to the position of the stressed syllable in the sentence. The closer the target is toward the sentence beginning, the larger the extent of expansion when the syllable is stressed.

We find that, under stress 4, the pitch range of the name syllable also expands to a considerable degree. However, for this stress pattern, the pitch range changes do not form a sharp contrast in F0 excursion. A stress pattern that realizes narrow focus, on the other hand, is often accompanied by an immediate fall of the F0 values after the stressed element.

Our investigation of the tonal contour of the syllable mǐng also suggests that the rising tone after a stressed syllable does not always lose its inherent tonal shape. We have examined the post-stress F0 values at points (c) and (d). The results indicate that, to a large extent, the F0 values at these points are determined by the tone of the preceding stressed syllable. When that syllable has a high-falling (T4) tone, the following rising (T2) syllable will exhibit a flat contour. When the preceding tone is low-dipping (T3), we will see a post-stress rising excursion of the F0 value. On the other hand, if the preceding stressed name syllable is high, but non-falling (T1 or T2), in most
cases, the following míng loses its inherent rising contour and becomes falling.

The phenomenon indicates that a pitch range reduction model is appropriate for describing stress in Mandarin Chinese. In this model, the reduction of the top line mainly depends on the stress conditions and the tone of the stressed syllable. A stressed high tone tends to reduce the top line of the following syllable more than does a stressed low tone. The explanation of the post-stress rising, flat and falling F0 excursions basically depends on the starting F0 point and the top line of the syllable. T3 differs from other tones in that it has a low target tone. When this tone is stressed, the following tone starts on a much lower F0 level. At the same time, it has a relatively higher top line. This results in a post-stress rising F0 contour. When a T4 is stressed, the target high tone is immediately followed by a low tone which serves as the starting point of the syllable míng. However, because of a much reduced top line, not much pitch range is left for the inherent rising tone. When a T1 or T2 is stressed, the following syllable starts off on a higher F0 level. However, the reduction of the top line is so much that the F0 excursion glides directly toward the top line and bypasses the low target of the syllable. This results in a falling F0 excursion for the syllable míng that is inherently rising.
Lastly, we have found that stress 3 is largely related to stress 4. In both cases, the overall F0 excursion is rather flat. The difference between the two stress patterns mainly shows up in the pitch ranges of the name syllable and the syllable liáo. Under stress 4, the name syllable has a larger pitch range. In contrast, the pitch range for the syllable liáo is greater under stress 3. Since the two stress patterns are related to broad versus narrow focus, the difference may well be related to the semantic weight associated with the particular focus pattern. Under broad focus, the name syllable carries a heavier information load. On the other hand, under narrow focus, the last syllable liáo is more significant in terms of the semantic information it carries. This explanation accounts for the acoustic difference between the two stress patterns.
Chapter V

Duration and intensity

In the last chapter, we have examined the first acoustic parameter, F0, of stress. It was found that the parameter is largely responsible for various pitch excursions related to different stress patterns. In the present chapter, we will investigate the second and the third acoustic correlates and see what effect they have on stress.

5.1 Duration

In this section, we will consider the second acoustic correlate of stress for the three target syllables. It has been noted in various places in the literature (Chapter Two) that duration is an important factor for stress. Our major aim in this section is to find the general as well as particular tendencies in our data with respect to the temporal domain of stress and see if conclusions drawn from there are compatible with previous studies.
5.1.1 The name syllable

In this section, we will examine the name syllable. We begin our discussion by analyzing the variance of the data, the results of which are presented in the following table.

Table 5.1 Analysis of variance for the duration of the name syllables

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>279795.6</td>
<td>93265.2</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>40452.0</td>
<td>13484.0</td>
<td>13.33</td>
<td>0.001</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>139573.3</td>
<td>46524.4</td>
<td>16.35</td>
<td>0.001</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>9102.2</td>
<td>1011.4</td>
<td>1.37</td>
<td>0.249</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>25611.0</td>
<td>2845.7</td>
<td>3.86</td>
<td>0.003</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>4440.6</td>
<td>493.4</td>
<td>0.67</td>
<td>0.729</td>
</tr>
<tr>
<td>subject*tone*stress</td>
<td>27</td>
<td>19912.2</td>
<td>737.5</td>
<td>1.64</td>
<td>0.028</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>115359.6</td>
<td>450.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>634246.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated

The above table shows that the duration of the target syllable is related to tone (F [3, 256] = 13.33, p = 0.001) and stress (F [3, 256] = 16.35, p = 0.001). In addition to the two main effects, interactions between subject and stress (F [9, 256] = 3.86, p = 0.003) and among the three variables (F [27, 256] = 1.64, p = 0.028) are also found to be significant.
The phenomenon that syllable duration is related to tone is well documented in the literature (Lin 1988, Tseng 1981). The analysis of variance shows that the effect of tone on the duration of syllable is also present in our data. It is found, however, that tone does not interact with stress ($F_{[9, 256]} = 0.67$, $p = 0.729$). In other words, for every tone, stress exerts an equal amount of effect on the duration of the target syllable. But this is not necessarily the case with every subject since the three-way interaction in our data is found to be significant.

In the following discussion, we will mainly limit ourselves to the main effect of stress and the three-way interaction. First we present the mean values of duration for the target name syllable with respect to stress in the following table.

<table>
<thead>
<tr>
<th>Stress 1</th>
<th>Stress 2</th>
<th>Stress 3</th>
<th>Stress 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>206.18 (44.91)</td>
<td>176.4 (38.64)</td>
<td>150.21 (34.02)</td>
<td>162.25 (39.93)</td>
</tr>
</tbody>
</table>

The above table shows that when the target syllable is stressed, that is, under stress 1, it has the greatest mean
value. The significance for the mean values in the above table is further examined through the multiple post-hoc Tukey comparisons. The results show that the mean difference between the durations of the name syllable under stress 1 and the corresponding values under stress 2 is significant. The 95% confidence interval does not include zero (29.78±16.06 ms). When the same procedure is done to stress 3 and stress 4, we find that, in both cases, the results are also significant (55.97±16.06 ms between stress 1 and stress 3 and 43.93±16.06 ms between stress 1 and stress 4). Thus, our initial conclusion for the duration parameter is that it is related to stress. When the name syllable is stressed, it is significantly longer than the cases when the syllable is not stressed.

Now we investigate the interaction of subject, tone, and stress since the analysis of variance yields the result of a significant three-way interaction (F [27, 256] = 1.64, p = 0.028). In the following figure, I present the mean values of duration for the name syllable with respect to the three variables (the actual corresponding values are also summarized in Appendix F).
Figure 5.1 Mean values of duration (ms) of the name syllable with respect to subject, tone, and stress
The above figure shows a clear tendency that the name syllable has the longest duration under stress 1, that is, when the syllable itself is stressed. The smallest target values of the four stress conditions in most cases occur under stress 3. For subject 3, stress 4 tends to induce the shortest duration for the target syllable. However, under stress 1, the syllable is still the longest for the same subject.

The figure also shows a general tendency that the mean values of stress 2 are in most cases located between those of stress 1 and stress 3. This indicates that, when the following syllable míng is stressed, the name syllable tends to be longer than if the syllable liáo is stressed. Although in neither cases is the name syllable stressed, the duration can still be affected by its relative position from the place of stress.

In order to test if the mean duration values of the name syllable presented in Figure 5.1 are significant when they are compared with each other, multiple post-hoc Tukey comparisons are performed with every subject. The results are presented in Table 5.3 below. It can be observed from the table that the durations of the target syllable under stress 1 are, in most cases, longer than the corresponding values under the other stress conditions. However, for subject 2, the target values are not always longer under
stress 1. The 95% confidence intervals between the mean values of the name syllable under stress 1 and those under stress 2 and stress 4 actually include zeros. This shows that for him the target values may not necessarily differ under those stress conditions. In spite of this inconsistency, we can still say that the name syllable has the longest duration when it is stressed.

Table 5.3 Pairwise Tukey comparisons of the duration of the name syllables

<table>
<thead>
<tr>
<th>Stress Diff.</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1 - s2</td>
<td>18.89, 73.91</td>
<td>-7.03, 27.73</td>
<td>10.75, 34.14</td>
<td>16.93, 62.87</td>
</tr>
<tr>
<td>s1 - s3</td>
<td>54.29, 109.31</td>
<td>9.27, 44.03</td>
<td>28.90, 52.29</td>
<td>51.83, 97.77</td>
</tr>
<tr>
<td>s1 - s4</td>
<td>36.69, 91.71</td>
<td>-1.88, 32.88</td>
<td>33.11, 56.50</td>
<td>28.23, 74.17</td>
</tr>
</tbody>
</table>

The above table also shows that the intervals for the comparisons between stress 1 and stress 3 are in all cases located within a higher range than those between stress 1 and stress 2. This indicates that the difference between the target values under stress 3 and those under stress 1 tends to be greater than the corresponding difference between stress 2 and stress 1. Again, we can use the relative position of the target from the stressed syllable to account
for this result. Under stress 3, the syllable liáo is stressed. The duration of the name syllable under this stress condition tends to be shorter than if the immediately following syllable míng is stressed. Although in both cases, the duration of the syllable would be much longer if the syllable itself is stressed.

To summarize, the duration of the name syllable may depend on two factors, stress and its relative position from another stressed syllable. Generally, we can say that stressing the name syllable can significantly extend its duration. However, this is not the only way that this syllable can get lengthened. In some cases when the following míng is stressed, the duration of the name syllable can also get affected and becomes longer.

5.1.2 The syllable míng

In this section, we investigate the effect of stress on the duration of the syllable míng. First, we present the results of the analysis of variance for the target values in Table 5.4 below. It is clear from the p-values that stress has a significant effect on the duration of the syllable míng (F [3, 256] = 15.26, p = 0.001). Of all the interactions, only the subject-stress interaction is found to be significant (F [9, 256] = 4.33, p = 0.001). In the
discussions below, we mainly present analysis of the stress effects. The mean values are summarized in Table 5.5 below.

Table 5.4 Analysis of variance for the duration of the syllable mǐng

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>189272.0</td>
<td>63090.7</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>4750.5</td>
<td>1583.5</td>
<td>2.48</td>
<td>0.127</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>73191.9</td>
<td>24397.3</td>
<td>15.26</td>
<td>0.001</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>5739.7</td>
<td>637.7</td>
<td>1.73</td>
<td>0.131</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>14393.0</td>
<td>1599.2</td>
<td>4.33</td>
<td>0.001</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>2835.7</td>
<td>315.1</td>
<td>0.85</td>
<td>0.576</td>
</tr>
<tr>
<td>subject<em>tone</em>stress</td>
<td>27</td>
<td>9970.2</td>
<td>369.3</td>
<td>1.35</td>
<td>0.122</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>70024.0</td>
<td>273.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>370176.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated

Table 5.5 Mean values of duration of syllable mǐng with respect to stress (N = 80)

<table>
<thead>
<tr>
<th>Stress 1</th>
<th>Stress 2</th>
<th>Stress 3</th>
<th>Stress 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>148.80 (27.68)</td>
<td>178.36 (33.06)</td>
<td>140.86 (28.95)</td>
<td>142.65 (32.59)</td>
</tr>
</tbody>
</table>

Table 5.5 shows that under stress 2, that is, when the target syllable mǐng is stressed, it has the longest duration. When the syllable is not stressed (at stresses other than 2), it is relatively short. The results of the
multiple post-hoc Tukey comparisons support this observation. It is found that, under stress 2, the target mean values are significantly greater than those under either stress 1, stress 3 or stress 4 (29.56±12.44 ms for stress 2 minus stress 1, 37.5±12.44 ms for stress 2 minus stress 3 and 35.71±12.44 ms for stress 2 minus stress 4). None of the 95% confidence intervals include zero. This indicates that stress has a general effect on the duration of the syllable *míng*. When the syllable is stressed, its duration will become significantly longer.

When each subject is considered in turn for the stress effect on the target values, we find the same pattern. Figure 5.2 below shows that, for every subject, the values under stress 2 are longer than those under the other stresses. The interaction comes from the fact that subjects do not have the same proportion with respect to the difference of values under stress 2 as opposed to the other stresses. For example, for subject 1, the mean values under stresses 1, 3, and 4 are smaller than those of subject 2. However, under stress 2, subject 1 has greater mean values. So, for him, the difference between values under stress 2 on the one hand and those under the other stresses on the other are greater than those of subject 2.
To summarize, we find that the duration of the syllable mǐng is significantly longer when it is stressed. Although subjects do not have the same behavior with respect to the stress effect under various stress conditions, the general tendency that the syllable mǐng becomes longer when it is stressed is the same across all subjects.

5.1.3 The syllable liáo

In this section, we will examine the duration of the syllable liáo with respect to stress. The results of the
analysis of variance of the target values are first presented in the following table.

Table 5.6 Analysis of variance for the duration of the syllable liáo

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>105171.9</td>
<td>35057.3</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>812.3</td>
<td>270.8</td>
<td>0.83</td>
<td>0.508</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>31809.8</td>
<td>10603.3</td>
<td>11.41</td>
<td>0.002</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>2920.5</td>
<td>324.5</td>
<td>0.89</td>
<td>0.546</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>8364.6</td>
<td>929.4</td>
<td>2.55</td>
<td>0.029</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>1705.0</td>
<td>189.4</td>
<td>0.52</td>
<td>0.847</td>
</tr>
<tr>
<td>subject<em>tone</em>stress</td>
<td>27</td>
<td>9831.1</td>
<td>364.1</td>
<td>2.14</td>
<td>0.001</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>43497.2</td>
<td>169.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>204112.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated

The above table shows that the duration of the syllable liáo is related to the main effect of stress, the interaction of subject and stress and the interaction of subject, tonal context, and stress. In the following discussion, we mainly focus on the main effect of stress. The mean values are first summarized in the following table.
Table 5.7 Mean values of duration of the syllable liáo with respect to stress (N = 80)

<table>
<thead>
<tr>
<th>Stress 1</th>
<th>Stress 2</th>
<th>Stress 3</th>
<th>Stress 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>198.66 (23.69)</td>
<td>197.24 (25.68)</td>
<td>222.30 (20.81)</td>
<td>204.74 (22.97)</td>
</tr>
</tbody>
</table>

The above table shows that, when the syllable liáo is stressed (stress 3), it has the greatest mean value of duration. Multiple post-hoc Tukey comparisons show that the duration of this syllable under stress 3 is significantly longer than the corresponding values under the other stress conditions. None of the 95% confidence intervals include zero (23.64±9.48 ms for stress 3 minus stress 1, 25.06±9.48 ms for stress 3 minus stress 2 and 17.56±9.48 ms for stress 3 minus stress 4). Thus, it can be generally concluded that, when the target syllable liáo is stressed, its duration will become longer.

5.1.4 A comparison of stress 3 and stress 4

In this section, we will examine the mean values of duration for the three target syllables with respect to stress 3 and stress 4. We first consider the name syllable. Presented in the following table are mean values of this syllable under the four stress patterns together with the results of the two-tailed t-tests.
Table 5.8 Comparisons of the mean duration (ms) for the name syllable (N = 20) under stresses 3 and 4

<table>
<thead>
<tr>
<th></th>
<th>stress 3</th>
<th>stress 4</th>
<th>2-tailed t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>subj 1</td>
<td>113.7 (19.2)</td>
<td>131.3 (32.2)</td>
<td>-2.10/0.044/31</td>
</tr>
<tr>
<td>subj 2</td>
<td>146.3 (19.0)</td>
<td>157.4 (18.0)</td>
<td>-1.91/0.064/37</td>
</tr>
<tr>
<td>subj 3</td>
<td>150.4 (15.5)</td>
<td>146.1 (16.4)</td>
<td>0.83/0.41/37</td>
</tr>
<tr>
<td>subj 4</td>
<td>190.6 (26.6)</td>
<td>214.1 (29.4)</td>
<td>-2.66/0.011/37</td>
</tr>
</tbody>
</table>

*t-value/p-value/degree of freedom

The table shows that, for subjects 1 and 4, the duration of the name syllable is significantly longer under stress 4 than it is under stress 3. Although for subject 3 the mean value under stress 3 is slightly longer, the difference is not significant. Generally speaking, the data suggest that, under the stress 4 condition, the duration of the name syllable tends to get longer than it does under stress 3. This finding is compatible with the one in Section 4.6. Stress 4 is related to broad focus. Under this focus pattern, the name syllable carries more information. This can account for the tendency of longer duration. However, it should be noted that the differences between the two stress conditions are rather trivial. In some cases, longer
duration is found under stress 3 rather than stress 4. This can be illustrated by the following scatter plot.

![Scatter plot of duration of the name syllable (N = 5) with respect to stresses 4 and 3](image)

Figure 5.3 Scatter plot of duration of the name syllable (N = 5) with respect to stresses 4 and 3

The above figure shows that all data points are very close to the equal-value line. Although some of the points fall on the side of stress 3, the general tendency is that the duration values under stress 4 (broad focus) are slightly longer.

Now, we examine the syllable míng. In Section 4.6, we found that this syllable does not show any difference in the F0 parameter with respect to the two stress conditions. How
about duration? In the following table, we present the mean values and the results of the two-tailed t-test.

Table 5.9 Comparisons of the mean duration (ms) for the syllable míng (N = 20) under stresses 3 and 4

<table>
<thead>
<tr>
<th></th>
<th>stress 3</th>
<th>stress 4</th>
<th>2-tailed t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>subj 1</td>
<td>104.85 (12.22)</td>
<td>102.30 (12.66)</td>
<td>0.65/0.52/37</td>
</tr>
<tr>
<td>subj 2</td>
<td>132.30 (8.95)</td>
<td>135.65 (12.75)</td>
<td>-0.96/0.34/34</td>
</tr>
<tr>
<td>subj 3</td>
<td>152.00 (11.76)</td>
<td>148.15 (14.02)</td>
<td>0.94/0.35/36</td>
</tr>
<tr>
<td>subj 4</td>
<td>174.30 (19.14)</td>
<td>184.50 (16.02)</td>
<td>-1.83/0.076/36</td>
</tr>
</tbody>
</table>

't-value/p-value/degree of freedom

The above table shows that, like the F0 parameter, the durations of the syllable míng are not distinguished under the two stress conditions. All t-test results yield insignificant values. This indicates that the two stress conditions, stress 3 and stress 4, do not have significant effects on the duration of the syllable míng.

In the following scatter plot, data points are presented with respect to the two stress conditions. It can be observed that the distribution of the data is rather even along the equal-value line. This indicates a general
tendency of non-distinction between the duration values under the two stress conditions.

![Graph showing scatter plot of duration of the syllable míng (N = 5) with respect to stresses 4 and 3.](image)

**Figure 5.4** Scatter plot of duration of the syllable míng (N = 5) with respect to stresses 4 and 3

In Section 5.1.3, we have already noted that the duration of the syllable liáo tends to become longer under stress 3. Although in both cases, the stressed syllable is most probably liáo (judging by its acoustic behavior and the results of the perception test), the syllable is usually longer under stress 3 (narrow focus).

To summarize, the difference between the duration of the three target syllables under stress 3 and stress 4 reflects the same distinction found for the F0 values. Under stress
4, the name syllable tends to get longer; under stress 3, the syllable liáo is usually found to be longer.

5.1.5 Conclusion

In this section, the parameter of duration of the three target syllables has been considered with respect to the four stress conditions. We find that stressing the target syllable generally makes that syllable longer. This is true when the target syllables are narrowly focused.

We also find that, when a target syllable is not stressed, its relative position to the stressed syllable will exert some influence on its duration. The farther away the target syllable is from the stressed syllable, the shorter it tends to become. Thus, comparing the duration of the name syllable under stress 2 and stress 3, we will expect the one under stress 2 to be longer. This difference might be related to the duration of the stressed syllable. When a syllable is stressed, its long duration may positively influence its neighboring syllable.

Finally, when stress 3 is compared with stress 4, we find that, under stress 4, the name syllable has longer duration, while under stress 3, the syllable liáo is more influenced. This tendency turns out to be the same as the one in which F0 parameter is considered. In both cases, larger values are related to heavier information load. Under
stress 4, the name syllable has greater semantic values. On the other hand, under stress 3, greater values are associated with the syllable liáo. These provide evidence for a semantic connection between stress and focus.

5.2 Intensity

In this section, we will investigate how intensity, our third acoustic parameter, influences stress. As is noted in Chapter Two, the general consensus is that this parameter is not much related to stress. Again, the data will be examined with respect to the three target syllables.

5.2.1 The name syllable

We begin our discussion by examining the intensity data for the name syllable. Table 5.10 below summarizes the results of the analysis of variance. It shows that intensity of the name syllable is related to tone (F [3, 256] = 13.49, p = 0.001) and stress (F [3, 256] = 7.56, p = 0.008). The only significant stress-related interaction is the three-way interaction (F [27, 256] = 1.54, p = 0.048). Stress is involved in the other two interactions: between subject and stress (F [9, 256] = 1.04, p = 0.438) and between tone and stress (F [9, 256] = 1.00, p = 0.462). However, neither are found to be significant.
Table 5.10 Analysis of variance for the intensity of the name syllables

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>4102.04</td>
<td>1367.35</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>3137.95</td>
<td>1045.98</td>
<td>13.49</td>
<td>0.001</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>736.62</td>
<td>245.54</td>
<td>7.56</td>
<td>0.008</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>697.87</td>
<td>77.54</td>
<td>2.47</td>
<td>0.033</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>292.48</td>
<td>32.50</td>
<td>1.04</td>
<td>0.438</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>282.78</td>
<td>31.42</td>
<td>1.00</td>
<td>0.462</td>
</tr>
<tr>
<td>subject<em>tone</em>stress</td>
<td>27</td>
<td>846.66</td>
<td>31.36</td>
<td>1.54</td>
<td>0.048</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>5222.27</td>
<td>20.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>15318.67</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated

We first examine the main effect of stress. In the following table, the mean values for the intensity data with respect to stress are presented.

Table 5.11 Mean values of intensity (dB) of the name syllable with respect to stress (N = 80)

<table>
<thead>
<tr>
<th>Stress 1</th>
<th>Stress 2</th>
<th>Stress 3</th>
<th>Stress 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.80 (8.20)</td>
<td>25.72 (6.33)</td>
<td>26.82 (6.23)</td>
<td>26.85 (6.20)</td>
</tr>
</tbody>
</table>

Although the above table indicates that the mean value of the intensity of the target syllable under stress 1 is
greater than those under the other stresses, the differences are very small. Multiple post-hoc Tukey comparisons show marginal significance (4.08±2.75 db for stress 1 minus stress 2, 2.98±2.75 db for stress 1 minus stress 3 and 2.95±2.75 db for stress 1 minus stress 4). This suggests that, when the name syllable is stressed, its intensity is generally somewhat greater.

Now, we examine the three-way interaction. The mean values (N = 5) with respect to the three variables are first presented in Figure 5.5 below (see Appendix G for actual values).

It can be observed from the figure that, under stress 1, the name syllable generally has higher intensity values. However, there are many exceptions. For example, subject 1 and subject 2 actually have lower intensity values at stress 1 for tone 3. This subject-stress interaction further weakens the claim that the target syllable has higher intensity when it is stressed.

In order to further evaluate the significance of the intensity values under stress 1 against those under the other stresses, multiple post-hoc Tukey comparisons are performed for every subject. The results are summarized in Table 5.12 below.
Figure 5.5 Mean values of intensity of the name syllable with respect to subject, tone, and stress (N = 5)
Table 5.12  Pairwise Tukey comparisons of the intensity values of the name syllable

<table>
<thead>
<tr>
<th>Stress Diff.</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1 - s2</td>
<td>-2.49, 8.00</td>
<td>-1.26, 5.73</td>
<td>1.57, 10.09</td>
<td>-0.38, 11.44</td>
</tr>
<tr>
<td>s1 - s3</td>
<td>-3.53, 6.95</td>
<td>-1.57, 5.42</td>
<td>0.60, 9.12</td>
<td>-2.46, 9.35</td>
</tr>
<tr>
<td>s1 - s4</td>
<td>-3.18, 7.30</td>
<td>-2.67, 4.33</td>
<td>2.65, 11.17</td>
<td>-3.87, 7.94</td>
</tr>
</tbody>
</table>

The results in the above table indicate that intensity is generally not related to stress for the name syllable. For three subjects, the mean target values are not significantly different under stress 1 from those under the other stresses. Only subject 3 have confidence intervals that do not include zero. For her, the target values tend to increase when the syllable is stressed.

It is a well-known acoustic phenomenon that intensity is largely related to fundamental frequency. A higher frequency sound tends to have higher intensity value. This is related to the greater respiratory effort involved in increasing the vibration rate of the vocal folds (Lehiste 1970). Whalen and Xu (1992) shows that amplitude contours can be used by the listeners as cues to identify different tones. The results of their study indicate that F0 parameter may be positively correlated with amplitude, that is, the changing of intensity over time.
In our data, intensity is highly correlated with T4. In Figure 5.5, it can be seen that T4 generally has the highest intensity values among the four tones under a given stress condition. In contrast to T4, the behavior of the other tones, especially T3, with respect to intensity is rather unpredictable. Since T3 is a low tone, it would be expected that its overall intensity level should be lower. However, this is not generally supported by our data. For example, subject 4 consistently shows a higher T3 value than T1 and T2 and subject 2 exhibits such a tendency most of the time.

As for T1 and T2, again no general pattern across the four stress conditions can be found. Although, for subject 3 and subject 4, stress 1 consistently produces higher intensity values, the other two subjects do not show the same tendency. It can be found that, in some cases, the intensity values of the name syllables are actually lower under stress 1 than those under stress 2.

To summarize, intensity is not generally related to stress. Although our data show that T4 tends to induce higher intensity values, no generalizable pattern can be found for the other three tones. The results of the pairwise Tukey comparisons also confirm that, for most speakers, intensity is not related to stress. Thus, our general conclusion is that intensity is not a reliable acoustic parameter for indicating stress for the name syllable.
5.2.2 The syllable míng

In this section, the intensity parameter for the syllable míng will be examined. We begin by presenting the analysis of variance of the target values in the following table.

Table 5.13 Analysis of variance for the intensity of the syllable míng

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>3509.90</td>
<td>1169.97</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>1815.20</td>
<td>605.07</td>
<td>9.27</td>
<td>0.004</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>456.54</td>
<td>152.18</td>
<td>3.64</td>
<td>0.057</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>587.34</td>
<td>65.26</td>
<td>6.07</td>
<td>0.000</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>375.77</td>
<td>41.75</td>
<td>3.88</td>
<td>0.003</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>144.11</td>
<td>16.01</td>
<td>1.49</td>
<td>0.202</td>
</tr>
<tr>
<td>subject<em>tone</em>stress</td>
<td>27</td>
<td>290.26</td>
<td>10.75</td>
<td>2.04</td>
<td>0.003</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>1350.46</td>
<td>5.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>8529.58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated

The above table shows that stress does not significantly affect the intensity of the our target syllable in question (F [3, 256] = 3.64, p = 0.057). However, we do find two stress related interactions: one between subject and stress (F [9, 256] = 3.88, p = 0.003) and the other among all the three variables (F [27, 256] = 2.04, p = 0.003). Thus, further examination is needed.
First, we investigate the interaction between subject and stress. The following figure summarizes the mean values of intensity of our target with respect to subject.

![Graph showing intensity (db) vs. stress for four subjects.]

Figure 5.6 Mean values of intensity of syllable míng with respect to subject and stress

It can be observed from the figure that the intensity values at stress 2 are not predominantly high among all the stress conditions. The greatest mean values usually do not occur under stress 2 when the target syllable is stressed. Thus, it can be said that, for each subject, the intensity level of the syllable míng is not related to stress.
Figure 5.7 Mean values of intensity of the syllable míng with respect to subject, tonal context, and stress
We now consider the three-way interactions in the data. Figure 5.7 above summarizes the mean target values with respect to the three variables (see Appendix K for actual values). It can be observed that the highest mean target values do not always coincide with stress 2, suggesting that stress conditions are not mainly responsible for the variations in the data. This further confirms our conclusion that intensity of the syllable míng is not related to stress.

5.2.3 The syllable liáo

In this section, we will investigate the behavior of our last target syllable liáo with respect to intensity and stress. First, we perform the analysis of variance on the data. The results are presented in the following table.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject</td>
<td>3</td>
<td>9270.39</td>
<td>3090.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tone</td>
<td>3</td>
<td>11.84</td>
<td>3.95</td>
<td>0.77</td>
<td>0.537</td>
</tr>
<tr>
<td>stress</td>
<td>3</td>
<td>8431.38</td>
<td>2810.46</td>
<td>21.42</td>
<td>0.000</td>
</tr>
<tr>
<td>subject*tone</td>
<td>9</td>
<td>45.89</td>
<td>5.10</td>
<td>0.62</td>
<td>0.771</td>
</tr>
<tr>
<td>subject*stress</td>
<td>9</td>
<td>1181.11</td>
<td>131.23</td>
<td>15.92</td>
<td>0.000</td>
</tr>
<tr>
<td>tone*stress</td>
<td>9</td>
<td>148.81</td>
<td>16.53</td>
<td>2.01</td>
<td>0.079</td>
</tr>
<tr>
<td>subject<em>tone</em>stress</td>
<td>27</td>
<td>222.62</td>
<td>8.25</td>
<td>0.60</td>
<td>0.944</td>
</tr>
<tr>
<td>Error</td>
<td>256</td>
<td>3522.06</td>
<td>13.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>319</td>
<td>22834.09</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no exact F-test can be calculated
The above table shows a main effect of stress ($F[3, 256] = 21.42, \ p = 0.000$) and an interaction of subject and stress ($F[9, 256] = 15.92, \ p = 0.000$). These will be discussed separately below.

We first consider the main effect of stress on the intensity of the target syllable. The following table summarizes the mean values of the syllable under different stress conditions.

<table>
<thead>
<tr>
<th>Stress 1</th>
<th>Stress 2</th>
<th>Stress 3</th>
<th>Stress 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.57 (4.54)</td>
<td>14.62 (5.75)</td>
<td>25.46 (8.93)</td>
<td>21.08 (6.98)</td>
</tr>
</tbody>
</table>

It can be observed in the table that the greatest mean value of intensity is located under stress 3. Although the value at stress 4 is very close to that at stress 3, multiple post-hoc Tukey comparisons indicate that the values under the two stress conditions are significantly different from each other ($4.38\pm2.74 \mathrm{~dB}$ for the mean difference of stress 3 minus stress 4). As for the other two stress conditions, the pairwise Tukey comparisons yield more significant results ($12.89\pm2.74 \mathrm{~dB}$ for the means of stress 3 minus stress 1 and
$10.84 \pm 2.74$ db for the same values under stress 3 minus stress 2). It can thus be concluded that when the target syllable liáo is stressed, it has higher values of intensity.

Although the mean intensity value for the syllable liáo under stress 4 has been found to be significantly lower than that under stress 3, it is still significantly higher than the corresponding values under either stress 1 ($8.51 \pm 2.74$ db for stress 4 minus stress 1) or stress 2 ($6.46 \pm 2.74$ db for stress 4 minus stress 2). Thus, generally we can say that when the syllable liáo is stressed either for narrow focus (stress 3) or for broad focus (stress 4), it has greater intensity.

![Figure 5.8 Mean values of intensity (db) of the syllable liáo for the four subjects with respect to stress](image)
Now we examine the interaction of subject and stress. The mean values are summarized in Figure 5.8 above. The figure clearly shows that, when the target syllable is under stress 3, the intensity tends to become greater. It can also be observed from the figure that the target values under stress 3 and stress 4 are generally higher than those under the other stress conditions. Both tendencies can be generalized across different subjects. Table 5.16 below summarizes the results of the pairwise Tukey comparisons.

Table 5.16 Pairwise Tukey comparisons of the intensity values of the syllable liáo

<table>
<thead>
<tr>
<th>Stress Diff.</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>s3 - s1</td>
<td>14.67, 23.10</td>
<td>5.38, 9.22</td>
<td>9.43, 13.15</td>
<td>10.79, 17.42</td>
</tr>
<tr>
<td>s3 - s2</td>
<td>10.16, 18.61</td>
<td>2.65, 6.48</td>
<td>8.54, 12.26</td>
<td>10.73, 17.36</td>
</tr>
<tr>
<td>s3 - s4</td>
<td>3.39, 12.33</td>
<td><strong>-0.30, 3.54</strong></td>
<td>3.29, 7.01</td>
<td><strong>-0.58, 5.95</strong></td>
</tr>
<tr>
<td>s4 - s1</td>
<td>6.55, 14.99</td>
<td>3.76, 7.60</td>
<td>4.28, 8.00</td>
<td>8.15, 14.78</td>
</tr>
<tr>
<td>s4 - s2</td>
<td>2.05, 10.50</td>
<td>1.03, 4.86</td>
<td>3.39, 7.11</td>
<td>8.09, 14.72</td>
</tr>
</tbody>
</table>

The results of the pairwise Tukey comparisons in the above table show that the target values under stress 3 are significantly greater than those under stress 1 and stress 2
for all subjects. Stress 4 is similar to stress 3 in that it also produces intensity values for the target syllable liáo which are relatively greater than the ones produced by either stress 1 or stress 2. As for the comparison between stress 3 and stress 4, subject 2 and subject 4 yield insignificant results. This indicates that, for these two subjects, stress 3 might not be different from stress 4 in terms of the intensity level of the target syllable. However, since zero is very close to the lower bounds of their confidence intervals, we may say that the two stress conditions are marginally different for these two speakers.

To summarize, we have found that, unlike the other two target syllables, the intensity of the syllable liáo is related to stress. The target values under stress 3 and stress 4 are found to be significantly greater than those under the other two stress conditions. However, stress 3 does not always produce higher intensity values than does stress 4. For two subjects, the two stress conditions are only marginally different, and so, may yield similar target values for the syllable liáo.

5.2.4 A comparison of stress 3 and stress 4

In this section, we compare the intensity values of the three target syllables under stress 3 and stress 4. We begin with the name syllable. The following table summarizes the
mean values of this syllable under the two stress conditions and the two-tailed t-test results.

<table>
<thead>
<tr>
<th></th>
<th>stress 3</th>
<th>stress 4</th>
<th>2-tailed t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>subj 1</td>
<td>28.86 (5.59)</td>
<td>28.51 (5.40)</td>
<td>0.20/0.84/37</td>
</tr>
<tr>
<td>subj 2</td>
<td>20.45 (3.53)</td>
<td>21.55 (4.41)</td>
<td>-0.87/0.39/36</td>
</tr>
<tr>
<td>subj 3</td>
<td>29.47 (5.05)</td>
<td>27.42 (4.93)</td>
<td>1.30/0.20/37</td>
</tr>
<tr>
<td>subj 4</td>
<td>28.49 (5.91)</td>
<td>29.91 (6.65)</td>
<td>-0.71/0.48/37</td>
</tr>
</tbody>
</table>

*t-value/p-value/degree of freedom

The above table shows that, under the two stress patterns being considered, the mean values of intensity for the name syllable do not differ. None of the test results turn out to be significant.

This situation can be further summarized by the scatter plot in Figure 5.9 below. Each data point in the figure represents an average value of five tokens. It can observed from the plot that the data points are evenly distributed on both sides of the equal-value line. This indicates that the target values under the two stress conditions are basically not different. We thus conclude that intensity values are
Table 5.18 Comparisons of the mean intensity (db) for the syllable mǐng (N = 20) under stresses 3 and 4

<table>
<thead>
<tr>
<th></th>
<th>stress 3</th>
<th>stress 4</th>
<th>2-tailed t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>subj 1</td>
<td>30.07 (3.35)</td>
<td>18.93 (4.49)</td>
<td>0.91/0.37/35</td>
</tr>
<tr>
<td>subj 2</td>
<td>12.30 (1.90)</td>
<td>11.85 (2.13)</td>
<td>0.70/0.49/37</td>
</tr>
<tr>
<td>subj 3</td>
<td>20.90 (4.25)</td>
<td>19.72 (4.74)</td>
<td>0.83/0.41/37</td>
</tr>
<tr>
<td>subj 4</td>
<td>15.58 (3.01)</td>
<td>16.91 (4.45)</td>
<td>-1.11/0.28/33</td>
</tr>
</tbody>
</table>

*t-value/p-value/degree of freedom

Figure 5.10 Scatter plot of intensity (db) for the syllable mǐng (N = 5) with respect to stresses 4 and 3

We have already discussed the intensity distribution of the two stress patterns for the syllable liǎo in the previous
section. When the intensity values of this syllable under stress 3 is compared with those under stress 4, the difference is marginal for two speakers and significant for the other two. So the general tendency seems to be that stress 3 induces somewhat greater intensity values than does stress 4.

To summarize, we find that the intensity values under stress 3 are not different from those under stress 4 for the name syllable and the syllable míng. For the syllable liáo, the general tendency is that intensity values under stress 3 are greater than those under stress 4. However, since the results for two speakers are marginal, the claim that the syllable liáo has higher intensity values under stress 3 is a weaker one.

5.2.5 Looking across the sentence

In this section, we examine the intensity data across the three target syllables. First, mean values of these syllables under the four stress conditions are presented in the following figure.
Figure 5.11  Mean values of intensity for the three target syllables with respect to stress

The values summarized in the above figure show that intensity for the three target syllables is a function of position. Sentence-initial syllables have higher values of intensity and those syllables located toward the end of the sentence have much lower intensity values. It can be observed in the figure that the intensity of the name syllable for all the four stress conditions is basically on a higher level. Although a slight peak can be observed for the name syllable under stress 1, this does not seem to be important since the intensity parameter for that syllable has been determined to be unrelated to stress. The behavior of
the syllable míng is entirely different. The mean values for the four stress conditions look rather even. No stress related peak for the intensity value can be found at all. Syllable liáo, on the other hand, shows a very conspicuous peak at the stress position. When it is not focused under either stress 1 or stress 2, the mean values are slightly lower than those of syllable míng under the same stress condition.

A general pattern that emerges from this description is that intensity parameter is related to stress when the target syllable is close to the sentence boundary. At the sentence-initial position, however, the stress effect cannot stand out very well due to the unmarked level of high intensity at this position. In Section 4.1, we observed that the general contour of a sentence is that it starts off with a higher F0 value and progressively falls off towards the end. Thus, it is not surprising that the intensity level at the beginning of a sentence tends to be high given the general correlation between intensity and fundamental frequency (Lehiste 1970, Whalen and Xu 1992). On the other hand, at the sentence-final position, the unmarked level for the intensity becomes very low. In this environment, when a syllable is stressed, its intensity will dramatically increase. As a result, a sharp contrast can be observed. Thus, stress effect at this position is the most conspicuous.
5.2.6 Conclusion

In this section, we have considered the parameter of intensity for the three target syllables with respect to stress. We find that the effect of this acoustic parameter on stress is generally related to the position of the target word in the sentence.

For the name syllable, we find that, although the mean values of intensity is significantly higher when the syllable is stressed, it turns out that the parameter is unimportant in determining stress. No obvious patterns can be found that are related to stress and can be generalized across subjects and tones.

For the syllable míng, we find that, when it is stressed, its intensity in most cases is not greater when the same syllable is under a different stress condition. Likewise, when it is not stressed, its intensity level may not be necessarily low. This shows that the intensity of the syllable míng, like the name syllable, is not related to stress.

Finally, our third target syllable liáo is located at the end of the sentence. The intensity does have a significant effect on that syllable. When the syllable is stressed, it exhibits higher intensity than if the name syllable or the syllable míng is stressed. We also find that, when the syllable is under stress 4, the mean intensity
value is also significantly higher than the corresponding values when the syllable is under either stress 1 or stress 2. However, the difference between stress 3 and stress 4 is only marginal for some speakers. When the target syllable is the name syllable or the syllable mìng there is no difference at all with respect to the two stress conditions.

When the three target syllables are compared across the sentence, we find that the intensity parameter for the stress effect is only important with respect to sentence positions. At sentence-initial or sentence-medial positions, intensity is not much related to stress. Only at the sentence-final position do we find high correlation between intensity and stress.
Note

By comparing the syllables across different sentence positions, we assume that the three target syllables have the same inherent intensity. The same comparisons are not performed for the F0 and the duration parameters. First, not all the three target syllables are controlled for tones. Second, the inherent duration for the three target syllables might not be the same, especially since the syllable liǎo has a medial which the other two target syllables do not have.
Chapter VI

Perception and stress

In the last two chapters, we have examined the three acoustic parameters - F0, syllable duration, and intensity - and their effects on stress. We find that both F0 and syllable duration are important correlates of stress. When a target syllable is stressed, it undergoes an increase in values for these two parameters. Although intensity is generally found not to be important for stress, at least for some target syllable (for example, the syllable liáo), this parameter turns out to be related to stress.

In Chapter Three, we presented the results of our perception experiment. It is found that subjects are generally able to perceive different focus patterns. Although focus 3 and focus 4 present some problems, the overall performance indicates that their judgment of the focus patterns is by no means random. To a large extent, the judgment is based on certain acoustic properties of the stimulus sentence. In this chapter, we will use multiple regression models to find the relationship between the
perception scores described in Chapter Three and the acoustic parameters. The result will be an indication of the relative importance of the different parameters in affecting the subjects' judgment of the individual tokens in the production data. Since top line reduction has been found to be related to stress, it will be added to our list. Altogether, we will consider four parameters: pitch range, top line, duration and intensity. For the same syllable, F0 values at point (a) or (b) is selected for indicating the top line depending on the tone of the syllable (point (a) for T3 and T4, point (b) for T1 and T2). For the other two target syllables, point (d) and point (f) are used for the same purpose. To simplify our analysis, we will disregard other possible factors that may have influenced the scores of the perception experiment.

6.1 Multiple regression

In this chapter, we use multiple regression models to analyze our data. Multiple regression analysis allows us to use more than one variable to find an equation that will provide the best fit of the relationship among data. We will use the subjects' score as the dependent variable and the four acoustic parameters as the independent ones. That is, we will be trying to predict the variability in the scores by the observed variability in the different independent variables. By using this model, a test statistic can be
calculated for each independent variable. The resulting F or t-value is an indication of whether the variable contributes significantly in accounting for the pattern of variation in the perception data.

I will use the stepwise regression procedure that is included in Minitab to calculate the model. This procedure first checks all the variables and fits them with the model one by one. The variable that yields the least residual of unexplained variability will be chosen. Then the rest of the variables will again be checked one by one and fitted to the model. The variable that gives the most significant improvement in the fit will be chosen. The same cycle is repeated until no further variables can be added. All the remaining variables are regarded as insignificant and will be subsequently dropped from the model.

I will use 4.0 as the F value to determine if a variable should be admitted into or removed from the model. In other words, a variable must at least have the F value 4.0 in order to enter the model. Once admitted into the model its F value must remain at or above 4.0 in subsequent fittings of the other variables. If at any point the F value falls below that critical point, the variable will be dropped.

In this study, a variable, unless later dropped, will be regarded as the first important if it is introduced into the model first and second important if it is introduced into the
model second and so on. This ranking of the variables will be used as statistical evidence to explain that certain acoustic parameters are more important than the others in influencing the subjects' judgment of a certain sentence as having a certain stress.

In regression analysis, the value $R^2$ indicates how well the model fits the data. This value can be thought of as the proportion of the observed variation in the data that can be explained by the model relationship. Thus, it is usually represented in terms of a percentage value. In any given cycle of model fitting, the higher the $R^2$ value, the better the model. In stepwise regression, the variable that yields the highest $R^2$ will be chosen in a particular variable selection cycle. Thus, that variable may be interpreted as the most important among all the other variables that have not yet been selected.

In the following analysis, we will first use the four acoustic parameters of individual syllables as the independent variables and the perceptual scores as the dependent ones. This will give us an indication of the relative importance of the three target syllables in accounting for the data. Next, we will include all the twelve parameters of the three target syllables in the model and see which parameter of which syllable is more significant in accounting for the variations in the perceptual scores.
6.2 Stepwise regression with four variables

In this section, we analyze the data by using four acoustic variables in the model. These will be pitch range, top line, duration and intensity of a particular target syllable. The three target syllables will again be the name syllable, the syllable míng and the syllable liáo.

First, we perform stepwise regression three times by using the four variables of the three target syllables against the value of score (a) ("who"). This is the sentence in which the name syllable is narrowly focused. The following table summarizes the results of the analysis.

<table>
<thead>
<tr>
<th>Target syllable</th>
<th>Pitch range</th>
<th>Top line</th>
<th>Duration</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name (44.0%)</td>
<td>(1)</td>
<td>(3)</td>
<td>(2)</td>
<td>dropped</td>
</tr>
<tr>
<td></td>
<td>34.84%</td>
<td>43.53%</td>
<td>40.14%</td>
<td></td>
</tr>
<tr>
<td>Míng (26.6%)</td>
<td>dropped</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.7%</td>
<td>23.85%</td>
<td>26.28%</td>
</tr>
<tr>
<td>Liáo (18.4%)</td>
<td>(3)</td>
<td>(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.31%</td>
<td>16.68%</td>
<td>dropped</td>
<td>12.72%</td>
</tr>
</tbody>
</table>

In the above table and in all the relevant tables below, the percentage shown immediately below the target syllable stands for the $R^2$ when all the available variables for that
syllable are entered in a regular multiple regression analysis. By comparing this value across the three syllables, it can be observed that the best model for the name syllable accounts for the highest percentage of the variance (44.0% as opposed to 26.6% for the syllable míng and 18.4% for the syllable liáo). This indicates that the name syllable is the most important if only one syllable is chosen to predict score (a). In this case, "who" represents focus 1, and therefore, the acoustic parameters of the name syllable contribute the most in predicting subjects' judgment for this focus pattern.

For the name syllable, pitch range is more important than duration. Intensity is not significant at all. Here the numbers between parentheses represent the relative order in which a variable is chosen to enter the model. The \( R^2 \) below the number represents the proportion of data that the model is able to account for when the associated variable is admitted into the model. For example, for this syllable, when only pitch range is entered, \( R^2 \) is 34.84%. This parameter is chosen first because it has the highest F or t-value among all the three variables. After pitch range, duration is chosen next. The \( R^2 \) value of 40.14% stands for the percent of the data the model is able to account for when both pitch range and duration are entered for the analysis. Next comes the top line of the syllable. Thus, in predicting
the focus 1 pattern, the top line of the name syllable is not very important. Rather, the pitch range of the name syllable is the most important. Intensity is dropped because it does not add much to the prediction of the data. Note, for the syllable míng, top line is the most important. This indicates that when the name syllable is stressed to represent a focus 1 pattern, F0 value at the top line for the syllable míng is the most important acoustic parameter for that syllable. For the syllable liáo, however, intensity appears to be the most important.

Now, we examine the results of analysis for score (b) (focus 2 that corresponds to choice "when") and score (c) (focus 3 that corresponds to choice "do what"). These are presented in the following two tables.

Table 6.2 Stepwise regression of the four variables of each target syllable against the score "when"

<table>
<thead>
<tr>
<th>Target syllable</th>
<th>Pitch range</th>
<th>Top line</th>
<th>Duration</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name (5.6%)</td>
<td>(1) 3.14%</td>
<td>dropped</td>
<td>(2) 5.21%</td>
<td>dropped</td>
</tr>
<tr>
<td>Ming (49.6%)</td>
<td>(1) 43.53%</td>
<td>dropped</td>
<td>(2) 48.94%</td>
<td>dropped</td>
</tr>
<tr>
<td>Liáo (18.9%)</td>
<td>(1) 17.66%</td>
<td>dropped</td>
<td>dropped</td>
<td>dropped</td>
</tr>
</tbody>
</table>
Table 6.3 Stepwise regression of the four variables of each target syllable against the score "do what"

<table>
<thead>
<tr>
<th>Target syllable</th>
<th>Pitch range</th>
<th>Top line</th>
<th>Duration</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name (25.6%)</td>
<td>dropped</td>
<td>(2)</td>
<td>(1)</td>
<td>(3)</td>
</tr>
<tr>
<td>Míng (29.4%)</td>
<td>(3)</td>
<td>(2)</td>
<td>(1)</td>
<td>(4)</td>
</tr>
<tr>
<td>Liáo (60.7%)</td>
<td>(1)</td>
<td>(3)</td>
<td>dropped</td>
<td>(2)</td>
</tr>
</tbody>
</table>

The above two tables both show that the stressed syllables are the most important in accounting for the perception data. In both cases, the $R^2$ of the best model for the stressed syllable is by far the largest. For focus 2 sentences, the stressed syllable is míng and, for focus 3, the corresponding syllable is liáo. When these syllables are stressed, the tables show that in both cases the most important parameter is the pitch range since this is the first variable considered by the stepwise regression procedure. We have observed above that pitch range is also the most important parameter of the stressed name syllable in the model of regression against the score (a) ("who"). Thus, it appears that when a syllable is stressed, the pitch range of the stressed syllable is the most important acoustic parameter for the corresponding focus pattern.
It can be observed in the above three tables that F0 parameters (pitch range and top line) are also the most important for the unstressed syllables if they immediately follows a stressed one. For example, when the name syllable is stressed (Table 6.1), the top line of the following syllable míng is selected first. On the other hand, the same parameters for a pre-stress syllable are not significant. In Table 6.2, the stressed syllable is míng. Although the pitch range of the preceding name syllable is still considered the most important, the data that this variable is able to account for are almost negligible. It can be observed that the R² in this case is only 3.14%, which is much smaller than the other cases where the same variable is ranked first.

Comparing the behavior of the unstressed syllables in the above three tables, we also find that, if they immediately follow a stressed syllable, they will have higher R² values, and are thus, relatively more important. For example, when the name syllable is stressed, syllable míng is more important than the syllable liáo. In Table 6.1, it can be observed that the R² for the syllable míng is 26.6%, which is higher than the corresponding value of 18.4% for the syllable liáo. When the syllable míng is stressed, the following liáo is more important than the preceding name syllable. So, in Table 6.2, the R² for the syllable liáo is higher than that of the name syllable.
We have observed in Chapter Four that the most salient F0 feature of a stressed syllable is the pitch range expansion of the stressed syllable together with a drastic reduction of the top line for the following syllables. This observation can naturally explain the reason why our data show that post-focus F0 is more sensitive to perception than is the pre-focus one.

Now we discuss the role of duration in the above three tables. It can be observed that basically it is the next important parameter after F0. First, when a syllable is stressed, duration is generally more important than intensity. For example, in both Tables 6.1 and 6.2, intensity is dropped for the stressed syllables. Second, in the pre-focus position, duration is also more important than intensity. In such a position, F0 is rather neutral in the sense that there is no noticeable rising or falling. Thus, it is not surprising that its distinctive function should get weakened to its minimum level. Of the other two candidates, duration and intensity, if they are equally important, the chance for them to be selected as the next variable should be the same. In Table 6.3 where the syllable liáo is stressed, duration is the most important parameter for both the syllable mín and the name syllable. In Table 6.2 where the syllable míng is stressed, for the pre-focus name syllable, duration is also more important than intensity. These
results all suggest that intensity comes after duration in terms of the relative importance for the perception data. Besides, in many cases, intensity is simply dropped from the best model.

Now we examine the syllable liáo. For this particular syllable, the parameter intensity plays an important role. First, when the syllable is stressed (Table 6.3), it ranks higher than duration since it is chosen second and duration is dropped. Second, when the name syllable is stressed (Table 6.1), intensity is the first parameter chosen for the syllable liáo. Only when this syllable immediately follows the stressed míng do we find that intensity is not important. This observation might be related to the fact that the target syllable is located towards the end of the sentence. Sentence-final lengthening is a well-known universal phenomenon for languages. Although the parameter duration for this syllable shows a main effect of stress (see Table 5.6), its efficacy as a perceptual cue might get somewhat weakened for its position at the end of the sentence. Intensity for this syllable, on the other hand, is different. It is expected that the intensity will fall towards the end of the sentence. As a matter of fact, this is part of the reason why the syllable liáo rather than the final yānɡ was chosen as the target syllable for the acoustic measurement. It was found that, for many tokens, the low intensity for
this syllable triggers breathy voice on the part of the subject with the result that F0 traces were not registered at all. This being the case, when the syllable liáo is stressed, the effect of intensity boost might well surpass the corresponding increase of duration.

Finally, we discuss score (d) ("what happened"). This response category corresponds to focus 4 (broad focus), which has a very peculiar status. For the other three focus patterns, a target of the stressed syllable can be identified. However, for this focus pattern, we are not absolutely sure which syllable is actually stressed.

In Chapter Three, we find that there exists profound confusion on the part of the subject in correctly differentiating focus 4 from focus 3. In Chapter Four, we find that stress 4 has a very similar general F0 contour pattern as stress 3. In both cases, the overall F0 excursion is high and flat. These observations provide strong evidence that the same syllable is stressed under both stress patterns.

However, after comparing the acoustic parameters of the two stress patterns in Chapters Four and Five, we also find that the two stresses are not exactly the same. First, under stress 3, the pitch range values for the syllable liáo are significantly greater for three subjects (Table 4.8). Second, the duration under stress 3 also tends to become
longer (Table 5.7). Third, intensity under stress 3 is also found to be higher for some subjects (Table 5.16). The evidence seems to suggest that, under stress 4, the syllable liáó is not stressed in the same way as it is under stress 3.

On the other hand, there is other evidence which shows that, under stress 4, the name syllable is also affected. First, its pitch range is relatively greater for most subjects under stress 4 (Table 4.9). Second, the duration of the name syllable tends to be slightly longer under stress 4 (Table 5.8). We have suggested that information focus may be used to explain the difference. Under focus 4, no presupposition is assumed on the part of the speaker. So, the information load for the name syllable is relatively heavier. However, under focus 3, the syllable liáó carries more information since it is narrowly focused.

From this viewpoint and from the general observation that, in narrow focus, stress always falls on the focused word, it seems to be logical to make three hypotheses about the location of the stressed word under broad focus (focus 4): 1. the syllable liáó is stressed; 2. the name syllable is stressed; 3. both syllables are stressed, but to different degrees.¹

The choices can be made relatively easy when we consider the results of the multiple regression analysis. In the following table, I first summarize the results of the
multiple regression analysis for the four parameters of the three target syllables against the score (d).

Table 6.4 Stepwise regression of the three parameters of each target syllable against the score "what happened"

<table>
<thead>
<tr>
<th>Target syllable</th>
<th>Pitch range</th>
<th>Top line</th>
<th>Duration</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name (16.9%)</td>
<td>(2)</td>
<td>dropped</td>
<td>(1)</td>
<td>dropped</td>
</tr>
<tr>
<td></td>
<td>16.69%</td>
<td></td>
<td>13.56%</td>
<td></td>
</tr>
<tr>
<td>Míng (23.3%)</td>
<td>(2)</td>
<td>(3)</td>
<td>(1)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>18.7%</td>
<td>20.25%</td>
<td>14.60%</td>
<td>23.32%</td>
</tr>
<tr>
<td>Liáo (20.2%)</td>
<td>dropped</td>
<td>(1)</td>
<td>(2)</td>
<td>dropped</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.22%</td>
<td>19.87%</td>
<td></td>
</tr>
</tbody>
</table>

We first examine the $R^2$ for the three target syllables. It can be observed that these syllables have very close $R^2$ values. Unlike the other three cases in which the best model has a distinctively higher $R^2$ value, none of the three syllables in the above table stands out alone. This fact is related to the focus pattern this stress condition stands for. We have already noted that, under broad focus, all elements are assumed to be new information by the speaker. This focus condition is realized by the stress 4 pattern in which no sharp phonetic contrast is found. This can explain why all the three target syllables under focus 4 have roughly the same $R^2$ value. Thus, the $R^2$ value of the individual
target syllable alone does not provide any clue as to which syllable is stressed.

However, if we consider the relationship among the acoustic parameters, we begin to see differences among the syllables. Although intensity is the least important for all the three syllables, duration is the first chosen for both the name syllable and the syllable míng, whereas for the syllable liáo, the F0 parameter top line is the most important.

We have noted above that, when a syllable is stressed for narrow focus, pitch range is the most important cue. In all cases, it is the first parameter to enter the model. Duration, on the other hand, is either dropped (for the syllable liáo) or selected after the pitch range (the name syllable and the syllable míng). Thus, in terms of relatively importance, duration obviously ranks lower than the F0 parameter. For the focus 4 pattern, we find in Table 6.4 that top line is the first parameter chosen for the syllable liáo. However, for the other two syllables, duration is. Although pitch range is dropped for the syllable liáo, top line is still an F0 parameter. It thus seems appropriate to pick liáo to be the stressed syllable under broad focus. But the two stress conditions (broad versus narrow) may not be exactly the same since pitch range expansion is not the primary cue for broad focus (focus 4).
However, in terms of the order in which an F0 parameter is chosen and together with other perceptual and acoustic evidence discussed in previous chapters, the syllable liáo appears to be the best candidate for the possible location of stress that represents broad focus.

On the other hand, since the name syllable under stress 4 also shows acoustic evidence of stress, it is likely that, under this stress condition, the degree of stress the syllable receives is heavier than the one under stress 3. Note that this is also supported by the perceptual evidence presented in Table 3.1. It is found that, under the focus 4 pattern, a considerable proportion of the score goes to the choice of "who", a focus condition in which the name syllable is stressed.

Thus, it seems likely that, under stress 4, both the name syllable and the syllable liáo are stressed. However, the primary stress falls on the syllable liáo. In terms of the degree of stress, we can say that the stress pattern is 231, that is, the name syllable is assigned degree 2, the syllable liáo degree 1.

To summarize, in this section, we have analyzed the results of the multiple regression analysis with individual perceptual scores as dependent variables and four acoustic parameters of a single target syllable as the independent variables. We find that, for narrow focus, the best model
for the stressed syllable always yields the highest $R^2$ value. This indicates that the narrowly focused syllable possesses the most salient acoustic features among all the three target syllables. Of the four acoustic parameters, we find that pitch range is the most important cue for identifying narrow focus since it is always the first parameter to be chosen by the multiple regression procedure. Duration is the second most important and intensity ranks the lowest on this hierarchy. In many cases, the parameter intensity is simply dropped. We have also discussed the possible locations of stress for broad focus. Our conclusion is that the primary stress for this focus pattern falls on the syllable lǐāo. At the same time, the name syllable also receives stress, but to a lesser degree. The evidence comes from the relative order the acoustic parameter is chosen in the multiple regression analysis and from the acoustic analysis presented in the last two chapters as well as the results of the perception experiment in Chapter Three.

5.3 Stepwise regression with twelve variables

In the last section, the regression model only includes variables from the same syllable. The relative importance of the three target syllables can only be evaluated in the light of the overall $R^2$ which is the result of regressing all parameters of these syllables against the dependent variable.
In this section, we will use all the twelve parameters from the three syllables as the independent variables of the regression model. Again, stepwise regression procedure will be used to determine the relative importance of all the parameters across different syllables. In the following table we first present a summary of the computer analysis.

Table 6.5 Stepwise regression of twelve parameters of the three target syllables against the four scores

<table>
<thead>
<tr>
<th>Variables</th>
<th>Focus 1 &quot;who&quot; (78.3%)</th>
<th>Focus 2 &quot;when&quot; (79.9%)</th>
<th>Focus 3 &quot;do what&quot; (72.4%)</th>
<th>Focus 4 &quot;what&quot; (48.6%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pitch range</td>
<td>(1) 34.84%</td>
<td>(3) 60.52%</td>
<td>dropped</td>
<td>dropped</td>
</tr>
<tr>
<td>top line</td>
<td>(3) 69.76%</td>
<td>dropped</td>
<td>(4) 66.38%</td>
<td>(3) 40.72%</td>
</tr>
<tr>
<td>duration</td>
<td>(7) 77.24%</td>
<td>dropped</td>
<td>(2) 47.53%</td>
<td></td>
</tr>
<tr>
<td>intensity</td>
<td>dropped</td>
<td>dropped</td>
<td>dropped</td>
<td>dropped</td>
</tr>
<tr>
<td>mǐng</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pitch range</td>
<td>(5) 76.23%</td>
<td>(1) 43.53%</td>
<td>dropped</td>
<td>dropped</td>
</tr>
<tr>
<td>top line</td>
<td>(2) 63.47%</td>
<td>(4) 69.28%</td>
<td>(5) 46.73%</td>
<td></td>
</tr>
<tr>
<td>duration</td>
<td>dropped</td>
<td>77.57%</td>
<td>dropped</td>
<td>14.6%</td>
</tr>
<tr>
<td>intensity</td>
<td>dropped</td>
<td>dropped</td>
<td>dropped</td>
<td>dropped</td>
</tr>
<tr>
<td>liǎo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pitch range</td>
<td>(6) 76.87%</td>
<td>(2) 53.92%</td>
<td>(1) 45.34%</td>
<td></td>
</tr>
<tr>
<td>top line</td>
<td>(4) 74.49%</td>
<td>(5) 73.37%</td>
<td>(3) 62.66%</td>
<td>(2) 30.09%</td>
</tr>
<tr>
<td>duration</td>
<td>dropped</td>
<td>dropped</td>
<td>70.26%</td>
<td>dropped</td>
</tr>
<tr>
<td>intensity</td>
<td>dropped</td>
<td>dropped</td>
<td>68.63%</td>
<td>dropped</td>
</tr>
</tbody>
</table>

*A previously chosen parameter is dropped when this is added to the list.
In the above table the percentage located immediately below the response category of each column stands for the total $R^2$ when all the twelve parameters of the three target syllables are used in the regression model. The first column ("who") corresponds to focus 1 and the second column ("when") to focus 2 and so on.

It can be observed that, for the first three focus patterns, the $R^2$ values for the best model are pretty high. We will examine those first since they form one separate group.

The first important observation about these focus patterns is that, for every response category, the first parameter selected by the regression procedure is always the pitch range of the stressed syllable. For example, for the category "who" that corresponds to focus 1, the pitch range of the name syllable is the first chosen. For "when", the corresponding parameter for the syllable míng is the first to enter the model. The same is true for the syllable liáo. This confirms our conclusion in the last section that pitch range is basically the primary cue for stress.

A second observation is that post-stress top line is also important for the perception. It can be noted that this parameter is the most important for the syllables míng and liáo when their preceding syllable is stressed. In the case of the syllable míng, when the name syllable is stressed, the
top line of the syllable turns out to be the second most important after pitch range of the stressed name syllable. For the syllable liáo in the same stress condition, top line is also more important than pitch range. When the syllable míng is stressed (focus 2), top line is the only parameter chosen for the syllable liáo. Here, it is interesting to observe that pitch range is actually chosen before the parameter top line. However, when duration is introduced into the model, the pitch range parameter is dropped. This shows that when the two most important parameters (F0 and duration) are both present for model fitting, the pitch range parameter for the syllable liáo becomes less important. It was shown in Chapter four that, when a syllable is stressed, the pitch range expansion is followed by the top line reduction of the following syllable. This claim is further supported by the results of the multiple regression analysis here.

As to the relative importance of the four acoustic parameters analyzed, it is clear that pitch range expansion of the stressed syllable is the most important cue for the stressed syllable. Top line reduction of the following syllable is the second most important for stress 1 and less so for stress 2. However, since both parameters belong to the same F0 category, we can say that by far F0 is more important than duration. For all the three focus patterns
considered here, the parameter duration of the stressed syllable is actually the last chosen after no other F0 parameters of the unstressed syllable are selected. At the same time, the duration of an unstressed syllable is not selected at all. However, duration is still more important than intensity since the latter fails to enter the model even when the target syllable itself is stressed. It can be observed that, in both models that best fit focus 1 and focus 2, no intensity parameters for any of the three syllables are chosen. It is selected by the model only in fitting the variations of the perceptual responses to the focus 3 pattern. When the syllable liáo is stressed, we find that intensity is chosen, but still after the parameters of pitch range and top line of the same syllable have been chosen. Although duration of that particular syllable is ranked somewhat lower, the generalization remains true that intensity is the least important among all the three acoustic parameters for stress, that is, F0, syllable duration and intensity.

Now we consider focus 4. It can be observed that the best fit model can only account for 48.6% of the data. This again suggests that this category is different from the other three focus patterns. A large proportion of the responses to this focus pattern is simply random, and so, have negatively influenced the fitting.
It was concluded in the last section that, for this focus pattern, both the name syllable and the syllable liáó are stressed with primary stress falling on the latter. The results of the regression analysis presented in Table 6.5 further supports the hypothesis. First, for the syllable liáó, both F0 parameters are chosen. Second, the first selected F0 parameter also belongs to the syllable liáó.

Comparing the "relative weights" of the parameters chosen for the name syllable and the syllable liáó, we can say that liáó is more important. First, the same parameter top line of the syllable liáó is chosen before the corresponding parameter of the name syllable. Second, for the syllable liáó, both F0 parameters are chosen, whereas for the name syllable, only one is chosen. These form the evidence that supports the hypothesis that the stress pattern for broad focus in Mandarin is 231.

To summarize, we have considered the results of the regression analysis that uses all twelve parameters from the three target syllables for the model fitting. The results strongly indicate that F0 parameter is the most important parameter for indicating stress. When a syllable is stressed for narrow focus, the primary cue is always the pitch range expansion of the stressed syllable. When the name syllable is stressed, the top line reduction of the following syllable míng is also very important. Duration is not so crucial for
indicating stress in our data. But, comparatively speaking, it is still more significant than intensity. The results also suggest that focus 4 is similar to focus 3 in that both of them have the corresponding stress falling on the last target syllable. However, for focus 4, the name syllable is also stressed, but to a lesser degree.

6.4 Conclusion

In this chapter, we have used the stepwise multiple regression procedure to analyze the three acoustic parameters for the target syllables. The general conclusion is that the parameter F0 is the most important for indicating stress. When a syllable is stressed, pitch range expansion and the top line reduction of the following syllable are more important than the other two parameters of the stressed syllable. Between duration and intensity, we find that duration is more important.

Based on the results of the multiple regression analysis and those in previous chapters, we have reached the conclusion about the difference between stress 3 and stress 4. Basically, for both stress conditions, the syllable liáo is stressed. Under broad focus, the name syllable is also stressed. However, the primary stress falls on the syllable liáo. The name syllable only receives secondary stress.
Note

Note that we cannot possibly hypothesize that the syllable mǐng is stressed under broad focus even though its information status is similar to the other two target syllables. First, the perception experiment shows that focus 2 is the least confused with focus 4 (Table 3.1). Second, none of the three acoustic parameters for the syllable mǐng are found to differ significantly when stress 3 is compared with stress 4.
In this study, we have explored in some detail the acoustic correlates of sentence stress in Mandarin Chinese and their relationship to perceived focus. Three acoustic parameters, F0, syllable duration and intensity, have been examined with reference to stress. Generally speaking, F0 and duration, but not intensity, are related to stress. This finding conforms to most studies in the past. The perceptual studies basically confirm and, in some cases, provide further insights into the phenomenon of stress in this language.

Acoustically, F0 is an important parameter for indicating stress. When a syllable is stressed for narrow focus, its pitch range will expand dramatically. For the name syllable and the syllable míng, this range expansion is also followed by a drastic fall of F0 contour overall. However, the inherent tone of the following syllable is not in every case overridden by the falling contour as is claimed by previous studies (Liao 1994, Tseng 1981). Our data show that the post-stress F0 excursion, to a large extent, is
determined by the inherent tone of the preceding stressed syllable. An inherent T2 following a stressed T3 syllable typically exhibits a rising contour. On the other hand, the post-stress rising tendency is much less when the preceding stressed syllable is T4. However, even in this case, an obvious break for the falling tendency in the form of a flat plateau is usually observed. Only when the stressed syllable is inherently T1 or T2 do we find cases in which the post-stress rising T2 contour is completely overridden by the falling intonation.

In Chinese linguistics, it is often claimed that sentence intonation not only reinforces the falling or rising pattern inherent to the tone of a syllable, but can also partly cancel or completely override the inherent tonal contour of the syllable. When the local tonal excursion glides in the same direction as the global intonation, reinforcement occurs. On the other hand, when the direction of tonal movement differs from the sentence intonation, partial or complete neutralization takes place (Chao 1968, Ho 1976a, 1977, Liao 1994, Shen 1990a, Tseng 1981).

It can be observed that the claim is not quite supported by our post-stress data. Since all our sentences are statements, it is reasonable to assume that they all have the same falling intonation. When the name syllable is stressed, the inherent tone of the stressed syllable takes precedence.
T1 and T2 are both high tones and non-falling. When they are stressed, one would expect that they will positively reinforce the rising contour of the following T2 syllable in the sense that the post-stress syllable would somehow preserve its inherent rising contour in the sentence environment. On the other hand, the local contour of such a syllable would be more subject to the falling tendency if the preceding stressed syllable is a high-falling tone (T4). In the same way, T3 is basically falling. Consequently, when this syllable is stressed, we should expect that it will negatively influence the rising tendency of the following syllable. However, our data show that, in the post-stress position, a rising tone following a stressed high-rising or high-leveling tone generally exhibits a falling contour, while the same tone following a stressed high-falling tone typically shows a flat plateau. In the case where the preceding stressed syllable is T3, the inherent rising contour of the following syllable is always completely reversed as a local rise in the sentence intonation.

Such a situation can be accounted for nicely if we assume a pitch range reduction model. F0 excursion of a post-stress syllable in such a model is determined by tonal coarticulation with the preceding stressed syllable and the reduced top line. When a syllable is stressed, its pitch range expansion is immediately followed by a top line
reduction of the following syllable. This reduced top line basically determines the high target of that syllable. The beginning point of the F0 excursion of the post-stress syllable, however, is conditioned by the tonal contour of the preceding stressed syllable. If that syllable ends in a high tone, the post-stress F0 excursion will begin on a high level, and in the case of a rising T2, will glide toward the top line of that syllable. However, because of the preceding stress condition, the top line is now located on a much lower level. If F0 directly glides toward that level because of tonal coarticulation, we see a falling contour. On the other hand, if the preceding stressed syllable ends on a lower level (e.g. T3 or T4), the post-stress F0 will begin its glide on a low level. In the case of a preceding stressed T3, that level will be lower than the reduced top line. This results in a rising F0 excursion if the post-stress syllable is T2 as is the case in this experiment. However, since this experiment only examines the T2 context in the post-stress position, further experiments are needed to investigate if the behavior of other tones can be explained by the model described here.

Besides F0, the results of our study also indicate that duration is another important acoustic parameter for stress. A syllable will typically become longer when it is stressed. This effect is true for every target syllable. However, the
same observation cannot be found for intensity. We find that this parameter is generally not a reliable cue for stress. For both the name syllable and the syllable míng, no significant stress effect can be found that forms a general pattern across subjects. Only at sentence-final position do we find intensity related to stress.

In this study, broad focus is also investigated. Both the results of our perception experiment and the acoustic analysis show that this focus pattern is closely related to focus 3 in which the last target syllable is stressed. The perception experiment shows profound confusion on the part of the subjects between focus 4 and focus 3. Our acoustic analysis indicates, however, that the two focus conditions are slightly different. For broad focus, the name syllable receives heavier stress, while for focus 3, the syllable liáo is more heavily stressed. Based on these results and the multiple regression analysis, we conclude that, for broad focus, primary stress falls on the syllable liáo and secondary stress on the name syllable. However, the degree of prominence is somewhat less than the corresponding narrow focus. In both cases, the general F0 contour is high and flat.

To analyze the correlation between the subjective judgment of the focus patterns in the perception experiment and the objective acoustic signals obtained from the
production experiment, we have performed a set of stepwise multiple regression analysis on four acoustic parameters – pitch range, top line, duration and intensity – of the three target syllables against the four perception scores. The results clearly show that the two F0 parameters are by far the most important for indicating stress. Whenever a syllable is stressed for narrow focus, its pitch range parameter is always selected first by the regression procedure. In the mean time, the top line parameter of the following syllable is also the most important for that syllable. This result statistically indicates that pitch range expansion as well as top line reduction are the most important cue for stress. Duration is related to stress on a much less significant level. It usually enters the best fit model after the F0 parameters. Lastly, intensity is found to be the least important for indicating stress. In most cases, it fails to be selected by the regression model even when the syllable is stressed.
Appendices
### Appendix A

Mean value of pitch ranges of the name syllables with respect to subject, tone, and stress (N = 5)

<table>
<thead>
<tr>
<th>Subject &amp; Stress</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1  SBJ1</td>
<td>98.48 (9.04)</td>
<td>86.96 (10.79)</td>
<td>110.48 (6.49)</td>
<td>88.04 (8.40)</td>
</tr>
<tr>
<td>S2  SBJ1</td>
<td>57.64 (10.50)</td>
<td>49.08 (8.67)</td>
<td>61.20 (14.49)</td>
<td>39.86 (21.35)</td>
</tr>
<tr>
<td>S3  SBJ1</td>
<td>54.24 (13.01)</td>
<td>61.72 (8.81)</td>
<td>45.10 (7.69)</td>
<td>16.74 (12.18)</td>
</tr>
<tr>
<td>S4  SBJ1</td>
<td>77.98 (8.46)</td>
<td>84.28 (4.58)</td>
<td>64.12 (18.18)</td>
<td>24.94 (16.38)</td>
</tr>
<tr>
<td>S1  SBJ2</td>
<td>42.66 (5.39)</td>
<td>41.40 (7.20)</td>
<td>37.64 (11.72)</td>
<td>49.00 (11.11)</td>
</tr>
<tr>
<td>S2  SBJ2</td>
<td>28.90 (2.38)</td>
<td>25.08 (4.68)</td>
<td>22.02 (5.24)</td>
<td>19.98 (7.86)</td>
</tr>
<tr>
<td>S3  SBJ2</td>
<td>28.40 (7.33)</td>
<td>28.90 (2.97)</td>
<td>27.88 (3.49)</td>
<td>15.28 (3.13)</td>
</tr>
<tr>
<td>S4  SBJ2</td>
<td>33.36 (4.45)</td>
<td>34.80 (4.23)</td>
<td>33.12 (6.10)</td>
<td>22.42 (1.04)</td>
</tr>
<tr>
<td>S1  SBJ3</td>
<td>95.68 (6.00)</td>
<td>89.36 (9.17)</td>
<td>137.9 (23.0)</td>
<td>98.90 (10.40)</td>
</tr>
<tr>
<td>S2  SBJ3</td>
<td>62.40 (6.86)</td>
<td>55.90 (2.30)</td>
<td>79.30 (6.97)</td>
<td>48.28 (6.64)</td>
</tr>
<tr>
<td>S3  SBJ3</td>
<td>57.12 (4.48)</td>
<td>56.62 (3.79)</td>
<td>66.88 (4.64)</td>
<td>32.24 (7.79)</td>
</tr>
<tr>
<td>S4  SBJ3</td>
<td>62.66 (5.02)</td>
<td>54.64 (3.56)</td>
<td>79.74 (14.19)</td>
<td>34.14 (5.83)</td>
</tr>
<tr>
<td>S1  SBJ4</td>
<td>76.78 (11.21)</td>
<td>77.00 (12.04)</td>
<td>147.5 (30.4)</td>
<td>101.86 (15.59)</td>
</tr>
<tr>
<td>S2  SBJ4</td>
<td>63.76 (4.26)</td>
<td>55.68 (5.06)</td>
<td>74.76 (10.12)</td>
<td>77.76 (10.16)</td>
</tr>
<tr>
<td>S3  SBJ4</td>
<td>61.86 (6.67)</td>
<td>55.12 (6.70)</td>
<td>77.80 (10.29)</td>
<td>63.10 (13.67)</td>
</tr>
<tr>
<td>S4  SBJ4</td>
<td>66.86 (5.75)</td>
<td>67.56 (4.67)</td>
<td>93.34 (12.41)</td>
<td>87.20 (7.10)</td>
</tr>
</tbody>
</table>
Appendix B

Mean values of pitch ranges of the syllable *mǐng* with respect to subject, stress, and tonal context (N = 5)

<table>
<thead>
<tr>
<th>Subject &amp; Stress</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 52.7 (24.8)</td>
<td>10.30 (9.49)</td>
<td>5.28 (7.40)</td>
<td>60.70 (9.37)</td>
<td>0.720 (0.99)</td>
</tr>
<tr>
<td>S2 26.44 (20.16)</td>
<td>17.06 (6.53)</td>
<td>34.96 (7.94)</td>
<td>19.40 (5.11)</td>
<td></td>
</tr>
<tr>
<td>S3 15.26 (11.55)</td>
<td>3.58 (8.01)</td>
<td>48.28 (6.56)</td>
<td>14.98 (10.56)</td>
<td></td>
</tr>
<tr>
<td>S4 6.760 (1.89)</td>
<td>3.480 (2.10)</td>
<td>19.08 (3.09)</td>
<td>13.90 (3.08)</td>
<td></td>
</tr>
<tr>
<td>S5 46.24 (4.37)</td>
<td>54.14 (11.82)</td>
<td>67.16 (4.45)</td>
<td>56.88 (11.42)</td>
<td></td>
</tr>
<tr>
<td>S6 52.00 (6.67)</td>
<td>51.16 (6.45)</td>
<td>57.22 (13.97)</td>
<td>52.16 (3.01)</td>
<td></td>
</tr>
<tr>
<td>S7 7.12 (7.37)</td>
<td>6.88 (10.48)</td>
<td>47.66 (9.96)</td>
<td>34.34 (2.92)</td>
<td></td>
</tr>
<tr>
<td>S8 9.38 (9.31)</td>
<td>7.26 (5.55)</td>
<td>47.76 (13.61)</td>
<td>46.92 (8.82)</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C

Mean values (Hz) at point (c) under stress 1 and the one-tailed t-test results

Table C.1  Mean values (N = 5)

<table>
<thead>
<tr>
<th></th>
<th>Context T1</th>
<th>Context T2</th>
<th>Context T3</th>
<th>Context T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subj 1</td>
<td>163.74 (12.57)</td>
<td>178.32 (12.70)</td>
<td>113.44 (3.62)</td>
<td>150.68 (10.80)</td>
</tr>
<tr>
<td>Subj 2</td>
<td>157.72 (1.26)</td>
<td>160.84 (5.27)</td>
<td>133.70 (3.30)</td>
<td>138.96 (5.31)</td>
</tr>
<tr>
<td>Subj 3</td>
<td>188.66 (7.33)</td>
<td>212.44 (3.95)</td>
<td>140.20 (8.90)</td>
<td>178.36 (3.83)</td>
</tr>
<tr>
<td>Subj 4</td>
<td>184.44 (4.30)</td>
<td>186.70 (6.52)</td>
<td>171.10 (5.75)</td>
<td>167.70 (6.36)</td>
</tr>
</tbody>
</table>

Table C.2  T-test results

(Insignificant values are highlighted and alternative hypothesis = values in T3 contexts are smaller)

<table>
<thead>
<tr>
<th></th>
<th>T3/T1</th>
<th>T3/T2</th>
<th>T3/T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subj 1</td>
<td>-8.59/0.0005/4*</td>
<td>-10.99/0.0002/4</td>
<td>-7.31/0.0009/4</td>
</tr>
<tr>
<td>Subj 2</td>
<td>-15.2/0.0000/5</td>
<td>-9.76/0.0000/6</td>
<td>-1.88/0.054/6</td>
</tr>
<tr>
<td>Subj 3</td>
<td>-9.40/0.0000/7</td>
<td>-16.59/0.0000/5</td>
<td>-8.80/0.0002/5</td>
</tr>
<tr>
<td>Subj 4</td>
<td>-4.15/0.0022/7</td>
<td>-4.01/0.0026/7</td>
<td>0.89/0.80/7</td>
</tr>
</tbody>
</table>

*t-value/p-value/degree of freedom

214
Appendix D

Mean values (Hz) at point (d) under stress 1 and the one-tailed t-test results

Table D.1 Mean values (N = 5)

<table>
<thead>
<tr>
<th></th>
<th>Context T1</th>
<th>Context T2</th>
<th>Context T3</th>
<th>Context T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subj 1</td>
<td>172.10 (12.64)</td>
<td>177.08 (6.28)</td>
<td>174.14 (6.79)</td>
<td>147.98 (5.51)</td>
</tr>
<tr>
<td>Subj 2</td>
<td>146.36 (5.09)</td>
<td>144.84 (3.79)</td>
<td>147.02 (2.41)</td>
<td>140.92 (6.22)</td>
</tr>
<tr>
<td>Subj 3</td>
<td>158.34 (3.73)</td>
<td>184.24 (4.02)</td>
<td>187.50 (7.88)</td>
<td>175.30 (3.88)</td>
</tr>
<tr>
<td>Subj 4</td>
<td>175.18 (3.99)</td>
<td>182.30 (8.81)</td>
<td>204.88 (8.67)</td>
<td>174.56 (6.69)</td>
</tr>
</tbody>
</table>

Table D.2 T-test results

(Insignificant values are highlighted and alternative hypothesis = values in T3 contexts are greater)

<table>
<thead>
<tr>
<th></th>
<th>T3/T1</th>
<th>T3/T2</th>
<th>T3/T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subj 1</td>
<td>0.32/0.38/6*</td>
<td>-0.71/0.75/7</td>
<td>6.69/0.0001/7</td>
</tr>
<tr>
<td>Subj 2</td>
<td>0.26/0.40/5</td>
<td>1.09/0.16/6</td>
<td>2.04/0.048/5</td>
</tr>
<tr>
<td>Subj 3</td>
<td>4.91/0.0022/5</td>
<td>0.82/0.22/5</td>
<td>3.10/0.013/5</td>
</tr>
<tr>
<td>Subj 4</td>
<td>6.96/0.0005/5</td>
<td>4.08/0.0024/7</td>
<td>6.19/0.0002/7</td>
</tr>
</tbody>
</table>

*t-value/p-value/degree of freedom
Appendix E

Mean values of pitch ranges of the syllable liáo with respect to subject, stress, and tonal context (N = 5)

<table>
<thead>
<tr>
<th>Subject &amp; Stress</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S1</strong></td>
<td>0.40 (0.90)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td><strong>S2</strong></td>
<td>0.00 (0.00)</td>
<td>0.720 (1.61)</td>
<td>0.30 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td><strong>S3</strong></td>
<td>29.48 (4.45)</td>
<td>24.68 (9.00)</td>
<td>33.62 (8.02)</td>
<td>32.50 (12.49)</td>
</tr>
<tr>
<td><strong>S4</strong></td>
<td>14.88 (5.33)</td>
<td>18.12 (11.83)</td>
<td>21.50 (6.19)</td>
<td>10.88 (7.88)</td>
</tr>
<tr>
<td><strong>S1</strong></td>
<td>3.56 (5.12)</td>
<td>6.30 (5.97)</td>
<td>4.50 (4.60)</td>
<td>7.92 (7.32)</td>
</tr>
<tr>
<td><strong>S2</strong></td>
<td>5.62 (6.44)</td>
<td>0.00 (0.00)</td>
<td>7.32 (5.14)</td>
<td>2.10 (4.70)</td>
</tr>
<tr>
<td><strong>S3</strong></td>
<td>10.10 (2.88)</td>
<td>6.84 (3.64)</td>
<td>10.42 (3.27)</td>
<td>10.52 (3.82)</td>
</tr>
<tr>
<td><strong>S4</strong></td>
<td>2.380 (2.21)</td>
<td>1.22 (2.73)</td>
<td>3.900 (2.18)</td>
<td>5.18 (1.83)</td>
</tr>
<tr>
<td><strong>S1</strong></td>
<td>7.98 (4.35)</td>
<td>9.92 (2.37)</td>
<td>1.400 (1.86)</td>
<td>11.22 (5.34)</td>
</tr>
<tr>
<td><strong>S2</strong></td>
<td>3.02 (4.04)</td>
<td>5.44 (4.14)</td>
<td>5.42 (6.41)</td>
<td>2.32 (5.19)</td>
</tr>
<tr>
<td><strong>S3</strong></td>
<td>27.760 (1.07)</td>
<td>26.56 (6.70)</td>
<td>26.16 (2.49)</td>
<td>24.94 (6.45)</td>
</tr>
<tr>
<td><strong>S4</strong></td>
<td>18.42 (4.26)</td>
<td>21.98 (4.34)</td>
<td>17.78 (4.00)</td>
<td>17.620 (1.10)</td>
</tr>
<tr>
<td><strong>S1</strong></td>
<td>6.60 (9.65)</td>
<td>9.78 (9.54)</td>
<td>7.46 (4.33)</td>
<td>3.26 (4.93)</td>
</tr>
<tr>
<td><strong>S2</strong></td>
<td>1.30 (2.91)</td>
<td>5.92 (6.35)</td>
<td>5.86 (7.46)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td><strong>S3</strong></td>
<td>27.94 (6.40)</td>
<td>29.88 (12.11)</td>
<td>19.68 (6.07)</td>
<td>28.18 (9.55)</td>
</tr>
<tr>
<td><strong>S4</strong></td>
<td>30.22 (8.46)</td>
<td>30.76 (9.61)</td>
<td>13.34 (10.66)</td>
<td>26.76 (8.97)</td>
</tr>
</tbody>
</table>
Appendix F

Mean values of duration for the name syllables with respect to subject, tone, and stress (N = 5)

<table>
<thead>
<tr>
<th>Subject &amp; Stress</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 184.6 40.9</td>
<td>171.6 37.4</td>
<td>203.2 47.9</td>
<td>222.60 20.06</td>
<td></td>
</tr>
<tr>
<td>S2 138.6 39.6</td>
<td>133.20 15.97</td>
<td>146.6 24.8</td>
<td>178.0 50.4</td>
<td></td>
</tr>
<tr>
<td>S3 102.4 22.4</td>
<td>109.8 26.2</td>
<td>118.00 4.69</td>
<td>124.60 13.83</td>
<td></td>
</tr>
<tr>
<td>S4 140.2 32.6</td>
<td>118.20 3.70</td>
<td>118.40 13.11</td>
<td>148.4 52.7</td>
<td></td>
</tr>
<tr>
<td>S1 160.6 29.7</td>
<td>151.60 17.97</td>
<td>180.60 8.38</td>
<td>198.80 10.13</td>
<td></td>
</tr>
<tr>
<td>S2 146.80 21.44</td>
<td>155.40 22.15</td>
<td>173.80 15.40</td>
<td>174.20 10.18</td>
<td></td>
</tr>
<tr>
<td>S3 130.80 9.63</td>
<td>134.00 17.04</td>
<td>152.20 8.70</td>
<td>168.00 11.64</td>
<td></td>
</tr>
<tr>
<td>S4 147.20 13.22</td>
<td>145.80 17.85</td>
<td>170.60 17.10</td>
<td>166.00 11.81</td>
<td></td>
</tr>
<tr>
<td>S1 187.80 8.26</td>
<td>190.80 8.70</td>
<td>186.00 7.31</td>
<td>199.20 14.65</td>
<td></td>
</tr>
<tr>
<td>S2 179.00 10.65</td>
<td>162.60 12.86</td>
<td>158.60 5.59</td>
<td>173.80 12.54</td>
<td></td>
</tr>
<tr>
<td>S3 141.80 9.42</td>
<td>142.0 24.3</td>
<td>159.40 5.90</td>
<td>158.20 7.98</td>
<td></td>
</tr>
<tr>
<td>S4 131.80 10.51</td>
<td>132.20 9.65</td>
<td>158.00 5.29</td>
<td>162.60 5.18</td>
<td></td>
</tr>
<tr>
<td>S1 266.00 12.27</td>
<td>273.2 44.5</td>
<td>243.20 20.64</td>
<td>279.00 18.73</td>
<td></td>
</tr>
<tr>
<td>S2 240.00 8.03</td>
<td>219.6 26.8</td>
<td>199.00 11.58</td>
<td>243.2 26.1</td>
<td></td>
</tr>
<tr>
<td>S3 187.40 17.33</td>
<td>156.00 12.59</td>
<td>199.20 14.96</td>
<td>219.60 6.07</td>
<td></td>
</tr>
<tr>
<td>S4 207.20 8.53</td>
<td>183.60 21.38</td>
<td>216.40 17.54</td>
<td>249.40 21.82</td>
<td></td>
</tr>
</tbody>
</table>
Appendix G

Mean values of intensity for the name syllables with respect to subject, tone and stress (N = 5)

<table>
<thead>
<tr>
<th>Subject &amp; Stress</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBJ1 S1</td>
<td>28.54 (7.29)</td>
<td>27.30 (1.77)</td>
<td>23.50 (3.00)</td>
<td>42.94 (5.41)</td>
</tr>
<tr>
<td>S2</td>
<td>24.58 (3.30)</td>
<td>27.60 (4.28)</td>
<td>26.60 (1.88)</td>
<td>32.50 (5.11)</td>
</tr>
<tr>
<td>S3</td>
<td>26.90 (5.83)</td>
<td>25.62 (3.46)</td>
<td>31.38 (7.22)</td>
<td>31.54 (3.32)</td>
</tr>
<tr>
<td>S4</td>
<td>26.02 (4.51)</td>
<td>32.26 (1.65)</td>
<td>24.42 (6.32)</td>
<td>31.34 (4.41)</td>
</tr>
<tr>
<td>SBJ2 S1</td>
<td>21.20 (1.69)</td>
<td>20.00 (4.85)</td>
<td>20.54 (3.28)</td>
<td>27.76 (4.41)</td>
</tr>
<tr>
<td>S2</td>
<td>17.58 (1.50)</td>
<td>16.76 (1.68)</td>
<td>22.92 (4.39)</td>
<td>23.30 (3.02)</td>
</tr>
<tr>
<td>S3</td>
<td>17.96 (3.23)</td>
<td>19.04 (1.54)</td>
<td>22.16 (3.98)</td>
<td>22.64 (3.31)</td>
</tr>
<tr>
<td>S4</td>
<td>19.12 (2.52)</td>
<td>17.86 (2.00)</td>
<td>22.28 (4.00)</td>
<td>26.920 (2.19)</td>
</tr>
<tr>
<td>SBJ3 S1</td>
<td>33.84 (6.46)</td>
<td>31.90 (4.21)</td>
<td>33.30 (4.64)</td>
<td>38.28 (2.27)</td>
</tr>
<tr>
<td>S2</td>
<td>29.92 (3.28)</td>
<td>25.38 (4.47)</td>
<td>25.26 (1.74)</td>
<td>33.46 (7.42)</td>
</tr>
<tr>
<td>S3</td>
<td>26.96 (2.62)</td>
<td>29.58 (3.06)</td>
<td>27.14 (3.97)</td>
<td>34.20 (6.81)</td>
</tr>
<tr>
<td>S4</td>
<td>27.46 (2.64)</td>
<td>23.00 (2.83)</td>
<td>27.74 (2.42)</td>
<td>31.48 (7.02)</td>
</tr>
<tr>
<td>SBJ4 S1</td>
<td>25.26 (4.45)</td>
<td>27.26 (7.07)</td>
<td>34.40 (5.05)</td>
<td>40.94 (7.11)</td>
</tr>
<tr>
<td>S2</td>
<td>20.84 (5.65)</td>
<td>24.76 (8.95)</td>
<td>28.22 (4.93)</td>
<td>31.82 (5.20)</td>
</tr>
<tr>
<td>S3</td>
<td>24.94 (3.10)</td>
<td>24.540 (1.22)</td>
<td>30.70 (8.17)</td>
<td>33.80 (3.37)</td>
</tr>
<tr>
<td>S4</td>
<td>26.34 (5.13)</td>
<td>23.56 (3.22)</td>
<td>32.96 (4.70)</td>
<td>36.76 (3.86)</td>
</tr>
</tbody>
</table>
Appendix H

Mean values of intensity for the syllable mίng with respect to subject, tonal context and stress (N = 5)

<table>
<thead>
<tr>
<th>Subject &amp; Stress</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBJ 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1 18.38 (2.86)</td>
<td>16.82 (3.30)</td>
<td>13.60 (1.82)</td>
<td>15.16 (1.05)</td>
<td></td>
</tr>
<tr>
<td>S2 19.76 (4.53)</td>
<td>21.80 (4.61)</td>
<td>18.68 (2.21)</td>
<td>19.84 (3.12)</td>
<td></td>
</tr>
<tr>
<td>S3 21.90 (3.17)</td>
<td>20.04 (3.73)</td>
<td>17.02 (1.77)</td>
<td>21.32 (2.89)</td>
<td></td>
</tr>
<tr>
<td>S4 20.10 (3.10)</td>
<td>22.84 (3.54)</td>
<td>14.40 (2.25)</td>
<td>18.38 (4.66)</td>
<td></td>
</tr>
<tr>
<td>SBJ 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1 12.98 (1.24)</td>
<td>13.04 (2.03)</td>
<td>7.780 (0.25)</td>
<td>11.12 (1.50)</td>
<td></td>
</tr>
<tr>
<td>S2 13.02 (0.40)</td>
<td>12.00 (0.44)</td>
<td>10.06 (1.58)</td>
<td>11.04 (0.82)</td>
<td></td>
</tr>
<tr>
<td>S3 12.92 (1.05)</td>
<td>13.44 (1.87)</td>
<td>9.760 (0.54)</td>
<td>13.06 (1.20)</td>
<td></td>
</tr>
<tr>
<td>S4 12.86 (0.43)</td>
<td>12.86 (0.93)</td>
<td>8.80 (1.20)</td>
<td>12.86 (1.92)</td>
<td></td>
</tr>
<tr>
<td>SBJ 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1 26.10 (1.35)</td>
<td>25.92 (1.42)</td>
<td>10.60 (1.66)</td>
<td>14.64 (0.81)</td>
<td></td>
</tr>
<tr>
<td>S2 19.82 (1.58)</td>
<td>21.68 (1.50)</td>
<td>13.68 (2.08)</td>
<td>20.42 (1.72)</td>
<td></td>
</tr>
<tr>
<td>S3 22.44 (1.85)</td>
<td>25.14 (1.72)</td>
<td>14.88 (1.31)</td>
<td>21.14 (2.66)</td>
<td></td>
</tr>
<tr>
<td>S4 23.52 (2.22)</td>
<td>23.52 (1.59)</td>
<td>12.90 (0.85)</td>
<td>18.92 (1.87)</td>
<td></td>
</tr>
<tr>
<td>SBJ 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1 11.50 (1.98)</td>
<td>12.32 (3.05)</td>
<td>8.40 (2.91)</td>
<td>7.640 (1.67)</td>
<td></td>
</tr>
<tr>
<td>S2 14.40 (1.16)</td>
<td>16.62 (2.23)</td>
<td>11.02 (1.91)</td>
<td>11.04 (1.76)</td>
<td></td>
</tr>
<tr>
<td>S3 16.84 (2.28)</td>
<td>18.88 (1.97)</td>
<td>13.64 (1.79)</td>
<td>12.96 (1.40)</td>
<td></td>
</tr>
<tr>
<td>S4 19.06 (5.45)</td>
<td>20.22 (4.06)</td>
<td>15.50 (1.94)</td>
<td>12.86 (1.28)</td>
<td></td>
</tr>
</tbody>
</table>
Bibliography


Chinese Linguistics (NACCL 4). University of Michigan, Ann Arbor.


Inkelas, Sharon and Draga Zec (eds.). 1990. The Phonology-Syntax Connection. Chicago University Press.


