

Categorical Perception of Lexical Tones: ERP and Behavioral Experiments

ZHENG, Hong-Ying

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in

Electronic Engineering

© The Chinese University of Hong Kong

January 2010

The Chinese University of Hong Kong holds the copyright of this thesis. Any person(s) intending to use a part or whole of the materials in the thesis in a proposed publication must seek copyright release from the Dean of the Graduation School.

To my parents,

Tianhe ZHENG and Meiyu CHEN,

To my husband,

Jianyong, CHEN

for their love, considerations, and supports.

ACKNOWLEDGEMENTS

This thesis is a synthesis of the research during my PhD study at both City University of Hong Kong (CityU) and The Chinese University of Hong Kong (CUHK). The most important person in my academic life is my supervisor, Prof. William ShiYuan WANG. All the words in the world do not suffice to thank him. His special wisdom led me into the beautiful realm of linguistics. His comprehensive lectures on speech perception inspired me to choose my topic on perception on lexical tones. He has guided me to learn how to appreciate the beauty and cherish the joy of doing research in basic science, and his encouragement has been all along supporting me to overcome various difficulties in research. He encouraged me to go into the new, exciting area of neuroscience. Without his recommendation and support, I would never have the chance to attend several international conferences and summer schools during the past four years or have the chance to visit the University of California, Berkeley during the summer in 2008.

I would like to extend my gratitude to Prof. Peter WM TSANG, who acted as my co-supervisor during my study in CityU, and sincerely supported me to continue my research when Prof. WANG leaves CityU. Especial thanks should be given to Prof. Richard IVRY and Prof. Paul KAY, who kindly hosted me when I visited the University of California, Berkeley. The working experience with them enriched my knowledge about psychology. More importantly, they helped me to raise the ability to do research work scientifically, critically, and creatively. I also want to thank Prof. Thomas Hun-Tak LEE, whose great knowledge about language acquisition and critical comments on my work are valuable for me to reconsider the experimental design.

I own a lot to the technical support and moral encouragement given by all the members of LEL. It was my great honor to collaborate with Dr. James W. MINETT and Dr. Gang PENG. Their rich experiences and critical challenge greatly inspired me to fulfill my research. I am grateful to Dr. Jin-Yun KE, Dr. Tao GONG, Dr. Francis Chunkit WONG, and Ms. Yao YAO to help me to improve my awkward English and to reorganize my paper. In addition, I am especially grateful to Dr. Tao GONG, Dr. Ching-Pong AU and Ms. Ruo-Xiao YANG who kindly helped me to collect some field work data presented in this thesis. Furthermore, I enjoyed a lot discussion with my other colleagues, Dr. Susan Lan SHUAI, Mr. Yingwei WONG, Mr. Raymond Weiwen WU, Mr. Guang-Ting MAI, Mr. Manson

FONG, Ms. Lin ZHOU, and Ms. Hui CHEN on various research topics. I am very fortunate to have these close amigos in my life.

In the end, I would like to thank the members in a bigger family, the Digital Signal Processing and Speech Technology Laboratory (DSPSTL). The leaders of this laboratory, Prof. Pak-Chung CHING, Prof. Tan LEE, and Prof. Ken Wing-Kin MA have generously supported me to perform my basic science research in this engineering laboratory. And I enjoyed working, growing, and improving my skills with the colleagues in DSPSTL, Dr. Meng YUAN, Ms. Yujia LI, Dr. Nengheng ZHENG, Ms. Ning WANG, Dr. Feng TIAN, Ms. Houwei CAO, Mr. Jiaxian PAN, Mr. Li QIANG, Dr. Yvone Siuwa LI, and Ms. Natalie Li ZENG. Without these friends, my PhD life in CUHK would have had lost many colors.

CONTENTS

| | |
|---|-------|
| ACKNOWLEDGEMENTS | III |
| CONTENTS..... | V |
| LIST OF TABLES | XI |
| LIST OF FIGURES | XIII |
| ABSTRACT..... | XVIII |
| 摘要..... | XX |
| CHAPTER 1 INTRODUCTION | 1 |
| 1.1. Categorical Perception and Its Importance | 1 |
| 1.2. Lexical Tone and Its CP Studies | 3 |
| 1.3. Event-Related-Potential and Its Applications in CP Studies | 6 |
| 1.4. Organization of the Thesis | 8 |
| CHAPTER 2 BACKGROUND | 1 |
| 2.1. Mechanisms Underlying the Categorical Perception..... | 1 |
| 2.1.1. Motor Theory | 1 |
| 2.1.2. Some Issues in the Motor Theory | 3 |
| 2.1.2.1. CP in nonspeech sounds..... | 3 |
| 2.1.2.2. CP in animals. | 4 |
| 2.1.2.3. Broca’s area and the motor cortex in speech perception..... | 5 |
| 2.1.2.4. Comments to the motor theory..... | 8 |
| 2.1.3. Factors Influencing Categorical Perception..... | 9 |
| 2.1.4. Multilevel and Multistore Processing Models | 11 |
| 2.1.4.1. Multilevel processing model | 11 |
| 2.1.4.2. Multistore processing model | 13 |
| 2.2. Behavioral Methods to Study Categorical Perception | 15 |
| 2.2.1. Identification Task | 15 |
| 2.2.1.1. Responses to each stimulus | 15 |

| | |
|---|----|
| 2.2.1.2. Boundary steepness and boundary position | 16 |
| 2.2.2. Discrimination Task..... | 20 |
| 2.3. ERP Components to Study Categorical Perception | 21 |
| 2.3.1. ERP Components Correlating with Categorical Perception | 21 |
| 2.3.2. MMN..... | 23 |
| 2.3.2.1. Characteristics of MMN..... | 23 |
| 2.3.2.2. Mechanisms and generators of MMN..... | 23 |
| 2.3.2.3. Applications of MMN | 25 |
| 2.3.3. P300 | 27 |
| 2.3.3.1. Characteristics of P300 | 27 |
| 2.3.3.2. Mechanisms and generators of P300..... | 27 |
| 2.3.3.3. Applications of P300..... | 29 |
| 2.3.4. Comments to MMN and P300 | 31 |
| 2.3.5. Neurophysiological Bases for Speech Perception | 31 |
| CHAPTER 3 BEHAVIORAL EXPERIMENTS ON CANTONESE LEVEL TONES..... | 32 |
| 3.0. Purposes of the Experiments..... | 32 |
| 3.1. Experiment I: to Examine the Context Effect..... | 32 |
| 3.1.1. Method | 33 |
| 3.1.1.1. Subjects | 33 |
| 3.1.1.2. Stimuli | 33 |
| 3.1.1.3. Identification task..... | 36 |
| 3.1.2. Data Analysis | 37 |
| 3.1.3. Results..... | 38 |
| 3.1.4. Summary and Discussion..... | 42 |
| 3.2. Experiment II: to Reexamine the Context Effect..... | 44 |
| 3.2.1. Method | 44 |
| 3.2.1.1. Subjects | 44 |

| | |
|---|-----------|
| 3.2.1.2. Stimuli | 44 |
| 3.2.1.3. Identification task | 46 |
| 3.2.2. Data Analysis | 46 |
| 3.2.3. Results | 47 |
| 3.2.4. Summary and Discussion | 50 |
| 3.3. Experiment III: to Check the Effect of Syllable and Pitch Transition | 53 |
| 3.3.1. Method | 53 |
| 3.3.1.1. Subjects | 53 |
| 3.3.1.2. Stimuli | 53 |
| 3.3.1.3. Identification task | 54 |
| 3.3.1.4. Discrimination task | 55 |
| 3.3.2. Data Analysis | 55 |
| 3.3.3. Results | 55 |
| 3.3.4. Summary and Discussion | 62 |
| 3.4. Experiment IV: to Check the Effect of Linguistic Background | 64 |
| 3.4.1. Method | 64 |
| 3.4.1.1. Subjects | 64 |
| 3.4.1.2. Stimuli | 64 |
| 3.4.1.3. Discrimination task | 64 |
| 3.4.2. Data Analysis | 65 |
| 3.4.2. Results | 65 |
| 3.4.4. Summary and Discussion | 70 |
| 3.5. Summary of chapter 3 | 71 |
| CHAPTER 4 BEHAVIORAL AND ERP EXPERIMENTS ON MANDARIN AND CANTONESE TONES | 73 |
| 4.0. Purposes of the Experiments | 73 |
| 4.1. Experiment I: MMN | 74 |

| | |
|--|----|
| 4.1.1. Method | 74 |
| 4.1.1.1. Subjects | 74 |
| 4.1.1.2. Stimuli | 75 |
| 4.1.1.3. EEG recording..... | 75 |
| 4.1.1.4. Behavioral task..... | 76 |
| 4.1.2. Data Analysis | 77 |
| 4.1.2.1. ERP data..... | 77 |
| 4.1.2.2. Behavioral data..... | 79 |
| 4.1.3. Results..... | 79 |
| 4.1.3.1. ERP results | 79 |
| 4.1.3.1. Behavioral results..... | 81 |
| 4.1.4. Discussion | 82 |
| 4.1.4.1. Cognitive stages of categorical perception of lexical tones | 82 |
| 4.1.4.2. Representation of lexical tones in the brain | 85 |
| 4.1.4.3. Double MMNs | 86 |
| 4.1.5. Summary | 88 |
| 4.2. Experiment II: P300 | 89 |
| 4.2.1. Method | 91 |
| 4.2.1.1. Subjects | 91 |
| 4.2.1.2. Stimuli | 91 |
| 4.2.1.3. Procedure during EEG recording..... | 92 |
| 4.2.1.4. Procedure during behavioral task..... | 93 |
| 4.2.1.5. EEG recording..... | 93 |
| 4.2.2. Data Analysis | 94 |
| 4.2.2.1. ERP data..... | 94 |
| 4.2.2.2. Behavioral data..... | 95 |
| 4.2.3. Results..... | 96 |

| | |
|---|-----|
| 4.2.3.1. Rising set (#1,#4,and #7) | 96 |
| 4.2.3.2. Level set (#4,#7,and #10)..... | 103 |
| 4.2.3.3. Identification results | 111 |
| 4.2.4. Discussion | 112 |
| 4.2.4.1. Category boundary | 112 |
| 4.2.4.2. Difference between speech and nonspeech..... | 115 |
| 4.2.4.3. Influence of language experiences on CP | 117 |
| 4.2.4.4. Brain activities and overt behavioral responses..... | 120 |
| 4.2.4.5. Hemispheric lateralization | 122 |
| 4.2.5. Summary | 126 |
| 4.3. Experiment III: Behavioral Test | 127 |
| 4.3.1. Method | 127 |
| 4.3.1.1. Subjects | 127 |
| 4.3.1.2. Stimuli | 127 |
| 4.3.1.3. Identification task..... | 127 |
| 4.3.1.4. Discrimination task | 128 |
| 4.3.2. Data Analysis | 128 |
| 4.3.2.1. Identification data..... | 128 |
| 4.3.2.2. Discrimination data | 129 |
| 4.3.3. Results..... | 130 |
| 4.3.3.1. Identification curves..... | 130 |
| 4.3.3.2. Discrimination results | 135 |
| 4.3.4. Discussion | 137 |
| 4.3.4.1. Effect of language experience on category boundary | 137 |
| 4.3.4.2. Difference between speech and nonspeech..... | 138 |
| 4.3.4.3. Linguistic boundary and psychological boundary in discrimination curve..... | 139 |

| | |
|---|-----|
| 4.3.4.4. Comparison between identification and discrimination results | 141 |
| 4.3.5. Summary | 142 |
| 4.4. Summary of chapter 4 | 143 |
| CHAPTER 5 GENERAL DISCUSSIONS AND MAJOR CONTRIBUTIONS | 147 |
| 5.1. General discussions | 147 |
| 5.1.1. Factors influencing CP | 147 |
| 5.1.1.1. Intrinsic acoustic properties | 148 |
| 5.1.1.2. Position of target syllables relative to context | 149 |
| 5.1.1.3. Language backgrounds | 151 |
| 5.1.1.4. Carrier syllables | 153 |
| 5.1.2. Temporal dynamics of CP | 155 |
| 5.1.3. Multistage model for tone perception | 160 |
| 5.2. Major contributions of this work | 162 |
| 5.3. Future work | 162 |
| REFERENCES | 164 |
| APPENDIX 1 ABBREVIATIONS | 180 |
| APPENDIX 2 STATISTICAL ANALYSIS ON CHAPTER 3 EXPERIMENT III | 181 |
| APPENDIX 3 STATISTICAL ANALYSIS ON CHAPTER 3 EXPERIMENT IV | 184 |
| APPENDIX 4 PUBLICATION LIST | 187 |

LIST OF TABLES

| | | |
|-------------|--|----|
| Table 1.2.1 | Syllable structure of a Chinese dialect. Elements enclosed in [] are optional [Taken from (Wang, 1973)]..... | 4 |
| Table 3.1.1 | F0 value of stimuli in Experiment Ia (isolation form). Anchor syllables, which were used as filler, were naturally uttered speech. Synthetic syllables were based on stim#8, the base syllable. 33 | |
| Table 3.1.2 | Mean F0 value (Hz) of stimuli in Experiment Ib (with contextual sentence). Synthetic syllables were based on stim#7, the base syllable | 35 |
| Table 3.1.3 | Grouping information, grouped mean value, and statistical analysis of predicted discrimination scores in three continua of Experiment I. **: $p < 0.003$; * : $p < 0.017$. (significant level $p < 0.017$ after correction)..... | 41 |
| Table 3.2.1 | Mean F0 value and F0 dist. of stimuli in Experiment II (with contextual sentence). Synthetic sentences were based on stim#9, the base syllable..... | 44 |
| Table 3.2.2 | Grouping information, grouped mean value, and statistical analysis of predicted discrimination scores in two continua of Experiment II. **: $p < 0.005$; * : $p < 0.025$ (after correction for type I error). | 49 |
| Table 3.3.1 | Grouping information, grouped mean value, and statistical analysis of predicted discrimination scores in two continua (SSC and INT) of Experiment III (after correction). **: $p < 0.0025$; * : $p < 0.0125$ | 57 |
| Table 3.3.2 | One-way RMANOVA on effect of stimulus in Experiment III. **: $p < 0.001$; * : $p < 0.05$ | 58 |
| Table 3.3.3 | Grouping information, grouped mean value, and statistical analysis of obtained discrimination scores in two continua (SSC and INT) of Experiment III (after correction). **: $p < 0.0025$; * : $p < 0.0125$ | 60 |
| Table 3.3.4 | Grouping information, grouped mean value, and statistical analysis of obtained discrimination scores in two continua (SSC, NWS, CPT) of Experiment III (after correction). **: $p < 0.0025$; * : $p < 0.0125$ | 62 |
| Table 3.4.1 | One-way RMANOVA on the effect of stimulus in Experiment IV **: $p < 0.001$; * : $p < 0.05$ | 65 |
| Table 3.4.2 | Grouping information, grouped mean value, and statistical analysis of obtained discrimination scores in three conditions (CS, MS and FS) of Experiment IV (after correction). **: $p < 0.0001$; * : $p < 0.008$ | 68 |
| Table 3.4.3 | Grouping information, grouped mean value, and statistical analysis of obtained discrimination scores in two conditions (CN and FN) of Experiment IV(after correction). **: $p < 0.0025$; * : $p < 0.0125$ | 68 |
| Table 4.1.1 | Summary of major results in study..... | 86 |
| Table 4.2.1 | Mean value and latency of difference wave (deviant-minus-standard) and hit rate for deviants in rising set (#1, #4, and #7) . * : $p < 0.0125$ (significant level was set at 0.0125 after correction) 98 | |

| | | |
|---------------------|--|-----|
| Table 4.2.2 | Mean value and latency of difference wave (deviant-minus-standard) and hit rate for deviants in level set (#4, #7, and #10) . *: $p < 0.0125$ (significant level was set at 0.0125 after correction) | 105 |
| Table 4.3.1 | Boundary position and boundary width on speech and nonspeech continua from three language groups calculated by probit approximation. | 131 |
| Table 4.3.2 | Grouping information, grouped mean value, and statistical analysis of predicted discrimination scores for continuum spanning from level to rising tones in six conditions (2 SYLLABLE \times 3 LANGUAGE). * : $p < 0.004$; (significant level $p < 0.004$ after correction) | 131 |
| Table 4.3.3 | Grouping information, grouped mean value, and statistical analysis of obtained discrimination scores for continuum spanning from level to rising tones in three language groups. * : $p < 0.008$; (significant level $p < 0.008$ after correction) | 137 |
| Table 4.4.1 | CP effect (Across-category > Within-category) of level vs. rising tones in different cognitive stages for Mandarin listeners. | 145 |
| Table 4.4.2 | Presentation of psychoacoustic and linguistic category boundaries in the continuum spanning from the level to the rising tone | 146 |
| Table 5.1.1 | Summary of the effect of carrier syllables on discrimination performance | 154 |
| Table 5.1.2 | Presentation of general auditory processing and language specific processing in three stages tested in the oddball paradigm. | 159 |
| Appendix Table 2.1. | Statistical analysis on predicted discrimination scores on simple shifted (SSC) and interpolation (INT) continua..... | 181 |
| Appendix Table 2.2. | Statistical analysis on obtained discrimination scores on simple shifted (SSC) and interpolation (INT) continua..... | 182 |
| Appendix Table 2.3. | Statistical analysis on obtained discrimination scores on simple shifted (SSC), nonword sounds (NWS), and complex tones (CPT) continua..... | 183 |
| Appendix Table 3.1. | Statistical analysis on obtained discrimination scores on language effect on real word continua | 184 |
| Appendix Table 3.2. | Statistical analysis on obtained discrimination scores on language effect on nonspeech continua..... | 185 |
| Appendix Table 3.3. | Statistical analysis on obtained discrimination scores on interaction of real word and nonspeech continua by Cantonese and French subjects..... | 186 |

LIST OF FIGURES

| | | |
|--------------|--|----|
| Figure 1.1.1 | Illustrations of the spectrographic patterns for stimuli used in the experiment by Haskins Laboratories. Pattern 14, at the lower right, is complete in both time and spectral domains. (For discussion, see the text for details) [Taken from (Lieberman, Harris, Hoffman, & Griffith, 1957)] | 1 |
| Figure 1.1.2 | Plot of ideal categorical perception (CP). Blue and green solid lines represent identification scores of category A and B respectively. Dotted red line represents correct discrimination of every two adjacent stimuli pair. | 2 |
| Figure 1.2.1 | Mandarin tones. The right panel represents the pitch contours. The middle four characters in left panel represented the corresponding characters to the pitch contours. [Taken from (Wang, 1973)]..... | 5 |
| Figure 1.2.2 | Pitch contours for Cantonese tones. Each tone is uttered by a male speaker. [Taken from (Peng, 2006)] | 5 |
| Figure 1.3.1 | A schematic illustration of EEG recording procedure. See the text for details. [Taken from(Luck, 2005)]..... | 7 |
| Figure 2.1.1 | A multistore model of CP. It includes four memory stores: sensory memory trace, analyzed sensory memory, short-term categorical memory, and long-term categorical memory. Information is encoded in a hierarchical order but short-term categorical memory and analyzed sensory memory can be processed in parallel. All the sensory and short-term categorical components are subject to memory decay. The available memory traces after decay are input for decision-making. If long-term categorical memory is also available, it will interact with short-term categorical memory via both top-down and bottom-up mechanisms. All the memory components with relatively longer lifetime are involved in context-coding.[Taken from(Y. S. Xu, Gandour, & Francis, 2006)]..... | 13 |
| Figure 2.2.1 | Schematic plots of identification responses from two data sets, the average response over the whole continuum is shown below the plot. a) The two sets differ in category boundaries, and the average information indicates the difference. b) The two sets differ in boundary sharpness but not in the position of the boundary, and the average information cannot indicate the difference. | 16 |
| Figure 2.2.2 | Three constraints with different degrees of gradience representing possible phonotactic constraints. [Taken from (Frisch, et al., 2004)] | 18 |
| Figure 2.2.3 | CDF function $\Phi_{\mu\sigma^2}(x)$ with different combinations of σ^2 and μ values. | 18 |
| Figure 2.3.1 | Schematic illustration of the P300 context-updating model.[Taken from (Polich, 2003b)] | 28 |
| Figure 3.1.1 | Stimuli continua structure of Experiment I. The step in the continua was 5Hz. We anticipated TS in stimulus No.1 to be perceived as [fan ¹] and TS in stimulus No.11 to be perceived as [fan ⁶]. | 36 |
| Figure 3.1.2 | Pitch contours of LC (left context) and RC (right context) in Experiment I. We anticipated TS (target syllable) in stimulus No.1 (blue line) to be perceived as [fan ¹], TS in stimulus No.7(green line) to be perceived as [fan ³], and TS in stimulus No.7(red line) to be perceived as [fan ⁶]. T1: tone 1; T3: tone 3; T6: tone 6. The dashed vertical lines in each plot are the boundaries of each word..... | 36 |

| | | |
|--------------|--|----|
| Figure 3.1.3 | Results of natural speech based Cantonese level tones perceived by native subjects in Experiment I. The target syllables had three types of positions: in the isolation form (Isolation), at the beginning of the sentence (RC), at the end of the sentence (LC). A solid line with a cross represented predicted discrimination (Pred disc.) calculated from the identification results. A solid line with a circle, a dotted line with a square, and a solid line with a rectangle represented the proportion of the stimuli perceived as tone 1 (T1), tone 3 (T3) and tone 6 (T6) respectively..... | 40 |
| Figure 3.1.4 | The delta value of identification responses (Δp_i) in three continua with different positions of target in Experiment I. The solid line with a rectangle, the dash line with a diamond, and the dotted line with a square represented the Δp_i when it was in isolation form, in RC continuum and in LC continuum respectively..... | 40 |
| Figure 3.1.5 | Grouped value of predicted discrimination according to AC (Across-category) and WC(within-category) in three continua of Experiment I. **: $p < 0.01$; ns. Non significant, RC: right context, LC: left context; ISO: isolation. Error bar represented standard error..... | 42 |
| Figure 3.2.1 | Stimuli continua structure of Experiment II. The step in the continua was 7.0 mels. We anticipated TS in stimulus No.1 to be perceived as /si ¹ / and TS in stimulus No.13 to be perceived as /si ⁶ /. Stimulus No.9 was the base sentence, based on which the continua were synthesized..... | 45 |
| Figure 3.2.2 | Pitch contours of LC (left context) and RC (right context) in Experiment II. We anticipated TS (target syllable) in stimulus No.1 (blue line) to be perceived as [si ¹], TS in stimulus No.9 (green line) to be perceived as [si ³], and TS in stimulus No.9 (red line) to be perceived as [si ⁶]. T1: tone 1; T3: tone 3; T6: tone 6. | 46 |
| Figure 3.2.3 | Results of synthetic based Cantonese level tones perceived by native subjects in Experiment II. The target syllables had two types of positions: at the beginning of the sentence (RC), at the end of the sentence (LC). A solid line with a cross represented predicted discrimination (Pred disc.) calculated from the identification results. A solid line with a circle, a dotted line with a square, and a solid line with a rectangle represented the proportion of the stimuli perceived as tone 1 (T1), tone 3 (T3) and tone 6 (T6) respectively. | 47 |
| Figure 3.2.4 | The delta value of identification responses (Δp_i) in two continua with different positions of target in Experiment II. The solid line with a diamond and the dash line with a rectangle represented the Δp_i when it in RC continuum and in LC continuum respectively. | 48 |
| Figure 3.2.5 | Grouped value of predicted discrimination according to AC (Across-category) and WC (within-category) from LC(left context) and RC(right context) continua in Experiment II. **: $p < 0.005$; *: $p < 0.025$; ns. Non significant. Error bar represented standard error. | 49 |
| Figure 3.3.1 | Pitch contours used in Experiment IIIa. SSC: simply shifted continuum; INT: interpolation continuum; LC: left context; RC: right context; T1: tone 1; T3: tone 3; T6: tone 6. The dashed vertical lines in each plot are the boundaries of each word..... | 53 |
| Figure 3.3.2 | Results of two sets of Cantonese level tones perceived by native subjects in Experiment IIIa. SSC: simply shifted continuum; INT: interpolation continuum. The target syllables had two types of positions: at the beginning of the sentence (RC), at the end of the sentence (LC). A solid line with a cross represented predicted discrimination (Pred disc.) calculated from the identification results. A solid line with a circle, a dotted line with a square, and a solid line with a rectangle represented the proportion of the stimuli perceived as tone 1 (T1), tone 3 (T3) and tone 6 (T6) respectively..... | 56 |
| Figure 3.3.3 | Grouped value of predicted discrimination according to AC (Across-category) and WC (within-category) from left context and right context continua in Experiment III. SSC: simply | |

| | |
|---|----|
| shifted continuum; INT: interpolation continuum; **: p<0.0025; *: p<0.0125; ns. Non significant (after correction for multiple t-test). Error bar represented standard error. | 58 |
| Figure 3.3.4 Obtained discrimination scores of LC and RC in Experiment III. SSC: simply-shifted continuum; INT: interpolation continuum; NWS: non-word sounds; CPT: complex tones..... | 59 |
| Figure 3.3.5 Grouped value of obtained discrimination according to AC (Across-category) and WC (within-category) from left context and right context continua in Experiment III to study the effect of pitch transition. SSC: simply shifted continuum; INT: interpolation continuum; **: p<0.0025; *: p<0.0125; ns. Non significant (after correction for multiple t-test). Error bar represented standard error. | 60 |
| Figure 3.3.6 Grouped value of obtained discrimination according to AC (Across-category) and WC (within-category) from left context and right context continua in Experiment III to study the effect of carrier syllables. SSC: simply shifted continuum; NWS: non-word sounds; CPT: complex tones; **: p<0.008; *: p<0.002; ns. Non significant (after correction for multiple t-test). Error bar represented standard error. | 62 |
| Figure 3.4.1 Obtained discrimination scores of LC and RC in Experiment IV. CS: Cantonese + Speech; CN: Cantonese + nonspeech; MS: Mandarin + Speech; FS: French + Speech; FN: French + nonspeech..... | 66 |
| Figure 3.4.2 Grouped value of obtained discrimination according to AC and WC from left context and right context continua in Experiment IV. CS: Cantonese+speech; MS: Mandarin+speech; FS:French+speech; CN:Cantonese+nonspeech; FN: French+nonspeech; **: p<0.001; *: p<0.008; ns. Non significant (after correction for multiple t-test). Error bar represented standard error. | 66 |
| Figure 4.1.1 a) Schematic diagram for pitch contours of 11 stimuli. Stimulus #7 was the category boundary (following Wang, 1976). Stimulus #10, #8 and #6 were used as within-category-deviant, standard and across-category-deviant, respectively. b) The amplitude profiles of speech and nonspeech. | 74 |
| Figure 4.1.2 a) Schematic diagram for the oddball paradigm. b) EGI's 128-channel Channel layout and the clustered channels that correspond to the international 10/20 system and nine spatial locations for the mean amplitude and statistical analysis of the ERP and MMN waves. EGI: Electrical Geodesics, Inc. | 76 |
| Figure 4.1.3 a) Grand averaged, from 19 subjects, ERP and difference waves for speech and nonspeech conditions from Fz cluster. Dot-shadowed and slash-shadowed areas represent two time regions (128–168 ms and 228–268 ms) of MMNs. b) Top view of a topographic map for two types of difference waves (across-category minus standard; within-category minus standard) in speech and nonspeech conditions over the two time windows. * : p<.05 on at least one spatial location (details are shown in text)..... | 78 |
| Figure 4.1.4 a) Identification and part of discrimination scores in speech and nonspeech conditions. b) Discrimination scores along the completed continuum in speech and nonspeech conditions. Error bars represented standard error. | 82 |
| Figure 4.2.1 The spectrogram information of speech and nonspeech (complex tone) in Experiment II. | 91 |
| Figure 4.2.2 The location information of 129 channels. Three channels (Poz, Po3 and Po4) were picked up for P300 analysis..... | 92 |
| Figure 4.2.3 The EPR and difference waves from rising sets in Experiment II. The left two panels were data from Mandarin subjects, and the right two panels were data from Cantonese subjects. All the data were extracted on channel Poz. | 99 |

| | | |
|---------------|---|-----|
| Figure 4.2.4 | The topographic maps of difference waves from rising sets in Experiment II. The left two panels were data from Mandarin subjects, and the right two panels were data from Cantonese subjects. The topographic maps were centered at the peak of each condition..... | 99 |
| Figure 4.2.5 | Summary of (a) ERP difference waves and (b) hit rate in rising set for four conditions (2 syllable types \times 2 language groups). | 100 |
| Figure 4.2.6 | The value of CP (AC-WC) shown in (a) ERP and (b) hit rate data (see text for details). All data were from level set..... | 100 |
| Figure 4.2.7 | The value of difference waves on left and right channels in different conditions. a) Data for Mandarin subjects; b) data for Cantonese subjects. All data were for rising set. AC: across-category deviant; WC: within-category deviant. | 103 |
| Figure 4.2.8 | Summary of hemisphere lateralization pattern contributed by a) language experiences, b) types of deviants, and c) types of carrier syllables. All data were for rising set. AC: across-category deviant; WC: within-category deviant..... | 103 |
| Figure 4.2.9 | The EPR and difference waves from level set in Experiment II. The left two panels were data from Mandarin subjects, and the right two panels were data from Cantonese subjects. All the data were extracted on channel Poz. | 105 |
| Figure 4.2.10 | The topographic maps of difference waves from level set in Experiment II. The left two panels were data from Mandarin subjects, and the right two panels were data from Cantonese subjects. The topographic maps were centered at the peak of each condition..... | 106 |
| Figure 4.2.11 | Summary of (a) ERP difference waves and (b) hit rate in level set for four conditions (2 syllable types \times 2 language groups)..... | 106 |
| Figure 4.2.12 | The value of CP (AC-WC) shown in (a) ERP and (b) hit rate data (see text for details). All data were from level set..... | 107 |
| Figure 4.2.13 | The value of difference waves on left and right channels in different conditions. a) Data for Mandarin subjects; b) data for Cantonese subjects. All data were for level set. AC: across-category deviant; WC: within-category deviant. | 110 |
| Figure 4.2.14 | CP in speech and nonspeech conditions presented in two hemispheres (the level set) 110 | |
| Figure 4.2.15 | Summary of hemisphere lateralization pattern contributed by a) language experiences, b) types of deviants, and c) types of carrier syllables. All data were for rising set. AC: across-category deviant; WC: within-category deviant..... | 110 |
| Figure 4.2.16 | Identification results on stimulus #1, #4, #7, and #10 from Mandarin and Cantonese subjects. Error bar represented standard error. | 112 |
| Figure 4.3.1 | Results of identification of level vs. rising tone continuum (open circles) and their corresponding probit approximation (solid line), with the predicted discrimination results (dotted line) superimposed over the identification curves. Upper panel represented results from the nonspeech continuum and lower panel represented results from the speech continuum. From left to right, three panels represented the performance of Mandarin, Cantonese and German subjects respectively..... | 132 |
| Figure 4.3.2 | a) Boundary width of continuum spanning from level to rising tones in six conditions [2 SYLLABLE (speech and nonspeech) \times 3 LANGUAGE (Mandarin, Cantonese, and German)]. A smaller number of boundary width indicated a sharper category boundary. b) Difference of boundary width (speech–nonspeech) on three language groups. *: $p < 0.05$, ns: non significant (after Bonferroni correction). Error bar represented standard error. | 132 |

Figure 4.3.3 Grouped value of predicted discrimination of level to rising continuum according to AC (Across-category) and WC(within-category). a) In the nonspeech continuum; b) in the speech continuum. *: $p < 0.004$; ns. Non significant. Error bar represented standard error..... 134

Figure 4.3.4 Obtained discrimination data of the speech continuum spanning from level to rising tones by Mandarin, Cantonese and German subjects. a) Discrimination curves along the speech continuum; b) grouped values of within-category pairs (WC), across-category pairs (AC), and level end pair. 136

ABSTRACT

Speech sounds vary across different conditions and subjects; nevertheless, listeners perceive the phonemes without difficulties. *Categorical perception* (CP) occurs when listeners map the varying speech sounds into discrete phonemic categories. In CP, to discriminate a pair of stimuli that cross a category boundary is much easier than those that lie within the same category, even though both pairs are separated by an equal physical difference. CP is one of the important properties essential for speech perception.

Pitch contour or its acoustic correlate, *fundamental frequency* (F0), distinguishes lexical meanings in tone languages. Two topics on CP of lexical tones were studied in the thesis: (1) the factors influencing CP, and (2) the temporal process of CP. These two topics were investigated through both behavioral and *event-related-potential* (ERP) methods on Cantonese and Mandarin tones.

Four factors were studied. They were (1) intrinsic acoustic properties of pitch contours by comparison between continua of level tones and contour tones; (2) positions of target syllables relative to context (without contextual sentence, at the beginning and at the end of the contextual sentence); (3) language backgrounds by comparison between listeners with different tone experiences; and (4) carrier syllables (real word, non word, and nonspeech). Three temporal stages were studied in the same experimental paradigm. They were (1) the preattentive stage investigated through the *mismatch negativity* (MMN); (2) the attentive stage investigated through the P300; and (3) the overt response stage investigated through the hit rate data.

All these four factors influence the degree of CP. In the discussion, both general auditory processing and language specific processing are suggested to be responsible for the various types of exhibition of CP, although they have different weights for different factors. Different patterns of CP were also observed in three temporal stages due to different weights of these two types of processing. In summary, a multistage model which includes both general auditory processing and language specific processing is proposed to explain the CP of lexical tones. This model improves previous models by proposing that the weights of these two types of processing in speech perception depend on the types of factors, and the temporal processing stages.

Finally, for the first time in the literature, the thesis also reported that even though a tone contrast (i.e., level vs. rising) is present in both tone systems, the same contrast is perceived differently by the two groups of subjects by virtue of their different language experiences.

摘要

同一句話，不同人或者相同的人在不同場合下發出的語音信號都是不一樣的，但是聽眾總能夠毫不困難地識別出這句話。語音的最小單位是音素。當人們將所聽到的變化多端的語音信號對應到有限的幾個音素上時，一種稱為範疇感知的現象出現了。範疇感知意味著兩種類別交界處的分辨率遠遠高於類別內部的分辨率。範疇感知是語音感知的一個重要特性。

在聲調語言中，聲音的音高或者基頻曲綫也可以用來區別語義，因此也是一種音素。本論文討論兩個關於聲調範疇感知的議題：(1) 影響聲調範疇感知的因素 (2) 聲調範疇感知的時間動態特性。這兩個議題將通過行為和事件相關電位兩種實驗手段進行研究。研究的對象是普通話和廣東話（粵語）的聲調。

本論文研究了四種因素。它們分別是(1) 基頻曲綫的內在聲學特性，包含了平聲調和曲折調；(2) 目標音節在上下文句子中的位置，包含孤立詞（沒有上下文），在句子開頭，和在句子結尾這三種情況；(3) 被試的聲調語言背景；和 (4) 載波音節的類型（真字、假字和非語音）。本論文還在同一种試驗範式下，研究了範疇感知在三種不同時間階段的表現。這三種時間階段是(1) 前知覺階段，主要通過失匹配負波來探測；(2) 知覺階段，主要通過失 P300 來探測；(3) 外顯反映階段，主要通過命中率來探測。

研究結果表明所有測試因素都影響了範疇感知的程度。在論文討論部分，不同的範疇感知表現被認為是因為測試因素中兩種處理過程的不同加權的體現。這兩種過程是普遍聲音處理過程和語言特有的處理過程。範疇感知在三個不同時間階段，表現形式也不一樣。同理，這些不同也是因為上述兩種過程的不同加權的體現。總之，本論文提出一種包括兩種過程的新的語音感知模型來解釋聲調的範疇感知問題。這種新的模型與以前的模型一樣，都包含了普遍聲音處理和語言特有的處理過程，但是這個模型首次提出這兩種過程在不同的狀態和不同的時間階段有不同的加權表現。

最後，本論文首次發現一種新的語言現象：即使兩種不同的聲調系統都包含同一對聲調對照物，因為被試的母語有不同的聲調系統，因此他們對這一對相同的聲調對照有不同的感知模式。這種語言現象據我們所知從未被報道過。