



Neurocognitive processing of Mandarin two character compounds with a neutral tone

Zhongpei Zhang

► To cite this version:

Zhongpei Zhang. Neurocognitive processing of Mandarin two character compounds with a neutral tone. Linguistics. Université de Nanterre - Paris X, 2022. English. NNT : 2022PA100132 . tel-04131230

HAL Id: tel-04131230

<https://theses.hal.science/tel-04131230>

Submitted on 16 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Zhongpei Zhang

Neurocognitive processing of Mandarin two-character compounds with a neutral tone

Thèse présentée et soutenue publiquement le 09/12/2022
en vue de l'obtention du doctorat de Sciences du langage
de l'Université Paris Nanterre
sous la direction de M. Frédéric ISEL (Université Paris Nanterre)

Jury * :

Rapporteur :	Madame Fanny Meunier	Directrice de recherche, CNRS
Rapporteur :	Monsieur James Myers	Professeur, National Chung Cheng University, Taiwan
Membre du jury :	Madame Mireille Besson	Directrice de recherche, CNRS
Membre du jury :	Monsieur Nicolas Audibert	Maître de conférences, Université Paris 3 Sorbonne nouvelle
Membre du jury :	Madame Florence Villoing	Professeure, Université Paris Nanterre
Directeur :	Monsieur Frédéric ISEL	Professeur, Université Paris Nanterre

Neurocognitive processing of Mandarin two-character compounds with a neutral tone

By

Zhongpei ZHANG

ABSTRACT

This doctoral thesis examines the processing of Mandarin two-character compounds with a neutral tone in the auditory modality. By using a multi-speaker production approach, the acoustic properties (e.g., duration, fundamental frequency (F0), and intensity) of the second neutral tone were analyzed. In the first study, eighty minimal pairs differing in their tonal patterns of the second constituents, e.g., the neutral tone compound 云彩/yun2cai0/(cloud) (strong-weak) and the full tone compound 色彩/se4cai3/(color) (strong-strong) were selected. The acoustic measures using the PRAAT were performed on the second neutral tones. The second study also focused on the second neutral tone but complemented the first one by exploring the influence of the preceding and following tonal information. These latter measures allowed us to study both the anticipatory and the carry-over effects. In general, these acoustic measures confirmed that the duration of the second neutral tones was on average significantly reduced (around 70 to 80%) in relation to their full tone counterparts. Moreover, these acoustic measures revealed a carry-over effect from the preceding tonal context on the duration and intensity of the second neutral tone. The third and fourth studies were constructed to investigate the neurophysiological effect of the lexical frequency and prosodic pattern in the auditory processing of the neutral tone compounds. A cross-modal auditory-visual repetition priming paradigm with a lexical decision task was used in the third study. Both the lexical frequency of whole compound and initial constituent, as well as the tonal pattern of compound words, were systematically manipulated to examine whether these factors constitute the determining cues of morphological decomposition. The N400 repetition priming effect was applied as an indicator of decomposition, the size of this effect was greater when compounds were rare, the first constituents were low-frequency, and the compound words carried neutral tones on their second constituents. The activation of the first constituents was observed in the tasks, which is demonstrated by the priming effect in both the behavioral results and EEG results. And this can constitute indirect evidence of the morphological decomposition. What's more, the prosodic information of individual constituents was found to be coded in the lexical representation, at least for Mandarin compounds. In the fourth study, the role of syllable duration in the processing of neutral tone compounds was explored, and the individual constituents were lengthened to create prosodically incongruent stimuli. Results showed that lengthened second neutral tones elicited the

largest MMN, N2b, and N400 effects in relation to the standard forms, the semantic representation was hindered by the incongruous prosody pattern of the second neutral tones in the compound processing. Taken together, the main findings of this dissertation indicate that tonal information of the second constituents within two-character compounds, combined with the lexical frequency of the whole-compounds and first constituents, constitutes the determining cues of morphological decomposition in the processing of Mandarin neutral tone compounds. Moreover, the duration of the second constituent carrying a neutral tone seems to be a crucial factor in lexical representation. The interplay between these different cues remains an open question for further studies.

Keywords: Compound words, Mandarin, Prosody, Neutral tone, ERP.

Traitement neurocognitif des mots composés de deux caractères portant un ton neutre en Mandarin

Par
Zhongpei Zhang

Résumé

Le Mandarin est une langue tonale qui a quatre tons lexicaux appelés ‘ton plein’ (Ton 1/2/3/4), et une catégorie de syllabe « légère » qui ne porte pas d’accent de mot que l’on appelle ‘ton neutre’ (ou ‘ton zéro’). Les mots composés en Mandarin sont majoritairement constitués de deux morphèmes ou constituants, et environ 5% des mots composés dissyllabiques portent un ton zéro sur le deuxième constituant (C2) et un ton plein sur le premier constituant (C1). Cette thèse porte sur l’étude du traitement auditif des mots composés du Chinois Mandarin qui contiennent deux constituants, le premier portant un ton plein et le second un ton neutre ou un ton plein, chez des adultes natifs du Mandarin. Cela nous a permis d’opposer deux patrons prosodiques différents, c’est-à-dire Fort-fort vs. Fort-faible. Cette structure morphologique et phonologique spécifique nous offre un cas particulier pour étudier l’interaction entre la morphologie lexicale (composition) et la phonologie. Cette thèse est constituée de quatre expériences, deux expériences de production sur lesquelles ont été conduites des analyses acoustiques et deux expériences de réception EEG de mesures de potentiels évoqués qui représentent les réponses électriques des cellules pyramidales du cerveau mesurées après amplification à la surface du scalp.

Dans la tâche de production, une approche méthodologique multi-locuteurs a été choisie donnant lieu ensuite à l’analyse des propriétés acoustiques (e.g., durée, F0 et intensité) du mot composé de deux caractères portant un ton plein ou un ton neutre dérivé de ce ton plein. L’objectif de la première expérience était de décrire et de comparer les caractéristiques phonétiques des C2 qui portent soit un ton neutre ou soit un ton plein, c’est-à-dire [Fort-faible] ou [Fort-fort]. Quarante-paires minimales des mots composés ont été sélectionnées, son second constituant portant soit un ton neutre (e.g., 云彩/yun2cai0/, nuage), soit un ton plein (e.g., 色彩/se4cai3/, couleur). Les caractéristiques acoustiques du ton neutre ont été analysées en comparaison avec celles du ton plein homologue dans le contexte de quatre tons pleins (Ton 1/2/3/4). Nos résultats confirment les observations rapportées dans des études précédentes sur les caractéristiques acoustiques des tons neutres. Nous avons répliqué l’effet de réduction de la durée (environ 70 à 80% de réduction) et de l’intensité des constituants en position finale portant un ton neutre comparés aux mêmes constituants mais portant un ton plein.

La deuxième expérience avait pour objectif d'examiner les effets de la coarticulation entre les deux constituants des mots composés qui portent un ton neutre. Plus spécifiquement, nous avons commencé par décrire les caractéristiques phonétiques des deuxièmes constituants (C2) des mots composés, qui porte un ton neutre, selon les 4 tons pleins portés par les premiers constituants (C1). Ensuite, nous avons analysé les caractéristiques phonétiques des premières constituants (C1) en fonction des quatre tons pleins et du ton neutre porté par des deuxièmes constituants (C2). Nos résultats révèlent des effets de transfert du contexte tonal des premières constituants (C1) sur la durée et l'intensité des deuxièmes constituants (C2) des mots composés portant un ton neutre. De plus, nous avons aussi démontré que la valeur de fréquence fondamentale du début du deuxième ton neutre est déterminée par la valeur de la fin du premier ton plein.

La troisième expérience conduite en réception avec la technique EEG avait pour objectif d'examiner les effets de la fréquence lexicale du mot composé et du premier constituant, ainsi que le patron prosodique des mots composés (Fort-faible vs. Fort-fort). Nous souhaitons comprendre si ces facteurs jouent un rôle important pour déclencher la décomposition morphologique des mots composés et nous souhaitons aussi d'étudier leur interaction. Pour ce faire, nous avons utilisé le paradigme de répétition audio-visuel avec une tâche de décision lexicale. L'indicateur critique pour déterminer si une décomposition morphologique avait eu lieu lors du traitement auditif des mots composés est l'effet N400 d'amorçage de répétition entre le constituant initial et une cible visuelle présentée après l'audition du composé qui répétaient ou non le constituant initial. Dans la littérature, rappelons que la N400 est un indicateur neurophysiologique de l'accès au lexique. En tout premier lieu, les résultats comportementaux ont montré un effet d'amorçage sur le premier constituant (C1) des mots composés de deux caractères. Les analyses sur les effets N400 montrent que les mots composés rares, les mots composés portant un premier constituant rare, et les mots composés portant un ton neutre avec un premier constituant rare ont tous donné lieu à un effet d'amorçage de répétition N400 sur le constituant initial. Ce résultat suggère que ces trois facteurs jouent un rôle dans la décomposition morphologique comme l'atteste la trace d'activation de la représentation lexicale du premier composant dans nos données comportementales de temps de réaction et nos données neurophysiologiques.

La quatrième expérience avait pour objectif d'examiner comment la structure rythmique du mot composé portant un ton neutre influence la reconnaissance du mot. Mais au-delà de cette question, le point critique testé dans cette étude est de comprendre le rôle de la durée syllabique des deux constituants individuelles (C1, C2) sur le processus de la reconnaissance du mot composé. Pour ce faire, nous avons évalué l'impact de l'augmentation de la durée du C1 portant un ton plein et du C2 portant un ton neutre. Notre hypothèse est que l'effet N400 supposé refléter le traitement

lexical du mot composé devrait varier en fonction des manipulations réalisées sur la durée du C1 ou du C2. Plus spécifiquement, nous prédisons qu'un allongement du ton neutre, compte tenu du fait que ce ton constitue un indice de décomposition morphologique en Mandarin conduisant à une ré-analyse lexicale-sémantique du mot composé en entier, pourrait être le plus perturbant dans l'accès au mot composé. Afin de réaliser cette expérience, nous avons sélectionné soixante mots composés de deux caractères portant un ton neutre, présentant deux types de durées syllabiques (standard ou doublé) résultant en quatre conditions expérimentales : standard-standard (contrôle), standard-doublé, doublé-standard, doublé-doublé. Les résultats neurophysiologiques ont montré que c'est la manipulation acoustique du second constituant, c'est-à-dire le doublement de sa durée, qui a donné lieu aux effets MMN, N2b et N400 les plus larges par rapport à la forme standard. De façon intéressante, l'incongruité acoustique du premier constituant portant un ton plein a évoqué un effet de N400 avec une latence tardive (autour de 550ms post-stimulus onset). Prises dans leur ensemble, les données de l'expérience suggèrent que la durée du ton neutre porté par le constituant final des mots composés en Mandarin constitue un indicateur important d'accès à la représentation sémantique appropriée.

Pour résumer, les principaux résultats de cette thèse montrent que l'information tonale du second constituant des mots composés, en plus de l'information sur la fréquence du mot composé et de son constituant initial, constitue un indice déterminant de la décomposition morphologique dans le traitement auditif des mots composés. De plus, la durée du second constituant portant un ton neutre semble être un facteur critique pour guider l'accès lexical du mot composé en Mandarin et en particulier l'accès à la représentation sémantique des mots. L'interaction entre les différents indices reste une question ouverte que d'autres études devront examiner en détails.

Mots-clés : Mots composés, Mandarin, Prosodie, Ton neutre, ERP.

Table of Contents

Introduction	- 13 -
1 Lexical stress and Mandarin neutral tone.....	- 23 -
1.1 Lexical stress in speech processing.....	- 23 -
1.1.1 Definition of lexical stress	- 23 -
1.1.2 Typological properties of lexical stress	- 24 -
1.1.3 Assignment of lexical stress	- 25 -
1.1.4 Phonetic correlates of lexical stress.....	- 26 -
1.2 Tonal and intonational processing model	- 27 -
1.3 Mandarin phonology system and neutral tone.....	- 34 -
1.3.1 A brief introduction to the Chinese phonological system	- 34 -
1.3.2 Neutral tone in Mandarin	- 36 -
1.3.3 Acoustic correlates of neutral tone	- 38 -
2 Compounding in Chinese morphology.....	- 43 -
2.1 Mandarin compound words	- 43 -
2.1.1 Brief introduction to Chinese morphology	- 44 -
2.1.2 Definition and classification of Mandarin compounds	- 46 -
2.1.3 Two-character neutral tone compound.....	- 49 -
2.2 The framework of compounding processing.....	- 51 -
2.2.1 Multilevel interactive-activation model.....	- 51 -
2.2.2 Multi-layer representation model by Zhou & Marslen-Wilson (1999)	- 53 -
2.3 Language-related property in Mandarin compound processing	- 55 -
2.3.1 Semantic transparency.....	- 55 -
2.3.2 Lexical frequency effect	- 56 -
2.3.3 Morphological family size	- 57 -
2.3.4 Orthographic complexity	- 58 -
3 Word recognition in cognitive system	- 61 -
3.1 Word processing models in psycholinguistics	- 61 -
3.1.1 Full-parsing/decomposition model.....	- 62 -
3.1.2 Full-form representation model	- 63 -
3.1.3 Dual-route processing model.....	- 65 -
3.2 Contribution of prosody to speech perception	- 67 -
3.2.1 A brief introduction to speech perception	- 67 -
3.2.2 Lexical stress in speech processing	- 68 -
3.2.3 Prosody-Assisted Processing model (PAP).....	- 71 -
3.3 Behavioral evidence of prosody processing in word recognition	- 71 -
4 Electrophysiological evidence of rhythmic processing.....	- 77 -
4.1 Applying Electroencephalogram in language processing.....	- 77 -
4.2 Language-related Event-related potentials (ERPs) components	- 78 -
4.2.1 Auditory sensory potentials (N1 & P2)	- 78 -
4.2.2 MMNs (Mismatch negativities).....	- 79 -
4.2.3 LAN & ELAN	- 80 -
4.2.4 P300.....	- 80 -
4.2.5 N400	- 81 -
4.2.6 P600.....	- 83 -

4.3	ERP evidence of rhythmic pattern processing in word recognition	- 84 -
4.3.1	Previous studies in Romance and Germanic languages	- 84 -
4.3.2	Previous studies in Chinese Mandarin	- 88 -
5	<i>Acoustic-phonetic correlates and tonal coarticulation of the Mandarin two-character neutral tone compounds</i>	- 93 -
5.1	Experiment design	- 93 -
5.2	Acoustic correlates of lexical neutral tone within Mandarin two-character compounds in the four full tone contexts	- 95 -
5.2.1	Linguistic materials	- 97 -
5.2.2	Results	- 98 -
5.2.3	Discussion	- 105 -
5.2.4	Conclusion	- 107 -
5.3	Contextual tonal variation of neutral tone within the Mandarin two-character compounds	- 107 -
5.3.1	Carry-over effect	- 109 -
5.3.2	Anticipatory effect	- 115 -
5.3.3	Conclusion	- 120 -
6	<i>Role of frequency effect and prosody pattern in the two-character neutral tone compounds processing</i>	- 125 -
6.1	Introduction	- 125 -
6.2	Goal of research and hypotheses	- 127 -
6.3	Methodology	- 129 -
6.4	ERP recordings and analyses	- 133 -
6.5	Results	- 136 -
6.6	Discussion	- 142 -
6.7	Conclusion	- 145 -
7	<i>Influence of syllabic lengthening in Mandarin spoken neutral tone compounds processing</i>	- 149 -
7.1	Introduction	- 149 -
7.2	Goal of research and hypotheses	- 150 -
7.3	Methodology	- 152 -
7.4	ERP recordings and analyses	- 154 -
7.5	Results	- 156 -
7.5.1	Behavioral results	- 156 -
7.5.2	ERP results	- 157 -
7.5.3	Comparison of Standard (Condition 1) vs. C1 deviant (Condition 2)	- 163 -
7.5.4	Comparison of Standard (Condition 1) vs. C2 deviant (Condition 3)	- 164 -
7.5.5	Comparison of Standard (Condition 1) vs. C1 deviant C2 deviant (Condition 4)	- 165 -
7.6	Discussion	- 166 -
7.7	Conclusion	- 169 -
8	<i>General discussion and conclusion</i>	- 171 -
	<i>Appendices</i>	- 179 -

Appendices 1	- 181 -
Appendices 2-1	- 183 -
Appendices 2-2	- 184 -
Appendices 3-1: related pairs	- 186 -
Appendices 3-2: non-related pairs	- 187 -
Appendices 3-3: pseudo-character pairs	- 188 -
Appendices 4	- 190 -

Introduction

The spoken word recognition involves the processing of different sources of linguistic information, which could include the phonological, orthographical, morphological, syntactic, and semantic processing (Eisner & McQueen, 2018; McQueen & Cutler, 2010; Pisoni & Remez, 2008). In the auditory modality, lexical access requires a phonologically legal match between the acoustic information provided by speech signal and the lexical representation stored in the mental lexicon (Gaskell & Marslen-Wilson, 1996). Syllable length, amplitude, and fundamental frequency (F0) are the primary concerns of prosody structure in the processing of the suprasegmental information (Cruttenden, 1997; Ladefoged & Johnson, 2014; Lehiste, 1970a). In Chinese Mandarin, there are four lexical full tones, they are the level tone 1 (Tone 1), rising tone 2 (Tone 2), dipping tone 3 (Tone 3), and falling tone 4 (Tone 4), e.g., ‘ma 妈’ (T1) (mother), ‘ma 麻’ (T2) (hemp), ‘ma 马’ (T3) (horse) and ‘ma 骂’ (T4) (scold). Besides the four full tones, through the sound weakening and pitch contour change, a kind of light stressless neutral tone (Tone 0) syllable exists in the non-initial position within the word and phrase (Duanmu, 2007; Yip, 1980). The two-character neutral tone compounds consist of a first stressed full tone and a second unstressed neutral tone, which results in a trochaic prosody pattern [Strong-weak]. E.g., within the neutral tone compound ‘**xia** (T4) **chang** (T0) 下场’ (*bad end*), the first constituent ‘**xia**’(*descend*) is in a fourth tone (T4), the second constituent ‘**chang**’(*field*) is in a neutral tone (T0). The central question addressed in the present thesis is to understand the prosody-lexical morphology interplay during the morphological complex word recognition. Depending on the previous work (Isel & Shen, 2020), we know no studies concerning the function of prosody in the online recognition of two-character neutral tone compounds in Chinese Mandarin. Moreover, we have also noticed the lack of studies on the question of morphological decomposition in lexical access of neutral tone compounds. Several questions were thus raised from the literature, whether the two-character neutral tone compounds are decomposed during auditory processing? And at what point in the word-recognition process does the neutral tone syllable play a role? Is the processing system able to anticipate the occurrence of a neutral tone in the ongoing speech information based on the acoustic analysis of the first constituent of a compound? The goal of this dissertation is to investigate the procedure of two-character neutral tone compound recognition in the auditory modality. First, we ran a series of phonetic measures on the second neutral tone within neutral tone compounds to understand the differences between the neutral tone and its full tone counterpart syllables in the four tones patterns; In addition, we examined the coarticulation effects concerning the second neutral tone with its preceding and following tonal contexts within the neutral tone compounds.

Furthermore, based on acoustic analyses, by applying the electrophysiological techniques, two ERPs studies about two-character neutral tone compounds were constructed. These two studies aimed to examine the time course of the compound recognition. Particularly, to examine the decomposition process of compound words, we questioned to which extent the tonal information of the second (rightmost) constituent, the frequency of whole compound, and the frequency of first constituent can be the crucial factors to trigger the decomposition process of a compound, and the activation of individual constituents could be constituted as a strong test of morphological decomposition. In that sense, the auditory processing of the two-character neutral tone compounds is examined by taking the lexical frequency, tonal information, and syllable length into consideration.

The way in which the multimorphemic words are stored and accessed was widely examined by researchers. One of the central questions is whether a compound word is represented as a whole unit or by the constituent morphemes in the mental lexicon (Pollatsek & Hyönä, 2005). Over the past thirty years, numerous psycholinguistic studies have been conducted (Bock & Levelt, 1994; Collins & Loftus, 1975; Eysenck & Keane, 2015; Rayner et al., 2012), all these experimental data collected with different techniques have allowed scholars to design models of compound word recognition, both in the visual (Coltheart et al., 2001) and in the auditory (Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Tyler, 1981) modality. These models make a different assumption on the nature of the lexical representation of morphologically complex words. To a large extent, these models could be classified into three main categories. Firstly, the continuous recognition theory proposed left-to-right processing associated with a global lexical representation without accessing of the word's morphological structure (Cutler, 1984; Marslen-Wilson, 1987). Secondly, sequential discontinuous recognition theory proposed the prelexical decomposition processing associated with a morpheme-decomposed representation, the lexical access of a multi-morphemic word is via its initial syllable in the auditory modality, and the first constituent is considered in this approach as code to access the lexicon (Taft & Forster, 1976). Finally, dual-route recognition theory claimed that two possible routes exist to process a morphologically complex word, (a) a direct and global route to lexical representation in the mental lexicon, (b) a decomposition route in which the compounds are accessed by its constituent morphemes after the morphological decomposition procedure, (c) both two routes operate in parallel to realize the dual-route processing (Caramazza et al., 1988; Frauenfelder & Schreuder, 1992; Schreuder & Baayen, 1995). In these latter models, the decomposition process is assumed to be triggered by the lexical frequency of the whole word. While the frequent compound words are thought to be processed thanks to the processing route which allows direct access to the mental lexicon, rare compound words or even non-lexicalized

compounds are supposed to be decomposed so that the recognition procedure is carried out through a semantic decomposition of the two constituents within the compound. Another dual-route model is the PAP model (Isel et al., 2003), which proposed that the decomposition process could be driven by prosodic information of first constituents, this model proposed that additional factors like lexical frequency may also play a determining role. The aim of the current thesis is to better understand the role of different factors in the decomposition process. Moreover, we wanted to investigate the role of tonal information of compounds in the decomposition process.

To investigate the lexical representation of morphologically complex words, a common approach is to apply a priming paradigm (Isel & Bacri, 1999). This paradigm helps us to investigate the activation of a word or a part of the words (phoneme, syllable, embedded word, constituents within the morphologically complex words). More precisely, in the cross-modal (auditory-visual) priming paradigm, participants listen to a word called “prime” and then decide on the lexical status of the visual target (i.e., a lexical decision task), which is presented during the hearing of the priming or just at the end of hearing. The lexical decision to the visually presented string can be performed after listening to a word or a sentence (Swinney et al., 1979). The reaction time (RT expressed in milliseconds) can be modulated by the relationship between the prime and target. When the target presents a relation with the prime (e.g., repetition, semantic or associative relation), the facilitative repetition priming effect could be observed (Slowiaczek et al., 1984). This method was found to be particularly useful in the study of the individual constituent’s activation in both the visual and auditory lexical recognition. Word’s prosody pattern, in particular the lexical stress of the disyllabic English words has been claimed to be involved in the lexical access and lexical retrieval during the word recognition (Cutler et al., 1997). On the one hand, prosody plays a role in lexical activation. On the other hand, the lexical prosody can be a part of the phonological code which is stored in the mental lexicon to facilitate lexical retrieval (Cutler, 1995).

The recognition of spoken words was found to be modulated by the word stress patterns (Connine et al., 1987) and the lexical tones (Fox & Unkefer, 1985). Results from listeners of Dutch (Van Donselaar et al., 2005) and Spanish (Soto-Faraco et al., 2001) showed that there is a direct use of word stress information in the lexical access, and the mis-stressing of words could prevent the lexical activation. Some studies found that prosody works in the prelexical stage of the lexical entries, which can contribute to the word identification (Cutler & Van Donselaar, 2001; Friedrich, Kotz, Friederici, & Gunter, 2004). And the extraction of the suprasegmental cues (e.g., duration) is performed by the processing system of the listeners to analyze the prosodic structure of words, and the subphonemic durational differences between different segments work as recognition cue for listeners to access the spoken compounds (Shatzman & McQueen, 2006; Spinelli et al., 2003).

The lexical frequency effect has been much investigated in the field of word recognition, and a set of research on lexical access has accounted for the effect (Grainger, 1990; Grainger et al., 1992; Perea & Carreiras, 1998; Segui & Grainger, 1990). Word frequency affects the organization of lexical search in the accessing, it serves as a crucial predictor for the performance of participants in the lexical decision tasks and examines the morphological decomposition of compound words (Forster & Chambers, 1973). In the processing of compounds, the lack of constituent frequency effect combined with a reliable whole-word frequency effect could be viewed as evidence for the holistic processing of compounds (Zhou & Marslen-Wilson, 2009). Some of the word recognition studies claimed that the Mandarin compound words are processed as a whole unit (Cui et al., 2021; Janssen et al., 2008; Taft, 2003; Wong & Chen, 1999; Zhang & Peng, 1992). The frequency of the whole-word affects the lexical decision time. In other cases, the frequency of individual constituent combined with the whole-word frequency provided evidence of dual-route processing. By varying both whole-word frequency and morpheme frequency, Zhou and Marslen-Wilson (1994) found a robust whole-word frequency effect in the lexical representation of disyllabic Mandarin compound words, and the positive morpheme or syllable frequency was not observed. By monitoring the eye movements during the reading of the right-headed Finnish noun-noun compound words, Hyönä and Pollatsek (1998) showed that the gaze duration on the compounds was 100ms shorter when there was a frequent initial morpheme than a rare initial morpheme.

This thesis focuses on the auditory processing of two-character neutral tone compounds in Chinese Mandarin. Neutral tone compounds represent 5% of the whole lexicon in the Chinese Mandarin (Li & Thompson, 1989). Mandarin is a tone language, which applies pitch variation to convey semantic meanings, different tones are applied to syllables to express lexical distinctions in the speech. Neutral tone is a kind of weak and unstressed syllable compared to full tones syllables which are viewed as strong and stressed, it shares the acoustic attribution of being a mid-target syllable, its intensity is relatively lower and its syllable duration is shorter than the full tones syllable (Duanmu, 2007). The neutral tone is always in the non-initial position within the multi-morphemic words and utterances. Within the multi-character words, there are two categories of neutral tone (Shen, 1992). In the first category, they are inherently toneless, which are mostly prepositions or suffixes (i.e., diminutive suffixes ‘*zi*(T0), 子 little’; ‘*de*(T0), 的, of’). In the second category, neutral tones are positioned within the multi-morphemic words, in which the middle of the last constituent loses its intrinsic full tone and became a neutral tone syllable, i.e., (1) the second syllable within the disyllabic reduplicated words, i.e., ‘*ma*(T1) *ma*(T0), 妈妈 mother’; (2) the second syllable within the two-character neutral tone compounds, i.e., ‘*tou*(T2) *fa*(T0), 头发 hair’ (Hua & Dodd, 2000). When neutral tone exists within a two-character compound, it is in the second (rightmost) position, with

a full tone in the initial (leftmost) position, which presents the trochaic rhythmic pattern [Strong-weak] (Lin & Yan, 1990). A typical example of a neutral tone compound is ‘*ba*(T1) *zhang*(T0), 巴掌 (hand)’, the first constituent carries a Tone 1, and the second constituent carries a Tone 0.

In Mandarin, the basic writing unit is a character, and the phonological form of a character is one syllable paired with one tone. A two-character compound is constituted by two characters (two constituents) in the writing form, however, the number of characters (constituents) could not determine its morphological structure, it may be a morphologically monomorphemic word with two constituents (two characters) in the writing form. Because only 1300 syllable combinations exist in Mandarin, each syllable might relate to any one of the four full tones or a neutral tone, thus many homophonic characters carry the identical syllable structure with different tones. Hence, the ideogram and phonological context are necessary to differentiate the homophonic morpheme in spoken word recognition (Zhou & Marslen-Wilson, 1994). Such as this pair of compounds 嘴巴 /*zui3ba1*/ (*mouth*) and 下巴 /*xia4ba0*/ (*chin*), their second constituents are the 巴 /*ba1*/ and 巴 /*ba0*/, they are the same character which carry either a Tone 1 or a Tone 0. As for the /*xia4ba0*/, the first constituent carries a strong Tone 4, and the second constituent carries a weak Tone 0. Through a combination of a leftmost strong full tone and a rightmost weak neutral tone, the trochaic prosody pattern [Strong-weak] is formed in a neutral tone compound, in contrast to a full tone compound /*zui3ba1*/ [Strong-strong]. In Mandarin, the distinction between full tone syllables and neutral tone syllables is primarily based on the acoustic cues of the syllable duration and fundamental frequency (Shen, 1993; Sproat & Shih, 1987). A typical example is the minimal pair of two compounds 东西 /*dong1xi1*/ (*east and west*) and 东西 /*dong1xi0*/ (*thing*). /*dong1xi1*/ (*east and west*) is a coordinative two-character compound, /*dong1*/ means ‘east’ and /*xi1*/ means west, and /*dong1xi0*/ (*thing*) is a non-transparent two-character compound. Both two compounds have identical syllable structures and different tone patterns on their second constituents, and the contrastive prosody patterns between the full tone and neutral tone are signaled mostly by the duration and F0. Within a neutral tone compound, /*xi0*/ has mid-fall in F0 contour, and a shorter duration (about 50%) compared to the longer duration of its full tone counterpart /*xi1*/ . In the auditory word recognition, native listeners are sensitive to the contrastive suprasegmental information contained in syllables, when hearing the phonological form ‘/ *dongxi* /’, both the neutral tone form ‘/ *dong1xi0* /’ and the full tone form ‘/ *dong1xi1* /’ would be activated, and the correct lexical representation of two compounds is activated by the neutral tone /*xi0*/ . Accordingly, the prosodic information possessed by the neutral tone serves to select among the two alternative compound candidates in the competition process. Within a neutral tone compound, the alternation of stressed and unstressed syllables constitutes

opposition in the rhythmic structure [Strong-weak], since in the phonological system of Mandarin, most of the two-character compounds are phonetically realized in the two full tones in the spondee prosody pattern [Strong-strong], and the infrequent trochaic prosody pattern of the neutral tone compounds offers us a good case to analyze the interaction of prosody and morphology in lexical access, however, as we know that fewer studies have been constructed in previous research.

In this thesis, Chapter 1 will give an overview of lexical stress, including its definition and typology, we will introduce some tonal and intonational processing models of lexical stress with an introduction to Mandarin phonology. Concerning the system of lexical tones (the four full tones and neutral tone). Chapter 2 will be devoted to an introduction of the Chinese morphology with a discussion of some frameworks of compound representation. The definition and classification of compound words will be summarized, and we will also discuss some language-specific properties of Mandarin compound processing. Chapter 3 will present the main theories of word recognition in the cognitive system, some current models and frameworks will be introduced, moreover, some behavioral and neuroimaging studies concerning the lexical access of prosody will be discussed. In Chapter 4, we will introduce some ERP components which are thought to sign different processing steps in the domain of auditory word recognition, and some electroencephalographical studies will be illustrated concerning the compound word's processing in Romantic, Germanic, and Mandarin.

Chapter 5 concerns the experimental section of this current thesis. In Subchapter 5.2, by analyzing the recordings of six native speakers, we constructed an acoustic analysis to explore the acoustic features of the second constituents in both neutral tone and full tone compounds across the four full tones contexts. In Subchapter 5.3, another acoustic study in the same method was conducted to investigate the detailed tonal variation and the coarticulation effects between the two constituents within two-character neutral tone compounds. In Chapter 6, by using the cross-modal priming paradigm and electroencephalographical technique (EEG), an ERPs study was designed to investigate whether the neutral tone compounds are morphologically decomposed or not in the auditory processing, and to what extent the lexical frequency of both the whole compound and first constituent as well as the second neutral tone syllable are factors to trigger the decomposition processing of the whole compounds. In Chapter 7, another ERP study is presented, the goal was to investigate the influence of the unexpected rhythmic patterns in the lexical processing of two-character neutral tone compounds. Particularly, this latter experiment aimed to explore the role of syllable duration in auditory processing. In Chapter 8, a general discussion will be made according to four experiments, and it will be followed by a conclusion. Based on the present study, a summary will be firstly presented, and we will then describe how this study will contribute to future study as

to explore the role and the function of prosody information in the lexical access of Mandarin two-character neutral tone compounds.

Chapter 1

Lexical stress and Mandarin neutral tone

1 Lexical stress and Mandarin neutral tone

1.1 Lexical stress in speech processing

1.1.1 Definition of lexical stress

Stress is a relative prominence between different syllables, which is considered a kind of suprasegmental feature. The definition of ‘stress’ is firstly specified in respect of metrical structure, ‘Stress is the linguistic manifestation of rhythmic structure. In stress languages, every utterance has a rhythmic structure which serves as an organizing framework for that utterance’s phonological and phonetic realization’ (Hayes, 1995). Jones (1962) has defined strong and stressed syllables as what speakers consciously produced with a greater effort than other non-stressed syllables within the words or sentences. Regarding the term stress at the sentence or the word level, Bolinger (1958) has proposed that ‘stress’ should be regarded as the word stress. Phonetically, stressed syllables are generally higher in pitch, and have a longer duration and an increased intensity compared with unstressed ones. At the level of sentences, it is preferable to be defined as a pitch accent. To simply define the ‘lexical stress’, it could reflect the rhythmical structure, which is phonetically more salient than other weak syllables.

As for the languages which have linguistic stress like English, German, Italian and Spanish, stress patterns can be categorized into three types, lexical, paradigmatic, and positional. In the first category, the placement of stress is unpredictable in the system, speakers must learn and memorize which syllables within the words are prominent. Russian stress system was cited as the lexical stress system (Melvold, 1989). In the paradigmatic stress system, stress is determined by morphological information, such as the nature of words (verb or noun), accordingly, English belongs to the paradigmatic stress system. In the positional stress system, such as the Austronesian languages, the placement of stressed syllables is determined by some specific metrical rules (Zsiga, 2012).

In the case when the position of stress within words or sentences is relatively stable and predictable, the stressed syllables could draw the listener’s attention during the listening (Beckman & Pierrehumbert, 1986; Gussenhoven, 2004). This phonetic property of stress can facilitate word recognition. Firstly, in some languages, stress could be the segmentation marker in the continuous speech, and it can interact with the listener’s attention to detect the lexical boundary and segment the speech continuum into the discrete words in speech processing. Secondly, lexical stress has a contractive phonetic function, a sequence of two syllables forms the contrastive stress placement. The different positions of stress within the same word can distinguish the semantic meanings from

one to another, e.g., the position of stress within the verb ‘PROgress’ is different from the noun ‘proGRESS’, it is the stress pattern that *guides* the word identification.

In some stress languages, like the word ‘*Apalachicola*’ in English, the metrical structure can be created by the rhythmic alternation of stressed and unstressed syllables (Rietveld & Koopmans-van Beinum, 1987). In a linguistic hierarchy that consists of multiple levels of prosodic constituents (e.g., word level and phrase level), the metrical structure can determine the position of lexical stress, which makes it predictable based on phonological feature (Lieberman & Prince, 1977). At the lexical level, stress can be phonologically encoded in the lexicon to support the word recognition process. Because stress is predictable and underlying in the rhymical framework of languages, along the continuum of lexical input, the listener’s attention is guided to the locations of stressed syllables. This makes it possible to anticipate the occurrence of the next stressed syllable, which can facilitate comprehension (Large & Jones, 1999).

1.1.2 Typological properties of lexical stress

Stress possesses a distinct phonetic characteristic, which makes stressed syllables different from non-stressed ones, and some of the typological properties make it an excellent candidate for lexical processing.

(1) Stress is acoustically culminative. Stress is salient in the phonological representation, the prominence of the stressed syllable (e.g., increased intensity, higher pitch, and longer duration) can automatically draw the listener’s attention. This is defined as the culminativity of stress (Hayes, 1995), which means that each word or phrase has a single strongest syllable that bears the main stress, this culminativity may be found in different levels of languages (Lieberman & Prince, 1977).

(2) Stress is rhythmically distributed, the syllables which bear the same kind of stress can form an alternating pattern (Selkirk, 1984). This can be illustrated by the example in Fig.1.1, when the main stress occurs near the edges of a word, i.e., the beginning of a word, the secondary stress can be located on every other syllable after the syllable conveying the primary stress.

Maranungku (Australia)			
primary stress on the initial syllable, and secondary stress on every other syllable thereafter:			
'ti ralk	saliva	'lang ka ,ra te ,ti	prawn
'me re ,pet	beard	'we le ,pe ne ,man ta	kind of duck
'yan gar ,ma ta	the Pleiades		

Fig.1.1 Example of alternating stress pattern (Hayes, 1980).

(3) Stress is hierarchical, which belongs to the multi-leveled relational hierarchy. Speakers can distinguish the multiple relative levels of stress.

(4) Stress can be a segmentation cue. Because the main stressed syllables mostly occur near the edges of words, this metrical structure can provide important cues for word segmentation,

(5) Stress doesn't assimilate. A stressed syllable does not induce another stressed syllable immediately following it, and its preceding syllable cannot be stressed (Hayes, 1995). In Liberman (1975), stress was found to be only perceived by the contrastive relationship between the stressed and non-stressed syllables, and stress was described as the relative property which arises from the relationship between two (or more) syllables, rather than a property inherent in only one syllable.

1.1.3 Assignment of lexical stress

The typology of stress distribution concerns the rules of stress assignment. In many languages, stress is predictable, the relationship between the location of stress in different words at the sentence level is rule-governed. In general, three categories of location regularities of stress in English were reported in previous phonetic studies (Clark & Yallop, 1990; Hayes, 1995). The three categories of stress location are as follows.

Category 1: Free-location and fixed-location stress

In some languages, the location of the stressed syllable is fixed and predictable, which can be derived from some rules (such as in Finnish). Once the words have been stored in the lexicon (lexicalized units), listeners could predict the occurrences of stress syllables based on the metrical structure. That is the case in Finnish in which words always initiate with a stressed syllable as well as in Turkish known to have stress on the final syllables of the words in most of the lexical units (Comrie, 1987). On the contrary, if the stress's location is not predictable by morphological rules, stress can obtain an independent role within the language system. Free-location stress is therefore unpredictable, it must occur in a specific word. However, the notion of fixed or free location stress is not a definite distinction. E.g., in Polish, the penultimate syllable (the second to last syllable of a word) is mostly stressed, yet there is a set of borrowed words without penultimate stress.

Category 2: Rhythmic & morphological stress

In a linguistic system with rhythmic stress, stress location is determined by phonological factors, such as the syllable's weight, the distance between stresses and the distance between stress and word boundary. As for morphological stress, the primary syllable or lexical root within a word is stressed and the affixes are unstressed. A system of this sort is the productive morphology of English, where the word stem carries the main stress, and most of the affixes are subordinated to the main stress (Hayes, 1995). As for the relationship between rhythmic and morphological stress,

it is admitted that the notion of 'rhythmic' and 'morphological' is in a strong interaction within the same language system, and most languages possess a mixed stress distribution system of these two.

Category 3: Bounded and unbounded stress

In a bounded system, stress falls within a particular distance from the lexical boundary or another stress. Stress can function as a boundary marker, and the placement of stress on a certain syllable is determined with reference to the word. A typical example is from Czech, a bound stress is on the first syllable of a word (Lehiste, 1973). In an unbounded system, stress falls an unlimited distance from the lexical boundary or another stress, the main stress is drawn to heavy syllables.

1.1.4 Phonetic correlates of lexical stress

Unlike the other phonological properties, the lexical stress has a phonetic realization with variability. Speaker's physical effort, e.g., the increase of muscle activities involved in respiration and subglottal pressure could influence the production. Liberman et al. (1957) realized a phonetic analysis for 16 English speakers with 24 minimal pairs, which possess the lexical stress in different positions. Results showed that 90% of stressed syllables have a higher pitch, 87% are louder and 66% have a longer duration. Another study by Cooper et al. (1960) analyzed the acoustic correlates of stressed syllables in the verb-noun pairs (e.g., CONflict - conFLICT). Contrary to the unstressed counterparts, 72% of stressed syllables tend to have a higher F0, 90% have a higher peak envelope amplitude and 70% have a longer duration. Another crucial study was conducted by Liberman and Prince (1977), which proposed the results for the relative prominences between the stressed and unstressed syllables. Accordingly, the pitch (F0), syllable duration, intensity, and vowel quality are demonstrated to be the primary auditory correlates of stress. These phonetic features combined to make a syllable more prominent than others. Compared to unstressed syllables, stressed syllables have a higher pitch, greater intensity, longer duration, and full vowel quality (Lehiste, 1970b).

Duration

Duration of stress is one of the crucial cues which is subjected to instrumental phonetic study. Syllables that receive the prosody emphasis are significantly lengthened in many languages. Based on Fry (1958), duration and intensity are viewed as cues for stress judgments, results proved that duration is an overriding cue when the effects of duration and intensity are studied separately.

Fundamental frequency

Stressed syllables consist of a wider F0 range than non-stressed ones. In many languages, a higher fundamental frequency provides a strong acoustic cue for stress judgment. Some studies

confirmed that the realization of stress can be achieved by widening the pitch range of a syllable. Fry (1958) manipulated the pitch contour of a synthetic dissyllabic word, when the magnitude of F0 was operated, results showed that the higher syllable was more likely to be perceived as stress, and a difference of 5 Hz is sufficient to influence the perception of stress. Morton and Jassem (1965) confirmed the conclusion of Fry (1958), the variation in pitch produced a greater effect on stress perception, changes from 25% to 58% in the fundamental frequency (e.g., from 120 Hz to 151Hz or 190 Hz, or from 120 Hz to 96Hz or 76 Hz) are effective to distinguish the stress.

Intensity

The intensity was proved to be less important than duration and F0 in the stress judgments (Mattys, 2000; Rietveld & Koopmans-van Beinum, 1987; van Heuven & Menert, 1996). Word distinction tests by Lehiste and Peterson (1959) were realized for word pairs with contrastive stress placements (e.g., DIgest-diGEST), which indicated that the difference of amplitudes between the stressed and unstressed syllables is approximately 2dB. Fry (1958) has examined the importance of duration and intensity in the word stress perception tests. Listeners tend to consider the longer syllables rather than the louder ones as stressed. When the pitch was considered as a parameter, listeners were more likely to consider the higher syllables as stressed. Jassem et al. (1968) used the same tasks as Morton and Jassem (1965) on the Polish speakers, they found that until a difference of 6 dB was reached, it is possible to perceive syllable stress.

Vowel quality

Stressed syllables tend to have a nonreduced full vowel quality (Bolinger, 1961), yet the reduction in vowel quality is typically associated with unstressed syllables. The vowel quality is the phonetic term for the property that makes a vowel sound different from another. In American English, contrastive stress patterns are enhanced by vowel quality. The vowels of stressed syllables tend to keep a distinct spectral quality, whereas the unstressed syllables tend to show a substantial reduction, which is produced in a less distinct vowel quality (Rosner & Pickering, 1994). Another example was found in Mohanan (2012), which reported the changes in vowel quality of unstressed syllables in Malayalam, involving both duration reduction and vowel centralization. In general, a longer and more intense syllable is more likely to be perceived as stress.

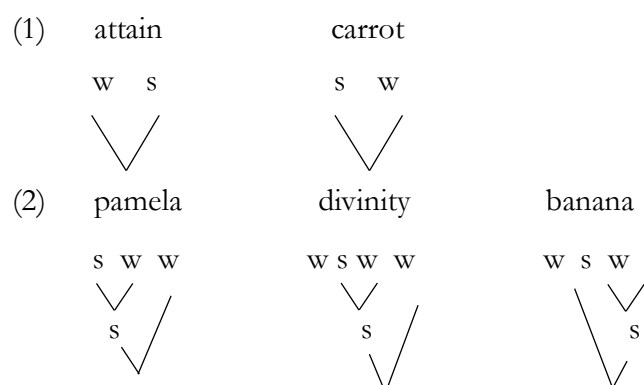
1.2 Tonal and intonational processing model

There are some important tonal and intonational models, which can be divided into two categories (Xu, 2005). The first category can be referred to as the direct pitch (F0) specification

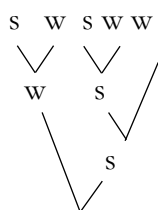
model, including the Pierrehumbert model (Pierrehumbert, 1981), the Tilt model (Taylor, 2000), and the Quadratic spline model (Hirst & Espesser, 1993), these models tried to describe the surface acoustic form, which relates the surface acoustic pattern directly with communicative function or phonological category. Their main assumption is that all the details of pitch contour are produced for linguistic specification or perception. The second category refers to the articulatory-oriented model, such as the Command-response model (Fujisaki, 1983), Soft-template model (Kochanski et al., 2003), and PENTA model (Xu, 2005), these models have focused on the simulation of the articulatory process which underlying the form of surface pitch contour. This chapter portrays the theory of Liberman and Prince (1977), Autosegmental-Metrical model (Pierrehumbert, 1981), and PENTA model (Xu et al., 2015) to provide an explanation of the tonal variation of compounds.

The framework of Liberman and Prince (1977)

In the early years, stress was viewed as a kind of absolute phonetic property that belongs to a single vowel (Chomsky & Halle, 1968). Since the work of Liberman and Prince (1977) and Liberman (1975), the metrical theory of the English stress system was well developed, it provided three hierarchical devices to represent various aspects of stress pattern in a sentence or word, stress represented as a relative prominence among different constituents at both the word and sentence level. This prominence is expressed by the metrical-tree structure, in which the pitch shape and prominence pattern are represented by nodes. A node is determined as strong because it is in the binary rhythmic relation with a weak node. Based on the timbre of constituents, two categories of nodes are firstly marked by applying the labels of ‘s’ (strong) and ‘w’ (weak) (Example (1)). For the multi-syllabic words which have more than two syllables, constituents and syllables are interpreted by their relative strength, as in (2). When a word has more than one stress, the main stress lands on the dominant syllable, and the other strong syllables receive the secondary stress, as in (3).



(3) sensibility



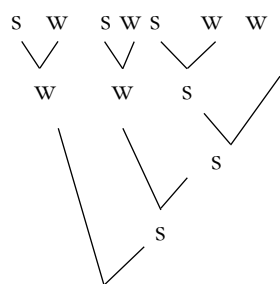
According to Liberman and Prince (1977), because of the existence of different degrees of stress, metrical stress in a language was defined as a hierarchical rhythmic structure based on the sister node pairs. The variant degree within the words is distinguished by different stress features (+ - stress) such as in (4):

(4) rabbi happy



hamamelianthemum

+ - + - + - -



In the example, a strong syllable is phonetically prominent with substantial pitch excursion, and weak syllables are conspicuously less prominent, with their energy dropping off and the pitch trailing way to the bottom of the speaker's phonetic range. There is no absolute prominence degree for 'weak' or 'strong', the concrete acoustic prominence is in terms of the relative metrical strength, the local acoustic property of individual syllables is not interpreted locally, but as a part of a relative metrical structure.

'Metrical grid' is a multi-layered sequence of beats to characterize the underlying rhymical structure of the timing of prosodic events in the metrical tree. Rhymical characteristics could be graphically represented in a bracketed metrical grid, which is placed above the words. In a grid structure, each grid marker indicates a level of prominence, and the degree of stress prominence is represented by the height of the grid column (the 'x' above the word). Grid columns are also called 'beats', a sequence of beats is used to depict the temporal structure, which is equally located in time and changed according to the syllable height and the relative prominence. The beats change

between several prominences, syllables and words generally tend to begin on the stronger beats, such as in (5).

(5) Mississippi Reconciliation

The theories of Liberman and Prince (1977) and Liberman (1975) proposed us a precise alignment between the rhymical structure and stress pattern in the metrical tree, and an alignment between the stress pattern and its timing in a metrical grid. Stress is defined as a syntagmatic feature of comparisons that cannot be distributed to a single location within the words and utterances. Prominence is thus a feature of a larger rhythmic pattern, which can comprise smaller constituents within the larger organizational structures (Beckman, 2012). The characteristics of this theory can be outlined in two points, firstly, stress is not placed to only one single syllable, it requires at least two syllables to create a stress prominence; secondly, the metrical structure is hierarchical, which possesses multiple levels of rhythmic salient structure in different features (Hayes, 1980).

Autosegmental-Metrical model

The Autosegmental-Metrical model (AM model) (Beckman & Pierrehumbert, 1986; Ladd, 2008; Pierrehumbert, 1981) postulates that tonal structure consists of a sequence of discrete local events in a segmental string. Pitch contour between the prosodic events is phonetically unspecified and can be described in terms of a gradual transition from one event to the next one. The most important events are the pitch accents and edge tones. While pitch accents refer to the prominent syllables in the string, edge tones associate with an edge (or a tone bearing unit at the edge), and they are a cue to the juncture between prosodic constituents. In this model, it is assumed that only the part of acoustic signals which directly reflects the phonological features is fully specified, the rest of them is presumed to be unspecified, and in this case, distinctive features can be presented phonologically. Only the syllables which bear the pitch accent are given the tonal targets, the weak elements don't have a specific target. The pitch (F0) of the unstressed syllables is derived from the interpolation between the adjacent pitch accents (Fig.1.2). As for weak elements, such as the schwa in English or the neutral tone in Chinese Mandarin, they are viewed as phonologically unspecified or underspecified, their phonetic features are assumed to derive from the interpolation between adjacent specified elements, or from the spreading of contextual phonological specified elements.

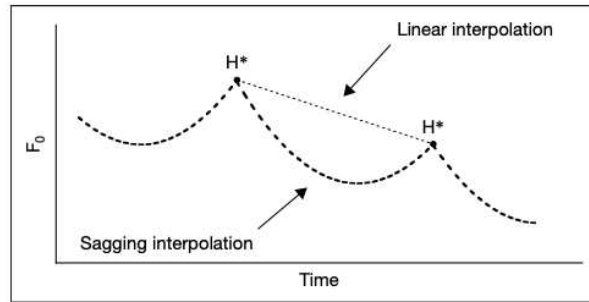


Fig.1.2 Illustration of ‘target connection’ through linear or ‘sagging’ interpolation between two H* accents as proposed in the AM theory of English intonation of Autosegmental-Metrical model (Chen & Xu, 2006a).

In the AM model, pitch accent is distinguished from stress (Beckman & Pierrehumbert, 1986). The prominence of accented syllable is a matter of metrical strength and/or dynamic stress of an individual word, it differs from the pitch accent which is the prominent syllable in the string. Pitch accent is specified into H, L, and different combinations of H and L, which are assumed to correspond directly with pitch contours, such as the F0 peak, valley, and turning point. There are six types of pitch accents, including the H*, L*, H*+L, L*+H, L+H*, and H+L* (Fig.1.3). The occurrence and transformation of phonological units are interpreted by a grammar of intonation, and communicative functions are related with the phonological units, the height of pitch contour is defined with the phonetic implementation rules (Fig.1.4). The pitch contour between the turning points is generated by the interpolation between the accentual tones. As for the over trends in the pitch contours, it is mostly reflected by the operation of localized but iterated changes in scaling factors.

$$(\%H) \left(\left(\left(\begin{array}{c} H^* \\ L^* \\ H^* + L^* \\ L^* + H^* \\ L + H^* \\ H + L^* \end{array} \right) \right)_1 \left(\begin{array}{c} H^- \\ L^- \end{array} \right) \right)_1 \left(\begin{array}{c} H\% \\ L\% \end{array} \right)$$

Fig.1.3 Illustration of tonal grammar (Beckman and Pierrehumbert (1986)).

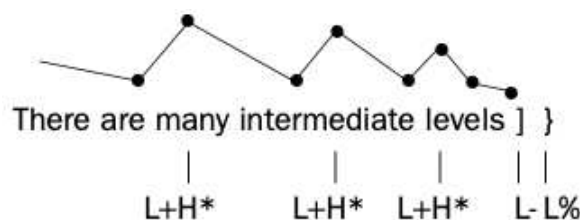


Fig.1.4 Illustration of tonal grammar in an example of a sentence (Beckman and Pierrehumbert (1986)).

Parallel encoding and target approximation model (PENTA)

PENTA model (Parallel encoding and target approximation model) is a framework about how multi-players of information are conveyed through prosody under the control of articulatory system. This model is constructed based on a tone language system (Mandarin) by a series of works (Xu, 2005; Xu et al., 2015; Xu et al., 2022). PENTA model proposed three layers between surface acoustic form and communicative functions: articulatory implementation, target assignment, and parallel encoding. In this model, the Target approximation theory (TA theory) is one of the crucial contributions, which explained the continuous tonal variation. Each syllable is assumed to possess an ideal underlying pitch target, it is not fully realized under the constraints of the articulatory mechanism. Tonal implementation is accomplished under various articulatory constraints, e.g., the speed of pitch change, and synchronization of laryngeal, and supra-laryngeal movements.

Firstly, the PENTA model has demonstrated the variation of surface tonal form between successive syllables in connected speech. The TA theory proposed that in different tonal contexts, there is an underlying trajectory contained in each syllable, which remains constant and continually approximates the adjacent following syllable. In Fig.1.5, when preceded by the second syllable in the four canonical full tones (H, R, L, F), the divergence of the third syllable is the largest at the beginning, and its converging configuration is variant. The underlying tonal target of the third syllable is constant in the four preceding tonal contexts of the second syllable, target approximation of each trajectory starts from the onset of the third syllable and ends at the offset of the third syllable.

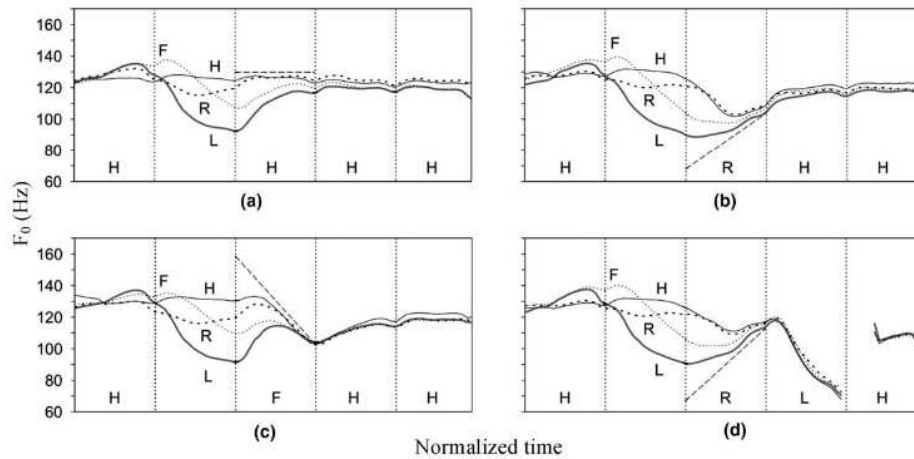


Fig.1.5 Time-normalized F0 contour of 4 sentences in Mandarin. Each sentence contains 5 syllables. In each panel, the tone of the second syllables alternates across Tone 1(H, high level tone 1), Tone 2 (R, rising tone 2), Tone 3 (L, low-dipping tone 3), and Tone 4 (F, falling tone 4), and the third syllable are held constant.

Then, Xu (2005) has proposed the syllable-synchronized sequential target approximation in the tonal implementation system (Target Approximation model), which claimed that syllables and tones are implemented in synchrony. Because a continually varied melodic pattern is realized by the articulator organs (human larynx), speakers would take a certain time to realize the transition movements, and the laryngeal movements (which generate pitch contour) must coordinate with the super-laryngeal movements (which generate segments). During the transition from one tone toward the next, the changes in the vocal fold's vibratory may influence the implementation speed of the two adjacent syllables, and the size of pitch changes will increase the transition time.

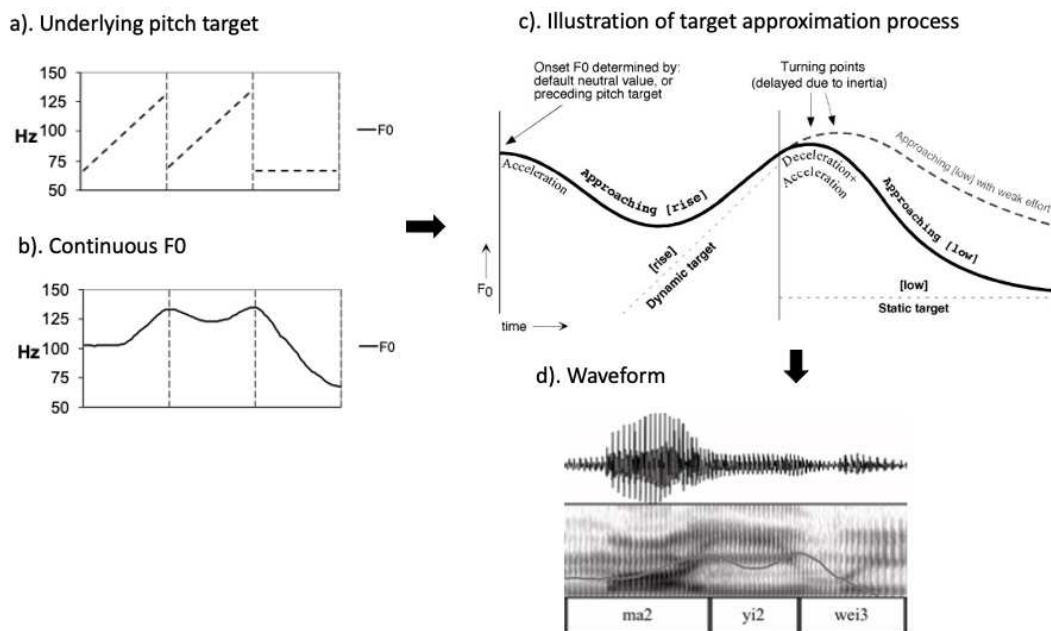


Fig.1.6. Target approximation model (Xu and Wang, 2001). 1.6(a) represents the hypothetical underlying pitch target; The continuous F0 contour is shown in 1.6(b); 1.6(c), the straight dash line is the ideal underlying pitch target, the thick curve line is the F0 contour which results from the asymptotic approximation of the pitch target, a vertical line is the onset and offset of syllables. Waveform, spectrogram, and F0 track are in the 1.6(d).

In the TA theory, each syllable is assumed to have an ideal local pitch target (either a static form [high], [low], or [mid] or a dynamic form [rise] or [fall]). Before its articulation, each syllable is assigned with both the segmental and suprasegmental goals as a control signal for the articulatory system, the aim of the implementation is to realize the pitch target in an asymptotical approach. In Fig.1.6, the pitch targets of two adjacent syllables are respectively dynamic [rise] (Tone 2, rising tone) and static [low] (Tone 3, low-dipping tone). In TA theory, the two primary implementation mechanisms are involved: underlying pitch target implementation for each syllable and syllable-synchronized sequential target approximation. In Fig.1.6, both the first syllable and second syllable are assigned with a specific underlying pitch target, the articulation starts and then implements all the pitch targets in tandem through the approximation. The global surface pitch contour over the

two syllables is a result of the sequential approximation. At the boundary between the successive syllables, the articulatory state of the offset of the preceding syllable is transferred to the onset of the following syllable, such transfer results in a delay of the surface alignment of the turning points, which can be a by-product of physical inertia (Fig.1.6).

1.3 Mandarin phonology system and neutral tone

1.3.1 A brief introduction to the Chinese phonological system

In tone languages, a crucial part of lexical representation is the tones distributed to syllables that could distinguish the semantic meanings. Mandarin is a tone language like most of the other Chinese dialects, the pronunciation of Standard Chinese is known as Mandarin, and this standard pronunciation is constructed based on the northern Chinese dialect of the capital city (Beijing). In Mandarin, the writing unit is called ‘character’ (Hanzi, ideograph). Every character has a tone pattern, and each constituent within a multi-morphemic word has a tone (Chen, 1984; Shei, 2014). A character could be a monosyllabic word (吃 /chi1/ ‘eat’) or one constituent morpheme within a multi-morphemic word (吃饭 /chi1fan4/ ‘have a meal’). In the structure of a multi-morphemic word, one syllable corresponds to one constituent morpheme, there is no ambiguity in the syllable boundary. Because of the lack of inflection and re-syllabication between constituent morphemes, most morphemes keep their primary phonetic shape in the combined syllables, there are a few phenomenon of interactions between the successive syllables, such as the tone sandhi of Tone 3 (Duanmu, 2014).

As for the syllable structure, Mandarin uses consonants, vowels, and tones to construct a syllable unit (Yip, 2002). Each unit consists of an initial consonant, a medial (on-glide), a vowel, a syllabic terminal (off-glide), and a supra-segmental tone, tone pattern of a syllable is always marked on the nuclei vowel (Fig.1.7).

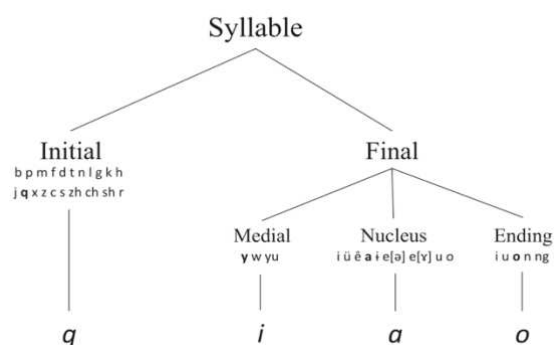


Fig.1.7 An example of the Chinese syllable structure, with all possible sounds in each level. In this example, the consonant is q, the medial is i[j], and the nuclei vowel is a[a] which can be marked with a tone pattern, and the syllabic terminal is o[o] (Shei, 2014).

A Romanized spelling system *PinYin* (拼音) is officially applied by Mandarin speakers for the annotation of sounds in Chinese script. Like the example 兔 /tu4/(rabbit), the pronunciation is represented by *tù* in italics *PinYin*, the consonant is *t* and the vowel is *u*, and the unit is in fourth tone (Tone 4, falling tone). 兔 is the character (ideograph) which is written in eight strokes.

In the phonological system, the pitch is the perceived height of F0. If the fundamental frequency is higher, the pitch will be higher. When the pitch is attached to different words, speakers can use it to distinguish the meanings, which is called the ‘tone’ (Huang et al., 2018). As a tone language, the variations of pitch contour are used to make the distinction of semantic meanings in Mandarin, so that tones could convey semantic, morpho-syntactic, or pragmatic information. In the Mandarin phonology system, there are four tone annotation markers corresponding to the four full tones: level tone 1 (T1), rising tone 2 (T2), dip-rising tone 3 (T3), and falling tone 4 (T4). Because neutral tone lost most of its tone shape, so T0 is used to indicate its pitch level (Fig.1.8).

T1	T2	T3	T4
/ma1/	/ma2/	/ma3/	/ma4/
‘妈, mother’	‘麻, hemp’	‘马, horse’	‘骂, scold’

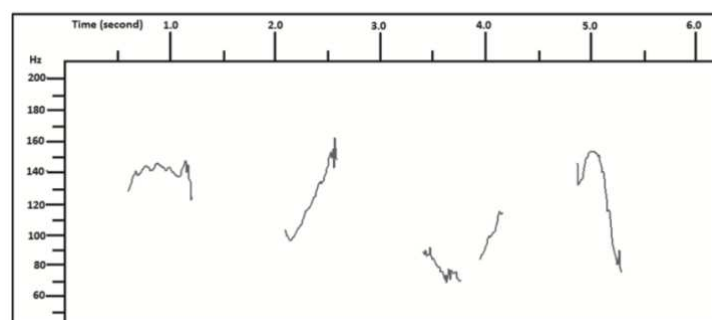


Fig 1.8. Pitch track illustration of four Mandarin full tone (Shei, 2014).

The Pitch contour of four lexical tones [55], [35], [214], and [51] are labeled as ‘high-level’, ‘mid-rising’, ‘mid-falling’ and ‘high-falling’ by (Chao, 1968), which have some variant phonological features (Fig.1.8). Tone 1 has a high-level pitch contour. Tone 2 has a rising pitch contour, it starts at a mid position and ends in a high position, slightly above the level of the first tone. Tone 3 has a circumflex pitch contour, by starting in the mid position, it falls abstractly down to low and sharply rises back to mid. Tone 4 has a falling pitch, it starts high and ends at a low position.

1.3.2 Neutral tone in Mandarin

In some stress languages such as English and Spanish, the stressed syllables underlying the metrical structure can facilitate lexical recognition. But linguistic stress in the Mandarin language system has been less discussed. Several reasons could account for this elusive nature of the Chinese stress (Chen, 1984; Duanmu, 2007; Hyman & Tadadjeu, 1976). Firstly, because there is no overall default stress value in the phonology system, even for native speakers, the relative degree of stress judgment is elusive, which is not as clear-cut as in the stress languages (Chao, 1968; Hoa, 1983). Secondly, a phonetic realization of stress in Mandarin is quite variable, the placement of stress is less ruled, and several hypotheses had been proposed to explain the stress assignment (Duanmu, 2007).

But we cannot ignore the importance of lexical stress in Chinese. A primary metrical stress representation is the neutral tone, which is closely connected to linguistic stress. Besides the four full tones, a syllable can be tone-neutralized through sound weakening and pitch contour change, they share the acoustic attributions of being the toneless weak syllables, their duration is relatively shorter, and their identity is reduced compared to syllables in the full tones. ‘Neutral tone’ is initially mentioned by Chao (1968), which is also called ‘zero tone’. Syllables in neutral tones belong to a small group of syllables, which are unstressed and weak without a fixed pitch pattern. Neutral tone appears mostly in the non-initial position within polymorphemic words and expressions, or they can occur as morphological suffixes in the sentence. Some of the neutral tone syllables possess the function to distinguish the part of speech and lexical meanings. According to the previous acoustic correlates studies, neutral tone is phonetically unstressed and light, their duration is generally 50% shorter than its full tone counterpart and 60% shorter of its preceding tone (Cao, 1992; Lin & Yan, 1990; Shen, 1992; Yip, 2002). Shen (1992) concluded that the tonal mapping of neutral tone results from the tonal assimilation, their pitch level varies which depends on the preceding citation tone, which means that the tonal feature of the preceding tone is spread from the left-to-right direction to the neutral tone, and neutral tone syllables lost most of their original tone values and become a short, mid-height and neutralized tone.

Although most of the neutral tone constituents were historically derived from the full tone counterparts, some of them are phonologically and semantically distinct from their counterparts today (Yip, 1980). For example, the noun suffix /zi0/ has a stressed full tone counterpart /zǐ3/ ‘child’ which bears a T3. However, considering some suffixes and other particles in a neutral tone, it is claimed by Shen (1992) that not all the neutral tones are derived from full tone counterparts, such as the question particles *ma*, *a*, and *ne*, which do not have any counterparts in lexical full tone. Based on the research concerning the categorization of neutral tones in Mandarin, the typology of

neutral tone syllables is mainly based on their tonal representation features between the tone and stress, and their morphological function within the sentences. Shen (1992) divided the neutral tone into two categories: underlying toneless and underlying toned. In the first category, there are some functional morphemes, such as suffixes, question particles, and genitive markers (Example (1)). Morphologically, neutral tone morphemes are attached to the notional morphemes, they can't be placed at the initial position of sentences or used by themselves. The morphemes remain inherently toneless and do not derive from any of the full tone counterparts.

- Example (1):

de genitive marker and nominalizer

le inchoative

me pause marker

er diminutive marker

zi nominalizer

ba adversative and question marker

zhe progressive marker

ma; ne question marker

a exclamation marker

In the second underlying toned category, the neutral tone morpheme is detonic, which is weak on the surface to differ morphologically and semantically from their full tone counterparts. Because the inherent tonal target of neutral tone morpheme is lost due to the absence of stress, their acoustic features are mostly determined by the preceding full tone syllable, which turns out to be underlying toned compared to intrinsically toneless neutral tone syllables in the first category (Shen, 1992, 1993). Neutral tone morpheme in the second category occurs in the polymorphemic words or the short phrases, such as disyllabic or trisyllabic compounds, they are mostly in the non-initial position. When a neutral tone exists on the second position and a full tone precedes it on the first position, a contrastive trochaic stress pattern [Strong-weak] is constituted, these words are called two-character neutral tone compounds, as the examples in Example (2),

- Example (2):

a. Tone 1 + Tone 0

玻璃 *bo li* 'glass'

b. Tone 2 + Tone 0

行李 *xing li* 'luggage'

c. Tone 3 + Tone 0

小姐 xiao jie ‘miss’

d. Tone 4 + Tone 0

漂亮 piao liang ‘pretty’

As for their acoustic correlates, neutral tone syllables in the first category are toneless and stressless, which occur mostly as clitics at the non-initial position. When they are pronounced in isolation, they are phonetically lower in pitch. In the second category, neutral tones are contextually conditioned by different preceding tonal context, their initial pitch height and duration is affected by preceding tones (Lin & Yan, 1990). Concerning the relationship between accents and lexical tones, they can interact together with the sentence intonation to modulate the semantic meanings of the sentence. In Mandarin, the stressed syllables are assigned with full tones or accents, neutral tone is not assigned a lexical tone (Duanmu, 2007).

1.3.3 Acoustic correlates of neutral tone

A neutral tone cannot occur at the initial position within polymorphemic words. According to the metrical pattern of disyllabic compounds, they are naturally divided into two main categories: ‘Strong-strong’ and ‘Strong-weak’. When two morphemes within the same word are both stressed full tones, it is a disyllabic compound with the spondee pattern ‘Strong-strong’, which constitutes most of the Mandarin disyllabic compounds. As for the dissyllabic neutral tone compounds, their stress pattern is ‘Strong-weak’, which presents a trochee pattern by a combination of two syllables.

(1) Mandarin Strong-Strong disyllabic word:

吃/chi1/ 饭/fan 4/ (have a meal)

Length long long

Reduction no no

Tone yes yes

(2) Mandarin Strong-weak disyllabic word:

本/ben 3/ 子/zi 0/ (notebook)

Length long short

Reduction no yes

Tone yes no

As for the two compounds above, in (1), both two syllables have full tones, they are long, and they have unreduced duration. In (2), the first syllable has a full tone (Tone 3), it is long and unreduced in duration, but the second neutral tone is short and reduced in duration. Compared to

their full tone counterparts at the initial position, the second neutral tones exhibit a marked degree of reduction in terms of the acoustic identity, such as the syllable duration, pitch (F0), and intensity.

Duration

The duration of the neutral tone is generally shorter than the full tone. A study by Lin and Yan (1990) compared the acoustic correlates of neutral tones with full tones within the minimal pairs. Disyllabic compounds in the phonological patterns of ‘Strong-weak’ and ‘Strong-strong’ are examined, the mean duration of neutral tone (T0) is about 50% shorter than full tones (the average duration of T1/2/3/4), the vowel centralization is a noticeable feature of neutral tone (Tang, 2014; Yip, 1980), the duration of neutral tone was found to be noticeably shorter than that of preceding full tone (the average duration of T1/2/3/4), the mean duration of preceding tone is 1.87 to 2.01 times longer than neutral tone (Lee & Zee, 2008).

Pitch (F0)

A crucial phonetic correlate of neutral tone is the fundamental frequency (F0). The neutral tone is phonetically ‘mid’ in height, the pitch is not stable, and it depends on the preceding tone. Because neutral tone is toneless and lost most of the underlying full tone target, a relatively lower pitch value is a phonetic reflection of the lack of lexical full tone (Duanmu, 2009). Concerning the pitch contour of neutral tone in continuous speech, previous studies firstly proposed that the pitch was determined by its preceding tone, which varies across the preceding tonal contexts, the tonal features of the preceding syllable spread to the neutral tone in connected speech, and the intrinsic full tone pitch target is mostly lost in the prosodic realization of a neutral tone, so that neutral tone exhibits the contextually conditioned realizations when preceded by different tones (Hoa, 1983). However, the views diverge regarding the tonal shape of neutral tones when preceded by different full tones. Gao (1980) claimed that the pitch contour of a neutral tone falls after Tone 1/2/4 and rises after Tone 3. Lin and Yan (1990) found that the pitch value of neutral tone is respectively L, L, H’, L when preceded by the preceding full tones in T1/2/3/4. By analyzing the tonal variation of neutral tone within the polymorphemic words in tonal combination patterns, the pitch contours were shown to be mid-falling after Tone 1, high-falling after Tone 2, mid-level after Tone 3, and low-falling after Tone 4 (Lee & Zee, 2008).

Since the years of 1990s, to describe the detailed phonological realization of a neutral tone, based on the dependence on the preceding-syllable hypothesis, researchers tried to construct more analysis, in regard of the tonal variation of neutral tone. Xu and his colleagues constructed a set of studies (Xu, 1999, 2004, 2005; Xu & Wang, 2001). Based on the PENTA model and the Target

approximation theory, even in the different tonal contexts, an underlying trajectory is contained in each syllable, which remains constant and continually approximates the adjacent following syllable. Chen et al. (2006) hypothesized that neutral tones have an independent tonal target in continuous speech, because of the weak strength, lower speed, and the other constraints during the articulatory procedure, the pitch target of neutral tone can't be adequately attained, which is greatly influenced by adjacent full tones in successive sequences during the target approximation procedure (Fig.1.9).

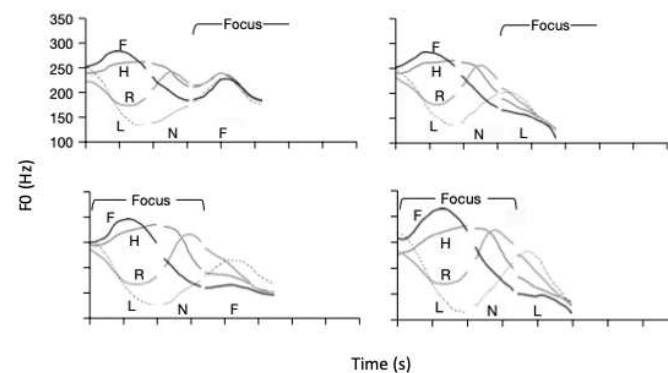


Fig.1.9 Mean pitch (F0) of neutral tone syllable (N) measured by 4 native subjects in different tonal contexts under two focus conditions (Chen & Xu, 2006a). H, R, L, and F represent respectively the Tone 1/2/3/4. In the upper figures, sentence focus is at the end of a phrase; In the lower figures, sentence focus is at the beginning of a phrase.

Intensity

The intensity of neutral tone has been revealed to be reduced compared with their full tone counterparts (Chao, 1968; Lin & Yan, 1990), which could be determined by the preceding tone. However, Lee and Zee (2008) found that the intensity of the neutral tone is not definitely lower than the preceding tones, and they suggested that the intensity co-varies with the pitch contour of the preceding full tone. The differences between the intensity of stressed and unstressed syllables in Mandarin were demonstrated to be nearly 8dB, which is used as a complementary cue with the duration in the stress judgment (Shen, 1993).

Chapter 2

Compounding in Chinese morphology

2 Compounding in Chinese morphology

2.1 Mandarin compound words

A compound is characterized by a complex word-internal morphological structure (Jarema, 2006). Lieber (2021) defined compound words as that, ‘compounds are words which are composed of two (or more) bases, roots or stems’ (page: 43). Compounding allows two or more free words to form one compound unit. According to Matthews (1991), a compound can be derived from the formation of two or simpler lexemes. Such as the compound noun ‘blackbird’, combined with ‘black’ (adjective) and ‘bird’ (noun), can be illustrated in the following formulate:

$$[X]_A + [Y]_N \rightarrow [X+Y]_N$$

In this example, we find that not only the form of the two simple words (‘black(adj.)’ and ‘bird(noun.)’) are compounding together, but also a new lexeme is derived from the compounding formation (‘blackbird (noun.)’). This productivity accounts for the type-frequency of compounds and compound types (Dressler, 2006; Fuhrhop & Endisch, 2000). The type-frequency signifies the number of compounds in a certain type and the family size of compounds within this type. The family size determines how many compound words which share the same morphemic member in the same position. Although productivity is an effective word formation method, compounds are still not fully productive, e.g., ‘blackbird’, ‘redbird’, or ‘writebird’ are not compounding lexemes.

A compounding lexeme can turn out to be a constituent lexeme in another compound word. Such as the word ‘fruit-juice carton’, the compound ‘fruit-juice’ is compounded by the noun ‘carton’ and turns into another compound noun ‘fruit-juice carton’. Another example would be the word ‘paper towel dispenser’, which consists of ‘paper towel’ and ‘dispenser’ (Fig. 2.1). From the syntactic viewpoint, there are two layers contained in the compounding structure, which might be drawn in the word tree proposed by Lieber (2021):

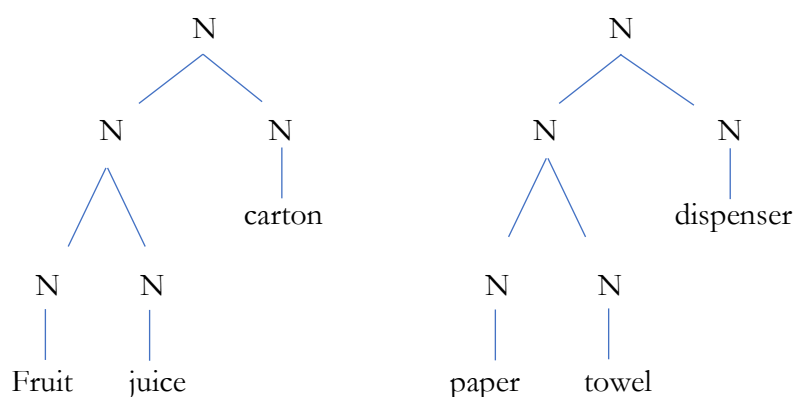


Fig.2.1. Word tree for the complex structure of compounding (Lieber, 2021).

Concerning the meaning of a compound, two aspects should be involved and specified: (1) the meaning of each lexeme; (2) the semantic relationship between the lexeme and compound. Some derived compounds turn out to be idiomatic, and their meaning compound can be totally distinct from original meanings of free lexemes (Matthews, 1991). E.g., ‘turncoat’ is constituted by ‘turn’ (verb) and ‘coat’ (noun), it means ‘deserter’ and does not relate with any of the two lexemes.

Classification of compounds is relevant to the basic properties of a word’s ‘head’. The primary classification of a nominal, verbal, and adjectival compound is defined by the word class of ‘head’, e.g., ‘blackboard’ is specified as a nominal compound depending on its head ‘board’. The same definition can account for the non-head compounds in the secondary category, such as the coordinate compounds in terms of noun-noun compounds (e.g., peace treaty, ‘a treaty to cease hostilities’). If a compound has its head within it, it is an endocentric compound (e.g., textbook), the head of exocentric compounds is outside (e.g., redhead, ‘person with red hair’) or to be inferred. The endocentric and exocentric compounds are morphologically different.

2.1.1 Brief introduction to Chinese morphology

Here we want to make a brief outline of the relationship between syllables, morphemes, and ideograms (characters) in Chinese Mandarin. The basic writing unit is a character, a character is written in several strokes. For example, the character 五 /wu5/, ‘five’ is constituted of 4 strokes. In the phonological system of Mandarin, one character corresponds to one syllable, which pairs with one tone. There are 21 consonants, 39 vowels, and 5 tone patterns (4 full tones and 1 neutral tone) in the phonological system (Shei, 2014), and about 1300 syllable combinations exist, each syllable can be attached with any one of four full tones or a neutral tone. For example, 股 /gu3/ (stock/thigh, etc.) is a character written in eight strokes, its syllable is composed of a consonant **g** and a vowel **u**, and its tonal pattern is Tone 3. 股 /gu3/ could combine with /pai2/ or /jian1jia2/ and turn to be a disyllabic (two-character) /a trisyllabic (three-character) compound word 排骨 /pai2gu3/(ribs) or 肩胛骨 /jian1jia2gu3/ (shoulder blade). When combined with /pi4/(fart), it is pronounced in a neutral tone /gu0/ within the disyllabic compound 屁股 /pi4gu0/ (fart).

Many characters with the identical syllabic structure can carry different tones, e.g., syllable /tu/ can correspond to 秃 /tu1/ (bold), 图 /tu2/ (picture), 土 /tu3/ (ground), 兔 /tu4/ (rabbit). In some cases, a character could be pronounced in two or more than two tones (polyphone character, which has several sounds), such as in the former example, 股 can be read in /gu3/and /gu0/. In this respect, we might notice the homophony problem in Chinese Mandarin, that a phonological form (a syllable) can correspond to several different morphemes (characters). Thus, in Mandarin,

there is a one-to-one correspondence between the morpheme and character, which means that each of the homophonic morphemes corresponds to only one character, the ideogram serves as a tool to differentiate homophonic morphemes. One morpheme relates to one specific character and one syllable structure. Additionally, speech context is another solution in the auditory modality (Zhou & Marslen-Wilson, 1994). When a neutral tone occurs within a compound word, due to the dependence on the preceding tone, it can be activated and accessed as the second component of the neutral tone compound in the mental lexicon (Fig.2.2).

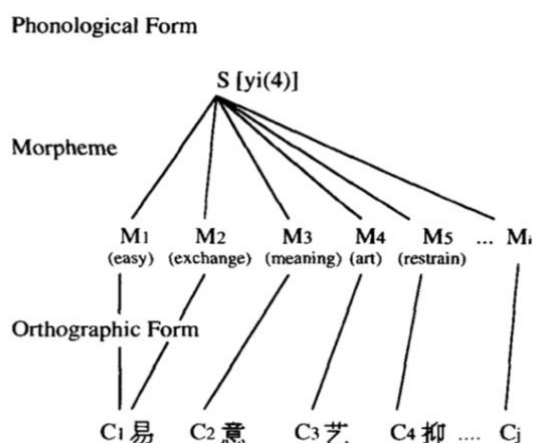


Fig.2.2 Illustration of relationship phonological form, morpheme, and orthographical form in Chinese Mandarin (Zhou & Marslen-Wilson, 1994). 'S', 'M', and 'C' represents respectively syllable, morpheme, and character. Syllable [yi(4)] consists of consonant y, vowel i, and tone pattern 4.

In the traditional view, the smallest unit which has grammatical and semantical functions could be defined as a morpheme (Hockett, 1958). In Mandarin, the basic writing unit is a character, but one character is not identical to one morpheme. When we aim to define a 'morpheme', because of the writing system (ideogram/character), the focus is the relationship among form, character, and meaning of words. A morpheme can be specified at both the semantic level (by the meaning of a lexical unit) and the phonetic level (a syllable). At the semantic level, a morpheme can be used alone as a monomorphemic word or combined with other morpheme and turn to poly-morphemic words, and no space is inserted to separate the characters within poly-character words. However, if one character doesn't correspond to one morpheme, a two-character or three-character word can be a monomorphemic word. Such as some borrowing words, 沙发 /sha1fa1/ (sofa) and 比基尼 /bi3ji1ni2/(bikini), there are two or three characters in one word, they are monomorphemic, they are not compounds because it is lack of free lexeme. The same case can be found in binding words, 蚯蚓/qiu1yin3/ (earthworm), neither of two characters within this word can have semantic

meaning when used alone. At the phonetic level, a morpheme can be distributed to exactly one syllable, a two-character/three-character word is thus a disyllabic /trisyllabic word.

2.1.2 Definition and classification of Mandarin compounds

Compounding is one of the most productive word formation ways in Chinese Mandarin, which involves a structural complexity at both semantic and morphological levels (Libben, 1998). The concept of words in Mandarin is viewed to be vague and blurred, and ‘compound’ is even quite difficult to be defined (Feng, 2017; Shei, 2014). Based on previous studies, Li and Thompson (1989) considered Mandarin compounds as ‘all polysyllabic units that have certain properties of single words and that could be analyzed into two or more meaningful elements or morphemes, even if these morphemes cannot occur independently in modern Mandarin.’ (Page: 46). Depending on their morphemic structures, M Taft et al. (1999) defined the compound as a combination of two stem morphemes. According to Shei (2014), a Chinese compound is a combination of two or more words, like the English compound words ‘strawberry’, and ‘football’.

Concerning the distinction between compound words and Chinese multi-lexical units, we may notice the possibility to insert another word between the components of a multimorphemic unit, e.g., 喝汤/he1tang1/ (drink soup) is a two-character verb-noun short phrase, we can add 碗/wan3/(bowl) between the verb and noun, which turns to 喝碗汤/he1wan3tang1/ (drink a bowl of soup). We could find that there might be a limited separability between the two components of a compound, no component can be inserted into the compounding. For example, as for the two-character coordinate compound word 开关/kai1guan1/ (‘open-close’→switch), no separability is allowed between two components of the whole compound. In the verb-object compound 毕业/bi4ye4/ (‘finish-study’→graduate), only a few words can be added between the two components, e.g., we might add the time aspect marker ‘了’ /le0/, 毕了业/bi4leye4/ (‘finish-marker of passed time-study’→graduated).

Concerning the statistical analyses, about 80% of Chinese words are polymorphemic (Lin, 1972), and over 70% of polymorphemic words used in Mandarin are disyllabic (Zhou et al., 1999). According to Lu (1981), among the 3000 commonly used Mandarin words, about 61% are two-character compound words. Compounding is one of the most important ways to form new words. Mandarin compounds possess different features in their word-internal morphological structure, componential structure, and semantic transparency.

Classification by morphological relationship between individual constituents

Based on the morphological complexity of the constituent morphemes within compounds, Mandarin compounds are classified into the three categories: subordinate, attribute and coordinate (Bisetto & Scalise, 2005; Ceccagno & Scalise, 2006; Dressler, 2006; Lieber, 2021).

1) Subordinate compound

A subordinate compound is also called ‘subject-predicate compound’, which is constituted of two components. As for the internal structure, the morphological relationship can be recursive, one member is semantically or morphologically subordinate to another.

(1). 地震 /di4zhen4/ (‘land-shock’ → earthquake), the first component (noun) is the subject of the second adjectival component (verb).

(2). 心酸 /xin1suan1/ (‘heart-sore’ → miserable), the first component is the subject, and the second component is a stative verb.

(3). 眼红 /yan3hong2/ (‘eye-red’ → covetous), the second morpheme (‘red’) is the predicate constituent of the first morpheme (‘eye’).

2) Attribute compound

An attribute compound consists of two components, the first component is the modifier of the second component (noun), which are mostly the endocentric modifier-noun compounds. Several categories exist based on their internal morphological structure:

(1). ‘Verb-Object compound’, the second constituent is the noun-object of the first verbal constituent. Chao (1968) established three rules, when a compound word accords with any one of the three rules, it could be a verb-object compound. 1). One or both two constituents are bound morphemes. 2). Idiomaticity of the compound’s meanings. 3). Inseparability or limited separability of the constituents. E.g., the compound word 效劳 /xiao4lao2/ (‘render-effort’ → serve) fits with the three rules, i.e., 效 /xiao4/ is a bound morpheme, the meaning of the compound is idiomatic, and it is not possible to insert any other constituent within the two constituents.

(2). ‘Verb-complement compound’, the second morpheme is the complement in the state or action of the first verbal component. For example, 说明 /shuo1ming2/ (‘speak-clear’ → explain), 厚实 /hou4shi2/ (‘thick-solid’ → thick).

(3). ‘Adjective-noun compound’, the relationship between the two components is viewed as ‘endocentric’, and the second morpheme is modified by the first morpheme. E.g., 飞机 /fei1ji1/ (‘fly-machine’ → plane), 香水 /xiang1shui3/ (‘fragrant-water’ → perfume).

(4). ‘Noun-classifier compound’, the second component is the classifier or the categorizer for the first component, the compounds belong to a specific category of two-character attributive

compounds. For example, 车辆 /che1liang4/ ('car-classifier'→flower), 纸张 /zhi3zhang1/ ('paper-classifier'→paper).

3) Coordinate compound

Components within the coordinate compounds present a binary relationship, which is constituted of two heads (the head of a compound is the stem that determines the semantic category of the compound). There are mainly three categories of coordinate compounds, including the parallel verb compound, parallel noun compound, and parallel adjective compound. Some coordinate compounds share the same syntactic and semantic properties of constituents; however, they are not always the same as each other. For instance, 死亡 /si3wang2/ ('dead-dead'→dead) is a noun-noun compound that is composed of two adjectives. In the lexical typology of classical Chinese, there is a predominance of monosyllabic (single character) words, like 幸 /xing4/ which is used for the example in the expression of 'fortunate'. Later in contemporary Chinese Mandarin, the number of coordinate compounds increased inasmuch as the polysyllabic tendency (Feng, 2017), e.g., 'fortunate' is now expressed by the coordinate compound 幸福 /xing4fu2/ ('fortunate + blessing'→ 'fortunate'). Depending on the relationships between the individual constituent and the whole compound, coordinate compounds can be divided into two categories:

Two morphemes share a close relativeness in meanings.

- (1). 身体 /shen1ti3/ ('body-body'→body), parallel noun compound.
- (2). 照顾 /zhao4gu4/ ('take care- take care'→take care), parallel verb compound,
- (3). 死亡 /si3wang2/ ('dead-dead'→dead), parallel adjective compound.

Two morphemes are distinct in meaning and combined to express a new meaning.

- (4). 开关 /kai1guan1/ ('open-close'→switch)
- (5). 东西 /dong1xi0/ ('east-west'→thing)
- (6). 冷热 /leng3re4/ ('cold-hot'→temperature)

Only one morpheme's meaning is kept, and another one's meaning is disappeared from the meaning of the whole compound.

- (7). 忘记 /wang4ji4/ ('forget-remember'→forget)
- (8). 动静 /dong4jing4/ ('move-quiet'→movement)

Classification by semantic relativeness between components and compound

Relationship between the meanings of a compound and the meanings of its constituents is varying from semantically non-related to totally related. Based on the semantic relativeness

degree between the individual component and the whole compound, three idiomatic classes of Mandarin compounds are proposed by Li and Thompson (1989).

1. Meaning of individual components is not semantically related to the whole compound, there is no apparent semantic correspondence between constituent and compound, compound's meaning is highly idiomatic, they are semantically opaque compounds.

- (1). 大意/da4yi4/ ('big-idea' → 'negligent')
- (2). 风流/feng1liu2/ ('wind-flow' → 'amorous')
- (3). 小说/xiao3shuo1/ ('small-speak' → 'novel')

2. Meaning of individual components is inferentially connected with whole compound, compound's meaning is not straightforward connected with its constituents, and a metaphorical relationship exists between them. They are only partially transparent compounds because only the meaning of one constituent relates to the meanings of the whole compound.

- (1). 天气/tian1qi4/ ('sky-gas' → 'weather'), only the meaning of 'sky' relates to 'weather'.
- (2). 电影/dian4ying3/ ('electric-shadow' → 'movie'), only the 'electric' is related to the meaning of the whole compound.

3. Meaning of individual components is directly related to the compound. An apparent close correspondence exists between them. They are semantically fully transparent compounds.

- (1). 干净/gan1jing4/ ('dry-clean' → 'clean')
- (2). 城墙/cheng2qiang2/ ('city-wall' → 'city wall')
- (3). 呼吸/hu1xi1/ ('exhale-inhale' → 'breath')

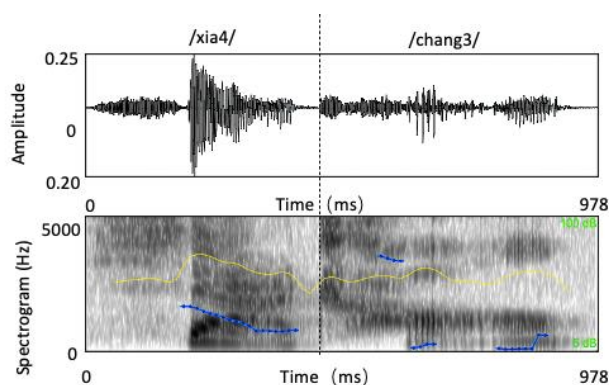
2.1.3 Two-character neutral tone compound

A set of disyllabic Mandarin compounds possess a neutral tone in the second position, they are called the 'neutral tone compounds'. They are constituted of two constituents, the first constituent is in one of the four full tones, while the second constituent bears a weak neutral tone. About 2000 two-character Mandarin compounds possess this trochaic prosody pattern, these compounds constitute a small group of disyllabic compounds in the Mandarin lexicon, which constitutes about 5% of compound words (Lin, 1972). According to lexical phonology, the second neutral tone within disyllabic compounds can be divided into two categories, including optionally neutral tone compounds and obligatorily neutral tone compounds. In the optional category, the second morpheme of a disyllabic word is optionally neutral-toned. For example, 把手(handle) can be pronounced both as /ba3shou3/ and /ba3shou0/. In the obligatory category, the second constituent must be pronounced in a neutral tone, and there are two subtypes in this obligatory

category. In the first subtype, neutral tone occurs at the second position of a word as the lexical suffixes (e.g., 子/zi0/, 头/tou2/), e.g., 儿子/er2zi0/(son). In the second subtype, a neutral tone exists in the final position of a compound like in 耳朵/er3 duo0/(ear). In this latter compound word, the first constituent 耳/er3/is in Tone 3, the second constituent 朵/duo0/ is in Tone 0, and it is a transparent dissyllabic compound, which possesses a rhythmic pattern [Strong-weak]. The alternation of strong full tone and weak neutral tone creates opposition in rhythmic structure within the dissyllabic neutral tone compounds.

As a tone language, tonal variation and prosodic information contained in the four lexical tones and neutral tone are exploited by Mandarin listeners to distinguish different morphemes and convey the semantic meanings. As for the second neutral tone within dissyllabic compounds, they are historically derived from the full tone counterpart morpheme through tone-neutralization, but most of them are both phonologically and semantically different from their full tone counterparts. Such as compound 下场 /xia4chang0/ (bad end), the second constituent can be pronounced in a third tone pattern (Tone3) and construct a full tone compound [Strong-strong] 下场 /xia4chang3/ (go off the stage) (Fig.2.3). Within these two compounds, the first constituent is the same /xia4/, and the second constituent is the same syllable /chang/ carried either a third tone or a neutral tone. The contrastive pitch target between the strong full tone and weak neutral tone is signaled mostly by the duration and pitch (F0). Within the neutral tone compound, /chang 0/ has a shorter duration and mid-fall in pitch contour (Fig.2.3(2)), in contrast to its full tone counterpart /chang3/ in Fig.2.3(1), which has a longer duration. In auditory processing, listeners are sensitive to the contrastive suprasegmental information contained in different syllables. When listeners hear the phonological form /xiachang/, both a neutral tone representation /xia4chang0/ and a full tone representation /xia4chang3/ are simultaneously activated, and the correct word representation is activated by the second neutral tone /chang0/.

(1)



(2)

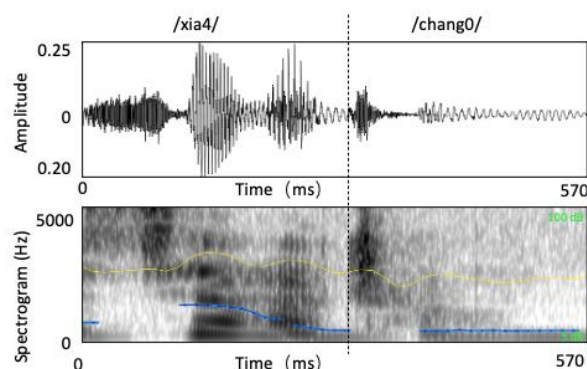


Fig.2.3 (1). (A). Full tone compound xia(T4) chang(T3) in citation form. (2). Neutral tone compound xia(T4) chang(T0) in citation form. The waveform is represented on the top part, and the spectrogram (gray scale), the F0 (blue line), and the intensity (yellow line) are represented on the bottom part.

2.2 The framework of compounding processing

2.2.1 Multilevel interactive-activation model

Adapted from the Interactive-activation model by Seidenberg and McClelland (1989) and McClelland and Rumelhart (1981), a set of studies by Taft and Zhu (1997); Zhou et al. (1995); Zhu and Taft (1994) and Taft and Zhu (1997) were constructed for the multilevel interactive-activation framework, which explained the morphemic processing of Mandarin compounds. The focus of the framework is on the hierarchical activation of compound words, morphemes, and sub-morphemic units in lexical processing. Chinese morphemes and words are represented both at the orthographical and phonological levels. As in Fig.2.4., the orthographical representation is described on the left side, from down to the top, a lexical unit is composed of strokes, which are viewed as the ‘features’ of characters. When some strokes are combined into radicals, the activation of the constituent (character) can be facilitated by the radicals contained in the morpheme. And morpheme could be associated to support the recognition of polymorphemic units, as to finally reach the semantic activation threshold of the whole compound. On the right side of Fig.2.4, the phonological representation is linked with both the morphemic level and polymorphemic unit (compound) level. During the visual recognition of compounds, activation on the orthographical level is hierarchically realized from the lowest stroke level to the highest meaning level, connection between these different levels is influenced by several language-related properties, such as the lexical frequency (Zhang & Peng, 1992). High-frequency compounds are activated more quickly than low-frequency compounds because of the stronger link between the levels. This could be

associated simultaneously with the phonological representation, the phonological information is activated in the time course of this process (Perfetti & Zhang, 1995).

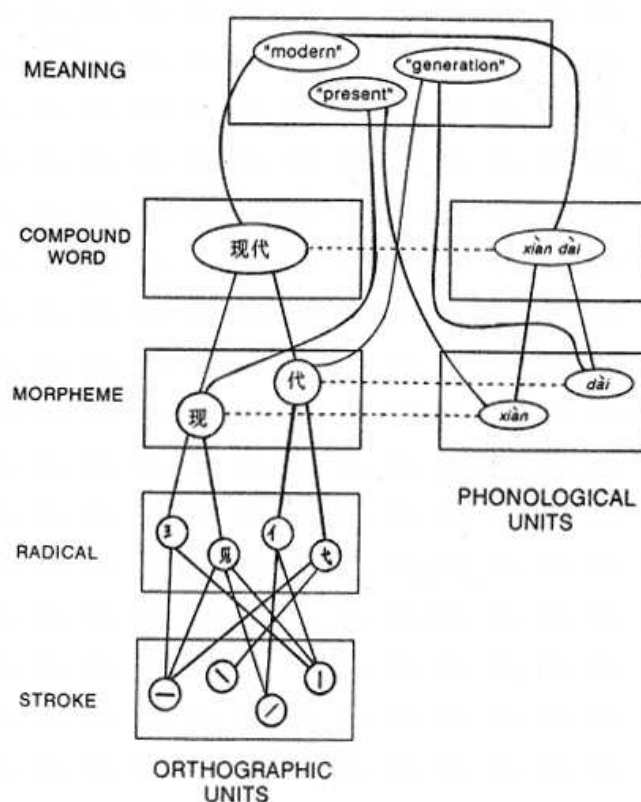


Fig.2.4. Illustration of activation of Mandarin compound ‘现代/xian4dai4/’ in the Multilevel interactive-activation architecture (Taft, 1994).

To precisely explain the phonological activation (in both morpheme and compound level) related to the orthographical level, the model was developed by adding the level of ‘lemma’, which is a kind of abstract concept node or hidden unit (Seidenberg & McClelland, 1989). In this case, the lemma serves as the modality-free unit and as a connection of the orthographical unit (character), phonological unit (syllable), and semantic unit. Lemma captures the relation between input form and meanings. As in Fig.2.5, an orthographical unit represents a character, each lemma unit can represent a morpheme, a phonological form, or a polymorphemic word. During the visual processing, the character 马(/ma3/, horse) can send activation to the units in the lemma which represents the morpheme, the pronunciation /ma3/, the meaning ‘horse’, and the semantic units which are horse-related, as well as the lemma units which represent the polymorphemic units with the character 马, e.g., compound word 马掌(/ma3zhang3/, horseshoe) (the combination of 马 (horse) and 掌 (horseshoe)), and lemma units which contain information to distinguish 马掌 from

the other polymorphemic units. Lemma can automatically develop the appropriate links to realize the processing of morphemes.

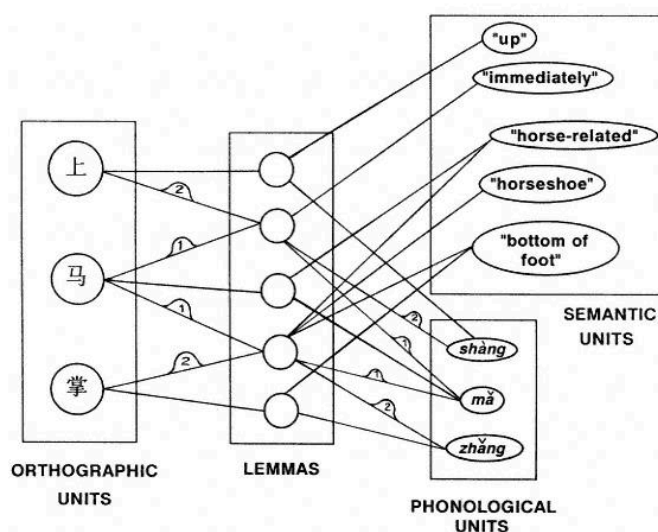


Fig.2.5 Illustration of multilevel interactive-activation model (Zhu & Taft, 1994). Example compound word 现代/xian4dai3/(modern).

2.2.2 Multi-layer representation model by Zhou & Marslen-Wilson (1999)

A group of studies was conducted on this model, e.g., Zhou and Marslen-Wilson (1999a, 1999b); Zhou et al. (1999); Zhou and Marslen-Wilson (1994, 1997, 2000); Zhou et al. (1995). They have investigated the lexical processing of Chinese compounds in the mental lexicon. The main purpose of studies is concerned with the representation of compounds in the mental lexicon, about how lexical knowledge is accessed, as to investigate the interaction of morphological, orthographic, and phonological information in compound recognition.

Based on these works, a multi-layer lexical presentation model was developed, and three layers exist in this model, including the morphemic layer, phonological layer, and semantic (whole compound) layer. In these three layers, the representations of compounds and constituents interact and overlap with each other. The representation in the orthographic layer for each constituent links directly to both the phonological layer and semantic layer, and for the representation of the compound in the semantic layer, contains the semantic representation of each constituent. The morphological structure of a compound is marked at the three representation levels, an excitatory connection exists between constituents and compounds during the lexical activation, and an inhibitory connection exists between homophonic constituents at the morphemic level.

During the visual recognition of a compound word, morphological structure impacts the interaction between the form and meanings of each constituent. The visual input is guided by their morphological structure and mapped into the mental lexicon, and the morphological structure of a compound is decomposed into the smaller constituent units to activate the orthographic form and phonological form. Semantic representation exists for both constituents and compounds, and the semantic activation of compounds is dependent on the constituent. The degree of semantic features shared by compounds and the constituent morphemes reflects the semantic transparency or opacity of the compound, it determines the excitatory or inhibitory effects between the different representation levels.

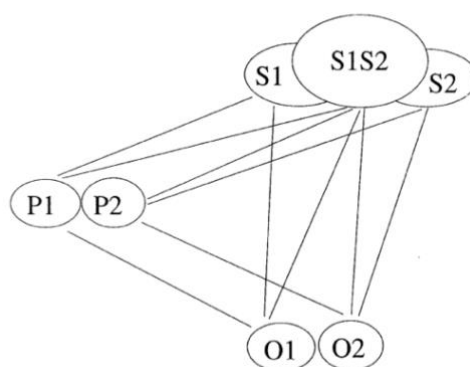


Fig.2.6 Multi-layer model of Mandarin compound representation by Zhou et al. (1999). In this illustration, ‘O’, ‘P’, and ‘S’ are prototypical for orthographical, phonological, and semantic representations. ‘O1O2’ represents a two-character compound, e.g., 小心, ‘P1P2’ represents the phonological form /xiao3/ and /xin1/, ‘S1S2’ represents the semantic meaning ‘be careful’, which overlaps with the meaning of S1 ‘small’ and S2 ‘heart’.

In this multi-layer representation model, constituent morphemes which share the same phonological form are linked to the same phonological form (prototypical: P) and compete with the other homophonic constituents, the access of visual input (prototypical: O) at the word level is succeeded through morphological decomposition to mark the constituents’ boundary (Fig.2.6). As for the representation in orthographical layer and whole-word layer, Zhou and Marslen-Wilson (1997) defined this relationship as semantic in nature. The relationship between the representations of morphemes and the representation of words depends on the semantic features which are shared by them, depending on the semantic compositionality (the degree of transparency) of compounds. The phonological form (syllable) corresponds to both words and morphemes, and the semantic representation of words and morphemes is arranged at the same level (Zhou & Marslen-Wilson, 1999a).

2.3 Language-related property in Mandarin compound processing

Compound word processing can be affected by some language-related properties, some previous studies explored to what extent the recognition of compounds in the mental lexicon can be affected by these linguistic properties. This sub-chapter is organized into four language-related properties, including the semantic transparency, lexical frequency effect, morphological family size, and orthographical complexity in Chinese compound processing.

2.3.1 Semantic transparency

Compound words are stored in the mental lexicon, due to the idiosyncratic changes brought by lexicalization, the meanings of compounds are not directly from their components. If there is more lexicalization, the compound will be opaquer, and less lexicalization means more transparency. In this respect, compounds can present different degrees of semantic transparency, which can be related to lexical productivity, transition probability, and family size (Myers, 2006). Based on Libben (1998), there are four fundamental degrees of morpho-semantic transparency, 1). Transparency of both two constituents of a compound, e.g., doorbell; 2). Transparency of the head constituent, e.g., strawberry; 3). Transparency of the non-head constituent, e.g., jailbird; 4). The opacity of both two constituents, e.g., humbug.

As for psycholinguistic research, one of the major questions is whether the individual constituent of compound words is activated during lexical processing. Accordingly, the meaning of an individual constituent within a compound word will be recognized faster due to its direct connection with the meaning of the whole compound. While opaque compounds are turned to be represented as a whole unit, the meaning of individual constituents is unrelated to the meaning of the entire compound. The semantic unpredictability of constituents raises the fundamental issue of the lexical representation of opaque compounds. In the auditory modality, the semantic opacity of compound words can hinder lexical access. Studies of semantic transparency about compound word processing address the question of whether the individual constituents of a compound are activated during the word recognition. By manipulating the opacity degree of morphemes within compounds, we could investigate if the opacity can hinder lexical access. According to previous studies, the semantically transparent compound words turned out to be represented as separated constituent morphemes, and compounds that are composed of transparent morphemes can be recognized more quickly than opaque morphemes. For example, Sandra (1990) found that in German, the semantic priming effect was observed for the semantically transparent and partially transparent compounds and not for fully opaque compounds. Jarema et al. (1999) found that in

French, transparent compound words are processed much faster than opaque compounds with equivalent lexical frequency, however, the reverse result was found in Bulgarian.

As for the Mandarin compounds, some research obtained the same conclusion (Myers et al., 2004; Tsai, 1994), that lexical access to compounds is facilitated by semantic transparency. For example, as for the compound word 马虎/ma3hu0/(careless), its constituents 马/ma3/ and 虎/hu0/ respectively mean 'horse' & 'tiger', both the two constituents don't correspond to the meanings of the whole compound, this opaque compound would be responded slower in a lexical decision task than the semantically transparent compound 马车/ma3che1/ (horse-drawn vehicle). However, some other studies found the reverse results or no effect of the semantic transparency (Su, 1998). Regarding the other properties in lexical processing, the lexical frequency effect was reported to interact with semantic transparency in the compound access. The frequent and opaque compounds are assumed to be processed more efficiently in a full-listing way; The infrequent and transparent compounds are assumed to be decomposed in the lexical processing. When semantic transparency is controlled, a positive morpheme frequency was found for transparent words than for opaque words (Myers, 2006).

2.3.2 Lexical frequency effect

Word frequency effect

When the whole-word frequency is an experimental variable, it provides the facilitative effect in the lexical decision tasks, words that have a higher lexical frequency can be responded to more rapidly (Forster & Chambers, 1973; Myers, 2006). In auditory word processing, the lack of a constituent frequency effect combined with a reliable whole-word frequency effect can be viewed as evidence of the holistic processing of compound words. Some studies found that Mandarin compounds are treated as a whole unit at the recognition level (Cui et al., 2021; Janssen et al., 2008; Taft, 2003; Zhang & Peng, 1992). which have found that whole-word frequency has a facilitative effect on the lexical decision time, which is not the case for the character or constituent frequency. E.g., by varying the word frequency and the morpheme frequency, Zhou and Marslen-Wilson (1994) reported a robust word frequency effect in dissyllabic compound processing, which means that frequent compounds were identified faster than infrequent compounds in the word recognition, and the facilitative morpheme frequency or syllable frequency are not found, similar results were found by Wang et al. (2005), which showed a robust whole-word frequency effect in the online processing of two-character compounds.

Character frequency

Psycholinguistics and neurolinguistic evidence have brought support to the activation of an individual constituent in Chinese compound processing. Previous research has investigated the morphological processing of Mandarin two-character compounds by manipulating the morpheme frequency in the word recognition tasks, character frequency was demonstrated to modulate the processing of infrequent rare disyllabic compounds, and robust evidence was found for both the initial morpheme and the second morpheme (Cui et al., 2021; Packard, 1999; Rayner et al., 2012; Wang et al., 2005; Zhang & Peng, 1992). Zhang and Peng (1992) found that as for coordinative compounds (bi-headed), the frequencies of both two constituents can affect the processing, the recognition of disyllabic compound words involves an online sub-lexical decomposition into the constituent lexeme. Similar results have been found by Zhou and Marslen-Wilson (1994) and Zhou and Marslen-Wilson (1995), the robust word frequency facilitatory effect was proved, and results showed that the frequency of individual characters plays an important role in the recognition of polymorphemic words. However, in Packard (1999), no morpheme or syllable frequency effect was found in the Mandarin spoken compounds processing, which suggested that the individual morpheme is not systematically activated in the online processing, and lexical access of spoken compound is at some point as morphological whole-units.

2.3.3 Morphological family size

Morphological family is defined as the set of all words which contain this given constituent as their morpheme, including when this constituent appears by itself as a monomorphemic word, and for the compounds which share this given constituent as the same members (Schreuder & Baayen, 1995). The size of the morphological family was found to play an important role together with the lexical frequency in the word recognition (Burani & Caramazza, 1987; Burani et al., 1984; Schreuder & Baayen, 1995; Taft, 1979). Taft (1979) has reported that the visual word recognition time of English monomorphemic words is codetermined by the whole-word's frequency and the inflectional variant's frequency. Similar results were found for words in Dutch (Baayen, Dijkstra, et al., 1997) and in Italian (Baayen, Burani, et al., 1997), the summed frequencies of the singular form and the plural form could determine the recognition speed of a monomorphemic noun, and even the frequency of a derived monomorphemic word could codetermine the recognition latency in visual lexical decision task. In Mandarin word recognition, when a character can appear in many words, the readers and speakers must obtain more information about the word's internal structure to realize the word recognition. Within the multi-character words, characters with a large family size could potentially activate many other words in the same family that will compete for word recognition. In contrast, as for the characters that possess a smaller morphological family, these

characters could exist in fewer words, and in this case, less information about the word's internal structure is demanded in the recognition, and fewer resources are needed in the representation (Myers, 2006).

2.3.4 Orthographic complexity

Chinese orthographical representation involves the processing of writing form in visual recognition. Character is the basic writing unit, and a character is written in several strokes. Stroke number and the number of radical components (radical is constituted by several strokes) could determine the orthographical complexity degree of the different characters, which could affect the visual recognition time. Previous studies revealed that the increased stroke number of characters can slow down the recognition time (Peng & Wang, 1997; Yu & Cao, 1992). These results suggest that readers take more time to identify the characters written in more strokes. Moreover, another important concept of Chinese characters is radical, a radical consists of several strokes, and the radical component has been found to influence the recognition of characters. Firstly, the simple characters (without any radical components) were found to be recognized faster than compounded characters (with one or more than one radicals). Moreover, both the recognition time and naming are generally shorter in high-frequency radicals than in low-frequency radicals (Taft & Zhu, 1997; Zhang et al., 2012; Zhu & Taft, 1994). Secondly, the category of radicals contained in characters was found to affect the identification process. When the prime characters and the target characters share identical semantic or phonetic radicals, for example, when the prime 江 and the target 滴 share the same radical 水, a priming effect was found in the character recognition tasks (Chen et al., 2006; Feldman & Siok, 1999; Marcus Taft et al., 1999; Zhou & Marslen-Wilson, 1999a).

Chapter 3

Word recognition in cognitive system

3 Word recognition in cognitive system

3.1 Word processing models in psycholinguistics

Psycholinguistics assumes the existence of a mental lexicon, in which the lexical knowledge of a speaker is stored in the long-term memory (Colé et al., 1997; Segui, 1992, 2015). According to Segui (2015), words constitute the fundamental interface between the formal and conceptual levels. Moreover, some authors like Levelt (1992) proposed that a lexical unit is constituted of different linguistic properties, e.g., phonological/spelling, morphological, syntactic, and semantic, the properties are represented mentally, and they could be accessed for language understanding or production. One of the most used experimental paradigms to investigate the nature of the lexical representation and its access is the priming paradigm with a naming and lexical decision task. Other paradigms or experimental tasks, like the phonological lexical decision task, allow us to study the specific mechanism involved in reading, namely the conversion grapheme-to-phoneme.

The comprehension of the mechanisms, which underpin the reading system has been one of the primary questions for psycholinguistics. Word recognition is defined as the process enabling the identification of the word's form and meanings from the input of letter strings. Because the precise processing procedure which leads to the mapping between the spelling and phonological information and with semantic representation is not straightforward, it needs to be considered with precision. Compounds are a kind of morphologically complex words, which consist of more than one morphological constituent. One of the central questions concerning the morphological complex word's representation in the mental lexicon is whether compounds are represented as constituent morphemes and/or whole units in the mental lexicon (Isel et al., 2003). To recognize a spoken compound, like a monomorphemic word, listeners must first solve the computational problems at a sub-lexical level. Then, at the lexical level, different cognitive processes are involved to analyze the different linguistic properties inherent to each lexical unit, namely the phonological, orthographical, morphological, syntactic, and semantic access (Marr & Vaina, 1982). Concerning the mental representation of morphologically complex words, like compounds, at least two main categories of lexical representational views were proposed in the psycholinguistic literature (Isel et al., 2003). In the first view, the so-called decomposition approach (i.e., the full-parsing theory), the morphologically complex words are thought to be accessed through the separated morphemes or segments in the mental lexicon, and they are segmented in reading. In the second view, the full listing theory (whole-word), morphologically complex words are represented as the whole lexical units in the cognitive system. According to both the full-parsing theory (morpheme-segment) and

the full listing theory (whole-word), three categories of compound processing models are proposed, including the decomposition model, full-form representation model, and dual-route models.

3.1.1 Full-parsing/decomposition model

In the decomposition approach, rather than the full form, complex words must be encoded into their constituent parts in the lexical access. The initial stage of decomposition is to extract the word's morpheme, word processing is through the stem of polymorphemic words. Decomposition models assume that several levels exist in the mental lexicon, including the morpheme and word. The constituent morphemes of compound words are stored in the mental lexicon, and a prelexical decomposition into these morphemes exists during compound word recognition.

Taft (1979, 1981, 1985) and Taft and Forster (1975, 1976) have proposed this well-known affix-stripping model, which depends on the prelexical decomposition hypothesis. In this model, only the root morpheme is assumed to be coded in the mental lexicon, a polymorphemic word is hypothesized to be processed through its root morpheme, and the decomposition occurs at the prelexical process stage which segments the multimorphemic string into roots and affixes. For example, RE could be stripped from the words REVIVAL or REPAY automatically at the affix-stripping procedure. To process the incoming speech signals in the visual or auditory modality, a certain part of the input will be recognized as the stem of a word in the lexical entry, and the other parts will be viewed as affixes. In this approach, an affix can be considered as a code to access the lexicon in which roots are stored.

Another type of decomposition model can be termed the fully-decomposition model (Taft, 2004). The hypothesis of the fully-decomposition model is that, not only root or stem morphemes, but also each constituent morpheme of the complex word is accessed and represented in the word processing (Fig.3.1). In the full-decomposition model, each morpheme will be decomposed, which could lengthen the processing time course of compound recognition.

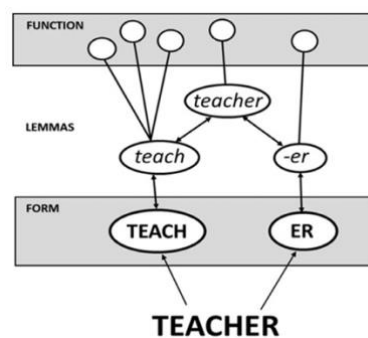


Fig.3.1. Obligatory morphological decomposition model is illustrated with the derived word 'teacher', it is represented by a lemma that is activated via lemma units representing its component morphemes (Taft et al., 2018).

3.1.2 Full-form representation model

Full-form representation model relies on the continuous processing view which postulates that words are represented in the mental lexicon as the entirety. The multimorphemic words are processed in a left-to-right direction, and the information contained in the input is continuously mapped into the mental lexicon, regardless of the internal structure of words. This view can be related to the full-listing hypothesis, that morphologically complex words are stored and accessed/recognized using a direct route. This underlines that morphologically complex words have an individual full-form representation that is stored in long-term memory. Several well-known word recognition models for auditory processing are the Cohort model (Zhou & Marslen-Wilson, 2000), the TRACE model (McClelland & Elman, 1986), and the Shortlist model (Norris, 1994).

The Cohort model is a full-listing model which is proposed by Marslen-Wilson and Welsh (1978) and Marslen-Wilson (1984) for online spoken word recognition. This model emphasized the ‘left-to-right’ nature of speech recognition. During the analysis of the 150 first milliseconds of the speech signal, the Cohort model postulates that the lexical representations which match with the incremental portion of the signal would be activated. These activated words constitute the so-called “initial cohort of words”. Then, progressively as the signal is incrementally presented in a left-to-right mode to the language processing system, based on the continuous mapping between the acoustic-phonetic information and the stored lexical representation, the representation of the word which is no longer compatible with the incoming input would be deactivated. The initial cohort will be progressively reduced because of the ongoing mapping. Finally, it will remain only one candidate in the cohort corresponding to the best match between the lexical representation and the segment of a speech signal, then the word recognition is realized. In English, the word CROCODILE is taken as an example by Taft and Hambly (1986), in the auditory modality, this word cannot be recognized any earlier than the ‘recognition point’ /d/, because there are no other English words that begin with ‘CROCOD’. Since the combination of phonemes CROCO could be related to the word CROCKERY, only CROC can’t be enough to realize word recognition. During the processing of the auditory input, the goodness of fit of the letter (or phonemes) strings determines the activation of word candidates, which helps to eliminate the other competitors.

McClelland and Elman (1986) have proposed the TRACE model, which is one of the most widely applied spoken word recognition models. It is based on the interactive activation framework of McClelland and Rumelhart (1981). The TRACE model is organized into three levels, feature, phoneme, and word, and there are feature detector/phoneme detector/word detectors in each of the three levels, and a number of units also exist at the feature level. Auditory input is presented as successive slices into the feature level, and the input is transmitted to the units, and the units

are involved in the processing of all the slices. The word “TRACE’ is used to specify the network of the units, because the spoken input is activated as the ‘trace’ within the three levels. In this network, the activation of phoneme nodes is realized from a series of feature detectors to the units, and then to realize the word activation through these phoneme features in the word level (Fig.3.2).

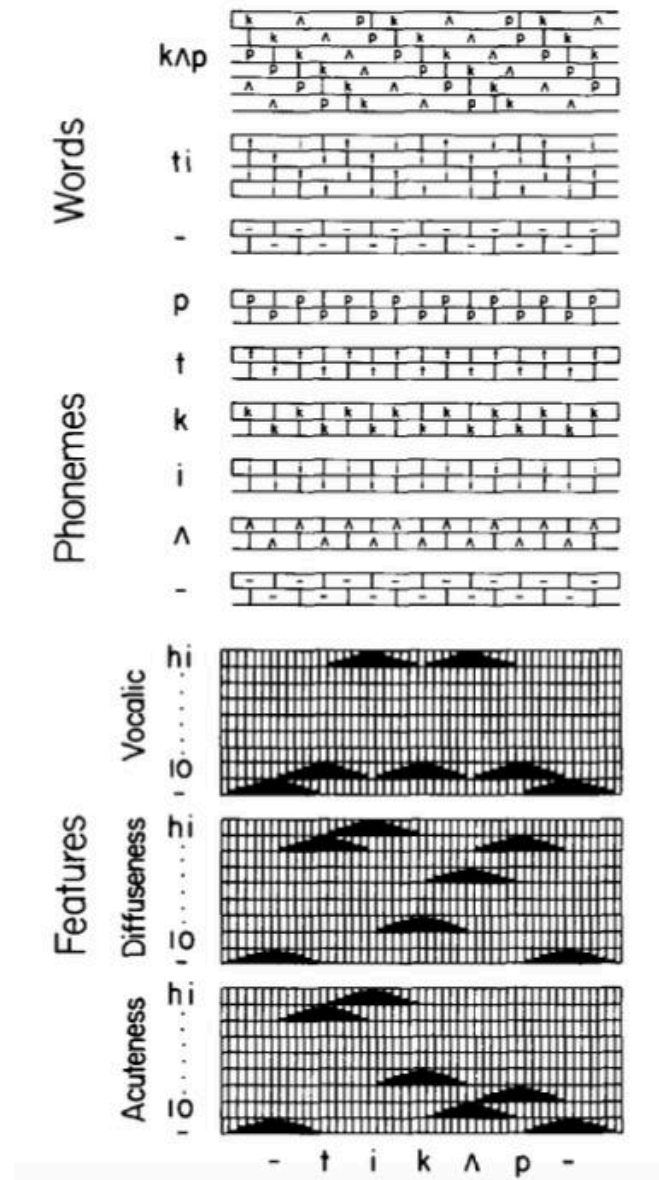


Fig. 3.2. Illustration of TRACE model (McClelland & Elman, 1986). The units are represented as the black rectangle in the figure. The auditory input is the word ‘teacup’. The horizontal labels indicate three levels, and the horizontal edge represents the portion of the TRACE spanned by each unit.

The Shortlist model (Norris, 1994) is a localist connectionist model. The model consists of two stages. Firstly, a short list of word candidates which match the auditory input will be derived in the exhaustive search of the first stage. In the second stage, these word candidates which have overlapping in the phonemes are wired into the constrained networks, they are connected because

of the overlapping in their phonemes, and the inhibition from the overlapping will eliminate the word candidates. The number of phonemes shared by the candidates determines the weight of inhibition links, and the inhibition effect is greater when more phonemes are overlapped. The activation of the word candidate node is realized by the bottom-up activation, which is determined by its goodness of fit with the auditory input. For example, when a spoken word CATALOG enters the auditory input, each of the phonemes /kæ/, /tə/, and /lɒg/ will be presented through the lexical search to update the candidate and wire the lexical network. The bottom-up activation score will also be updated by the input. Through some competition cycles in the lexical network, the activation of a candidate is realized.

3.1.3 Dual-route processing model

Both the prelexical decomposition and whole word access exist in the hybrid mode, such as the Augmented Addressed Morphology Model (AMM) (Caramazza et al., 1988; Caramazza et al., 1985), the Morphological Race Model (Frauenfelder & Schreuder, 1992) and the Dual-route model (Hyönä & Pollatsek, 1998; Pollatsek & Hyönä, 2005; Pollatsek et al., 2000). These models accounted for morphological decomposition based on different factors, one of the most important ones is the lexical frequency of the compound word. More recently, some authors proposed that besides the lexical frequency, some other factors might also play a determining role to trigger the decomposition route in dual-route architecture. E.g., Isel et al. (2003) proposed that the prosodic information provided by the initial constituent of the two-nouns transparent German compounds can be a factor of morphological decomposition. Based on the Addressed Morphology Model, the frequently used compounds are processed in the full-form representation route, whereas the rarely used compounds are accessed in the decomposition route. In the Morphological Race Model, both two processing routes work in parallel. ‘Race’ is assumed in the representation of morphologically complex words between the decomposition and whole-word processing, some polymorphemic words are more likely to be decomposed than others. Dual-route processing approaches consist of the direct whole-unit representation accessing and the decomposition representation processing. Polymorphemic words can be processed through a direct full-form route, which leads to a whole-word representation, and/or it could be accessed through the decomposition route that their constituent morphemes are activated.

In these models, complex words can be represented in dual routes, both decomposition route and/or direct whole-unit route. One issue correlated with these dual-route models is whether complex words are processed in both two routes, or whether only through one route and the alternate one is rejected. And more questions can be involved in this debate, e.g., if both two routes

can be selected, they operate independently or interactively, and how to determine which route operates at first in the processing, etc. In these respects, a crucial question raises according to the dual-route model is, whether a sort of polymorphemic word is preferentially processed via one route rather than the alternate one. The interaction effect of morphological features and lexical properties can produce some influences on the representation of complex words. Depending on previous studies, the word's internal structure and some lexical properties (e.g., word/morpheme frequency, semantic transparency, and prosody pattern) might determine which one of the routes the cognitive system will choose (Isel et al., 2003; Isel & Shen, 2020; Semenza & Mondini, 2010).

Previous studies examined the factors which might affect the processing of compounds in the dual-route models. As for the morphological structure of compound words, derived words are claimed to be processed in a whole-unit approach, whereas the inflected words are processed in a decomposition way (Niemi et al., 1994). From the view of morphological productivity, researchers claimed that a compositional approach will be applied in the lexical access by the cognitive system regardless of the impact of semantic transparency. And some researchers have found evidence of decomposition in the opaque compound processing (Badecker, 2001), lexical frequency can alter the recognition time course in the morphemically complex words, which might be correlated with 1) morpheme frequency. 2) whole word frequency. 3) both morpheme and word frequency. High-frequency compounds are more likely to be processed in the direct route than the low-frequency compounds. Moreover, semantically transparent compounds, that means, compounds in which each of the constituents possess a semantic relation with the meaning of the whole word (e.g., /'blakbɜ:d/, *blackbird* in English), would be processed in the decomposition route through their constituents (Mäkisalo et al., 1999). However, the triggering of the morphological decomposition of semantically transparent compounds also depends on the lexical frequency of compounds. That means, high-frequency compounds will be recognized as a whole unit in holistic processing, with no need to extract the constituent morpheme to access their meanings. More generally, concerning the interaction of lexical frequency and semantic transparency in compound processing, both the high-frequency opaque and transparent compounds would be accessed in the whole word units. On the other hand, low-frequency compounds would be processed more efficiently through the decomposition of their constituents. The semantic properties of the compounds will emerge from the combination of the meanings of each constituent because of the mechanism of the semantic recomposition (Libben, 2006).

3.2 Contribution of prosody to speech perception

Many studies have shown the role of prosody in spoken word recognition, and prosody was suggested to assist in the retrieval of segmental phonology. In the early stages of lexical access, word prosody could work as a framework to facilitate lexical segmentation at an acoustic-phonetic level, and word prosody is important in word retrieval (Levelt, 1992). Listeners can use the prosody knowledge stored in the mental lexicon of the cognitive system as to use them to process the input signals and realize the lexical representation (Cutler & Clifton, 1999). In this chapter, based on the brief introduction of the speech perception and lexical stress processing mechanism, the Metrical Segmentation Strategy (MSS) (Cutler & Norris, 1988) and the Prosody-assisted processing model (PAP) (Isel et al., 2003; Isel & Shen, 2020) will be illustrated, and some behavioral evidence from the previous studies concerning the prosody processing in the word recognition will be stated.

3.2.1 A brief introduction to speech perception

Speech perception is to apply the combined sources of segmental and suprasegmental information to recognize spoken words (Eisner & McQueen, 2018). In the continuous speech, the acoustic pattern of a speech sound might be influenced by the surrounding context and present variability. Listeners must segment the individual word in the incoming auditory input and extract correct words which might have embedded words within them (Repp & Liberman, 1987). Three main problems exist considering the complexity of input signals, including the variability problem, segmentation problem, and lexical embedding problem. With this respect, a model was proposed by Eisner and McQueen (2018) to explain how online input information is interpreted by listeners.

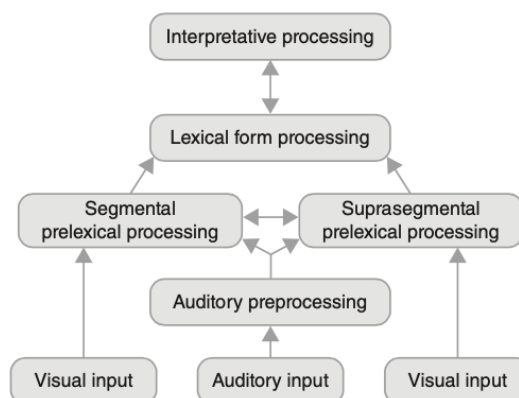


Fig.3.3. Illustration of speech perception, the flow direction of input information is marked by the arrows (Eisner & McQueen, 2018).

In Fig.3.3, three stages in speech processing are illustrated for the auditory/visual input processing. In the first stage, segmental prelexical processing means that segmental information, which is stored in the long-term memory, could be applied to realize the lexical retrieval in the computational process. Then in the second stage, suprasegmental prelexical processing refers to the analysis of the suprasegmental information contained in the word, which can assist the word retrieval process. For example, in English, the word **forbear** can be pronounced in different stress patterns depending on whether it is a noun (FORbear) or a verb (forBEAR). The contrastive stress patterns work as a crucial suprasegmental cue in the retrieval system. In the third stage, the correct phonological form of the words will be selected and recognized from the potential candidates depending on the information obtained in the latter stages. Lexical form processing distributes this information to interactive processing, including the syntactic, semantic, and pragmatic accesses to finally perceive the word's meanings.

3.2.2 Lexical stress in speech processing

The information of metrical structure is systematically stored in the mental lexicon of listeners, stress is represented as a part of lexical entry which facilitates word recognition (Cutler, 2010). The salient stressed syllable attracts the attentional resources in the online processing, the interaction between stress and the listener's attention can influence cognition. In this sub-chapter, two frameworks are illustrated to discuss the processing of stress in speech processing: Dynamic Attending Theory (Jones, 1976) and Attentional Bounce Hypothesis (Pitt & Samuel, 1990).

Dynamic Attending theory

Dynamic Attending Theory is originally proposed by Jones (1976) and then completed by a series of works (Mathewson et al., 2010; Schroeder & Lakatos, 2009). This theory assumes that a sequence of tones could present a regular rhythm and entrain attentional oscillations, which may assist the rhythm processing. The temporal predictability of auditory information can entrain the listener's attentional resources. Many kinds of sounds possess temporal regularities, such as music and some kinds of speech. Temporal regular stimulation has been demonstrated to generate strong temporal expectations for upcoming speech signals (Jones, 2010). The expectations are assumed to be created when sensory signals have a regular temporal pattern, which can guide the listener's attention to the rhythmic points. The expectations can be created by the regular auditory stimulus, which is considered a crucial part of input processing. Studies in the later stage have observed the preferential processing of the temporal expected signals (Large & Jones, 1999). The regularities in the signals can anticipate the predictable events in the auditory processing, rhythmic regularity can

modulate the perception of the subsequent target words, and the processing of stimulus words presented in the same rhythm will be facilitated by the rhythmic regularities.

Dynamic Attending Theory model postulates that the perception of metrical information is a result of the interaction between the acoustic input and the intrinsic neural dynamics, some additional evidence has demonstrated that the neural oscillations could be modulated by rhythmic regularity of stimulus (Henry & Herrmann, 2014; Schroeder & Lakatos, 2009; Lakatos et al., 2013; Arnal & Giraud, 2012; Rohenkohl & Nobre, 2012). Here, the neural entrainment refers to the presentation of temporal regular input, the expectable stimulus was found to be better perceived than non-expectable stimulus. As for the complex rhythm with multiple metrical layers of stimulus, in which each layer has a specific temporal regularity, the neural entrainment will be assumed to distribute the selective attentional resources to track the predictable events in each metrical layer, which forms a united hierarchical perception.

Attentional Bounce Hypothesis

Attention Bounce Hypothesis was firstly proposed by Pitt and Samuel (1990). According to the hypothesis, listener's attention is oriented automatically to the salient event in the continuous speech, and the predictability offered by the metrical structure of speech will enhance this effect during the word processing. The regular occurrences of stressed syllables will anticipate the future occurrences of the next stress, and the temporal regularity of metrically strong syllables facilitates the word segmentation. Studies that aimed to test this model were constructed by using phoneme monitoring tasks. For example, in English, the stress of the noun 'PERmit' is on the first syllable, and the stress of the verb 'perMIT' is on the second syllable, listeners were instructed to detect the target phonemes [p] or [m] in this pair. Here, the acoustic input is manipulated to put the two phonemes in an equal level of salience, behavioral results showed that the expectation of phoneme location can influence the decisions, the syllables which were presented in the previous rhythmic contexts as stressed facilitated the lexical identification, listener's attention could be preferentially distributed to the stressed syllables.

Metrical Segmentation Strategy (MSS)

In continuous speech, there is no reliable marker of a word boundary, and the problem of word segmentation arises for the listener's lexical processing. To confirm the position of a word boundary, listeners tend to use the metrical structure information from their previous linguistic experiences, as to mark the word onset. In the stress languages, such as English, according to the rhythmic pattern, there are stressed and unstressed syllables. Stressed syllable is strong with a full vowel and unstressed syllable is weak with a phonetically reduced vowel. In this case, the metrical

structure of the stress language is stress-based, words are mainly initiated by the stressed syllables. According to Cutler and Carter (1987), in English, about 90% of content words are either monosyllabic (stressed) or polysyllabic with their stressed syllable in the initial position. In English, stress can be considered a reliable prosodic feature to identify the onset of a word, and consequently to mark the left boundary of the words.

Cutler and Norris (1988) proposed the Metrical Segmentation Strategy (MSS) based on the metrical structure of English words. MSS predicted that listeners turned to insert a word boundary when they encountered strong syllables. In the word-spotting experiments, some real words (e.g., *thin/mint*) are embedded within the dissyllabic sequences, which end with either a strong syllable (e.g., *thintayf/mintayf*) or a weak syllable (e.g., *thinef/mintef*). It was shown in native listeners of English, that the average detection times are significantly longer for the embedded word *thin* from *thintayf* than from *thinef*. To explain this result, researchers proposed that the segmentation of *thintayf* is prior to the second syllable, which means that *thin* is segmented as *thin_tayf* within the *thintayf* which ends with a strong syllable, whereas no segmentation exists within the *thinef* which ends with weak syllable. According to the prediction of MSS, this means that the recognition of *thin* is apart from the other parts of speech material, however, the detection of *thin* in *thinef* is less difficult, because the second syllable is weak (schwa), which is described as non-independent and doesn't have to be segmented from preceding materials (Cutler, 1990). In continuous speech, listeners are assumed to apply this segmentation strategy to realize the word segmentation, stressed syllables are often considered as the onset of spoken words. Listeners search in the lexicon when encountering strong syllables and refraining from their research with weak syllables, the lexical access will be initiated by strong syllables.

Metrical segmentation strategy is a stress-driven accessing model, which can be viewed as a discontinuous lexical processing strategy, spoken words are not processed strictly from left to right, and strong stressed syllables will trigger the lexical access. A few empirical studies have been conducted concerning the MSS to explore the segmentation strategy in the lexical access (Cutler & Butterfield, 1992; Vroomen & De Gelder, 1995). Some studies analyzed the misperception of word boundary location in connected speech. Listeners were presented with the sentences which were out of the perception threshold, e.g., an audible sentence that consists of some altering strong and weak syllables (e.g., *conduct ascents uphill* → *the doctor sends her bill*). The results have found that more erroneous word boundaries were placed before the strong stressed syllables, and listeners tended to delete the word boundaries before weak syllables.

3.2.3 Prosody-Assisted Processing model (PAP)

The Prosody-Assisted Processing model (PAP model) is a compound processing model, which has been proposed by Isel et al. (2003), Isel et al. (2010), and Isel and Shen (2020). In Isel et al. (2003), four behavioral priming experiments with a lexical decision task were conducted to investigate the auditory processing of German noun-noun compounds, and the prosodic structure of the first constituent was manipulated. Results showed that the prosodic structure of the first (leftmost) morpheme of compounds works as the determining decomposition factor in the lexical activation. The PAP model is a dual-route cascading model, both two accessing routes (direct route and decomposition route) are postulated in the hybrid framework. These two accessing routes work in parallel in word recognition. In the auditory modality, the direct whole-unit representation is assumed as the default processing route. When the auditory input is sequentially unfolded, the prosodic analysis of the initial (leftmost) morpheme will trigger and assist the decomposition of a compound word. And among the acoustic parameters of the initial morpheme, the duration of the first syllable is viewed as a crucial role in the lexical access. When the listener encounters a spoken compound, the prosodic information contained at the onset of the word may firstly activate the abstract storage in the mental lexicon, simultaneously, an acoustic analysis of the first morpheme will be processed. If the length of the first morpheme could be perceived as the first syllable of a compound word by the listener's cognitive system, the decomposition processing approach will be automatically triggered to activate the morpheme within a compound word. In contrast, when the duration of the first morpheme is not matched with the length of first syllable of a compound, the direct whole-unit processing will be inferred by the cognitive system, which may be oriented to the monomorphemic word's accessing. Accordingly, the output of this acoustic analysis will notify the processor to determine the processing route, the decomposition route of the PAP model is assumed to be hierarchical, and a syntactically prelexical decomposition route is possible to exist in the discontinuous processing. E.g., the head morpheme of transparent noun compounds will be accessed in the primary position.

3.3 Behavioral evidence of prosody processing in word recognition

Spoken word recognition involves the procedure to activate the correct word among the various potential candidates. In tone languages, the suprasegmental information which contained in words could convey different semantic meanings, and the internal prosodic structure of words has been found to constrain the lexical activation in the early stage, which can assist the activation of the correct representation among several candidates (Cutler, 1995). Some behavioral studies

examined the role of prosodic structure in the early stage of word recognition, and one of the mostly applied experimentation methods is to measure the response time of participants in lexical decision/detection tasks. By manipulating the prosodic structure of auditory stimulus, researchers want to observe the effects of an ill-formed word's prosody on the lexical processing, and the main goal of this research is to understand how the prosody structure influence the word recognition process (Cutler et al., 1997). In this sub-chapter, some behavioral research that concerned the mis-stressed prosody pattern and mismatched syllable length in word recognition are illustrated.

Mis-stressed prosody pattern

Speech signal provides both segmental information and suprasegmental information, such as pitch, amplitude, contrastive durational differences between stressed segments and unstressed segments. Suprasegmental information is used by the listeners during spoken word recognition. Listeners are sensitive to the small subtle differences between the segments in the contrastive stress pattern, the contrasts are used to constrain the lexical access. The syllables within a word may have different phonetic realizations in variant prosodic contexts, e.g., the same segment would be longer and higher when it is stressed. This kind of variability is processed in the suprasegmental prelexical processing, in which the specific prosodic features are extracted to entail the abstract stress pattern, which helps the listeners to cope with auditory processing in the early stage of word recognition.

The mis-stressing effect was found in the stress languages, such as Dutch and English. In Dutch, the metrical structure of words is stress-based, the strong stressed syllables mark the lexical segmentation in the speech stream (Cutler & Norris, 1988). When the words were deliberately mis-stressed, their recognition process is delayed (van Heuven & Menert, 1996). Certain types of stress misplacement tend to impair word recognition. Cutler and Clifton (1984) switched the position of the stressed syllables within disyllabic compounds, e.g., *canTEEN* and *TURbine*. Behavioral results showed that the misplaced substitutions impaired the word identification of native listeners, the application of suprasegmental information in the stress-pattern processing has helped the listeners to solve word embedding and segmentation problems. Goodglass and Wingfield (1997) conducted the research with aphasia patients, word prosody can prime phonological access.

By using the cross-modal word fragment priming tasks, the word prosody was reported to modulate the behavioral response's latency (Cutler & Van Donselaar, 2001; Soto-Faraco et al., 2001). Participants' responses to the target word in the stress matching prime–target pairs (e.g., muS–muSEUM) are faster than to the target word in the stress mismatching prime–target pairs (e.g., muS–MUSIC). Soto-Faraco et al. (2001) found that mis-stressing targets inhibit the behavior results. In a Spanish study, by using the disyllabic fragments, slower responses were observed for the targets with mismatching prosody (e.g., *princi*–*PRINcipe*) than for completely unrelated targets

(e.g., *mosqui*–*PRIN*cipe), results indicated that the lexical representation of partially mismatching targets receives inhibition from the completely matching candidates (e.g., *prin**CI*prio [beginning]).

The evidence of the functions of stress in word recognition has been provided in previous studies, e.g., in Dutch (Reinisch et al., 2010), Italian (Tagliapietra & Tabossi, 2005), and Spanish (Soto-Faraco et al., 2001). Listeners have been found to be sensitive to the sub-lexical unit stress information in Dutch, suprasegmental stress is applied to resolve the competition problem, and they can distinguish *voor* taken from the word *VOORnaam* (first name) and *voor**NAAM* (respectable) (Cutler & Van Donselaar, 2001), Dutch listeners were found to use the stress information in the early stage of lexical recognition, before the delivery of the specific segmental disambiguation information. For example, in this eye-tracking study, *oktober* (October, stress on the second syllable) and *octopus* (octopus, stress on the first syllable) are accessed by listeners before the occurrence of the distinctive consonants /b/ and /p/ in the speech signal (Reinisch et al., 2010). The recognition of visually presented words is facilitated by prior auditory primes with the same first two syllables when they were appropriately stressed, e.g., *OKTOBER* can be primed by *okTO-*, in contrast to the inappropriate stressed primes (e.g., by *Octo-*), which produced inhibition in the recognition process (Van Donselaar et al., 2005).

Previous studies in Dutch have also confirmed that listeners use stress to resolve inter-word competition problems. For example, in Dutch, *zee* was detected more quickly in the word *luzee* than in the word *mu^zee* (which activates the Dutch word *museum*). However, when *mu^zee* is pronounced with a stress on the initial position as *MUsee*, the stress pattern is no longer matched with the word *muSEum*, it no longer inhibits the detection of the word *zee* (Cutler & Van Donselaar, 2001). Italian listeners relied on the lexical stress cues in online recognition. By using a cross-modal priming paradigm, listeners were tested with a lexical decision task on visual targets, results showed that the reaction time could be facilitated by these spoken bi-syllabic primes when the targets and primes are in the matched stress pattern, e.g., the target word *GOmito* (‘elbow’) is facilitated by the prime *GOmi* (Tagliapietra & Tabossi, 2005). Similar results were found for Spanish listeners by using the same cross-modal paradigm in Soto-Faraco et al. (2001), lexical stress is used by Spanish listeners in the same manner as segmental information to constrain word access.

Mismatched syllable length

Most of the spoken word recognition models claimed that the acoustic signal is decoded into a string of phonemes to be accessed in the mental lexicon. The effect of syllable duration on word recognition has been observed in previous studies. During lexical processing, the extraction of the suprasegmental cues is adopted by listeners to recognize the syllabic structure of a word or between the words to solve the lexical-embedding and syllabic segmentation problems. In some

cases, the different durations of the segments could be applied to disambiguate the similar-sound words. When two words are constructed by identical segments, the lexical competition problem might arise due to the variant competitive candidates in the mental lexicon, and the duration of segments can facilitate the recognition of correct segments and determine the word boundary in the continuous input of speech signals.

In English, when a word is monosyllabic, its duration is longer than when it's embedded within a polysyllabic word, e.g., *cap* is longer than in *captain* (Lehiste, 1972). And there are some ambiguous sequences, e.g., *tulips* and *two lips*, the /l/ are found to be longer on the initial position than on the non-initial position, listeners use this acoustic cue to recognize ambiguous sequences (Gow & Gordon, 1995). The subphonemic durational difference across multiple segments works as a recognition cue for listeners to process the ambiguous two-word sequences in Dutch. The different lengths of consonant /s/ within *een spot* (a spotlight) and *eens pot* (once jar) are used by listeners in the spoken word recognition (Shatzman & McQueen, 2006). Another study by Quené and Kager (1992) examined the sequence pairs like *diep in* (deep in) and *die pin* (that pin), Dutch listeners use the durational differences of consonant /p/ to segment word pairs. In French, the duration of consonant /r/ is extracted and employed for the distinction of liaison, results from cross-modal priming experiments revealed that /r/ in the liaison environment '*dernier oignon*' is shorter than consonants in the word-initial position '*dernier rognon*' (Spinelli et al., 2003). And the duration of the sub-lexical unit is also employed as an acoustic cue by Italian listeners to segment speech and access the lexical sequences which have the same initial syllables. For instance, with respect to the duration of /l/ in the final position of the first syllable like in *silvestre* is quite different from the /l/ in the initial position of the second syllable like in *si.lencio*, listeners could disambiguate the syllables between the minimal pairs without hearing the following segments (Tabossi et al., 2000). Evidence from some languages suggests that the information on syllable duration can be effectively exploited in the word recognition (Zhou & Marslen-Wilson, 1994). Salverda et al. (2003) conducted an object-naming study by monitoring the participant's eye movement. The participants were instructed to click on the object which was mentioned in the phrase, and the length of the first syllable of the object's name was manipulated by replacing them with monosyllabic words.

Chapter 4.

Electrophysiological evidence of rhythmic processing

4 Electrophysiological evidence of rhythmic processing

4.1 Applying Electroencephalogram in language processing

Electroencephalogram (EEG) is the electrophysiological recording of the oscillations of brain electric potentials preserved from the electrodes on the human scalp. A single electrode could observe an estimation of the postsynaptic actions averaged over the tissue masses which contain roughly 100 million to 1 billion neurons (Nunez & Srinivasan, 2006). Several neurons will be activated at the same time to create electrical fields which are large enough to propagate through the human brain. Scalp EEG recording is an important tool to observe and treat certain kinds of illnesses. Additionally, a major task for cognitive scientists is to determine the correlations between the potential recorded measurements and the underlying psychological processing.

Previous research in the psycholinguistic domain has revealed a correlation between EEG signals and human cognitive processes, such as mental calculation, working memory, and selective attention. The electrical activities are divided into two categories, the first category is spontaneous potentials, such as the alpha and sleep rhymes. Another category is constituted by the event-related potentials (ERPs), which means electrical potentials that are related/time locked with exogenous or endogenous events (stimuli). Some exogenous stimuli, e.g., a light flash or a piece of auditory recording, could be used to measure the event-related potentials. ERPs provide a direct temporal measurement for the state-dependent brain processing of these stimuli, the small changes in the electrical activities of the human brain are recorded from the scalp, and the activities change rapidly and have a spatially extended field. Averaged ERPs can measure the electrical potentials due to the postsynaptic activities of neural ensembles, which are generated in the extracellular fluid flowing across neurons (Kutas & Dale, 2013). The sensory evoked potentials could involve the auditory, somatosensory, and visual modalities in the language processing (Nunez & Srinivasan, 2006).

ERPs can provide a detailed quantitative measurement in the research of diagnosis and treatment of central nervous diseases, and an accurate temporal solution is important to study the brain activities from one millisecond to the next. As ERPs can provide the temporal dynamics of neural activities in response to the stimuli which evoked them, this makes it an ideal tool to observe the effect of a certain experimental manipulation. Because language processes devoted to different linguistic information occur quickly, ERPs could record the neural activities to explore the natural language processing, as well as the prediction research, because ERPs could provide a continuous monitor of brain responses that cover the full course of a discourse. The peaks of stimuli-locked ERP waveforms, i.e., the delay between the stimulus onset and the time at which the maximum electrical response (maxima) is observed, allow us to investigate in real time how the participant's

cognitive system responds to the manipulations of the stimulus. An auditory or visual stimulus can produce a potential variation in the midbrain, which could be measured on the scalp with suitable techniques (EEG).

4.2 Language-related Event-related potentials (ERPs) components

The decomposition of ERP waveforms in terms of their underlying component is trying to explain the different neural and cognitive processes in the linguistic recognition procedure. In psycholinguistic experiments, some ERP components are used to investigate the neurodynamic of different cognitive processes, unfolding from the auditory level to the pragmatic level. Most of the ERP components are referred to by a letter (N/P) indicating polarity (negative/positive), followed by a number indicating either the latency in milliseconds or the component's ordinal position in the waveform (Nunez & Srinivasan, 2006). Accordingly, ERP is characterized by three dimensions: the polarity (negative/positive ongoing voltage measured in microvolt), the latency measured in milliseconds, and the topography of the surface indicated by the localization of the electrodes at the surface of the scalp. Moreover, the ERP components are measured by assessing their amplitude and latency. Amplitude (μV) is measured by the mean pre-stimulus baseline voltage and the largest polarity peak of ERP waveform within a time window determined by the experimental condition. Latency (ms) is defined as the time from the stimulus onset to the point of the maximum polarity amplitude within the same time window (Luck, 2005). The different ERP components emphasize different aspects of the defining features, and some components, which belong to the same family might present the similar polarities and time courses to the experimental manipulations (e.g., P3a & P3b in the family of P3). In this section, some language-related ERPs components, e.g., auditory evoked potentials (N1, P2), MMNs, LAN, ELAN, P200, P300 (P3a, P3b), N400, and P600 in the investigation of language processing are illustrated.

4.2.1 Auditory sensory potentials (N1 & P2)

Development in the auditory system can be evaluated by auditory event-related potential (AERP) components. N1 is a negative polarity, and it peaks at 100ms latency from the stimuli onset, which is generated in the subcortical areas (Kaseda et al., 1994). As the early auditory sensory potentials, N1 can be quantified as the cortical activities in response to the auditory input (Nordby et al., 1994). P2 is a larger positive polarity whose peak latency is at 200ms post-stimuli. N2 is a larger negative polarity whose peak latency is at a frontocentral region in the 171ms - 271ms time window. As late auditory sensory potential, the response of P2 reflects the attention orientation,

decision-making, and activation of the immediate memory (Polich & Kok, 1995). Within the time window of 100ms - 200ms after the auditory stimuli onset, the N1/P2 biphasic ERP complex can be maximal at the scalp vertex, which can be in response to the processing of the physical attributes to the auditory inputs (Näätänen & Winkler, 1999). N1 and P2 could be modulated by the shift of the subject's attention and age. When the processing attention is reduced, the amplitude will be smaller (Kutas & Hillyard, 1980). And their amplitude will decrease with the subject's age (Mikkola et al., 2007).

4.2.2 MMNs (Mismatch negativities)

Auditory mismatch negativities (MMNs) are the negative fronto-central component, which peak latency is at 150- 250ms after the deviant stimulus onset. The MMNs are generally observed in an oddball paradigm (Dehaene-Lambertz, 1997), which reflect the neural mismatch between the deviant and standard stimulus, which are obtained in response to the infrequent deviant stimuli, that is interspersed in the unattended repetitive sequences of the standard stimuli (Näätänen et al., 2007; Näätänen et al., 2004). Their scalp distribution is frontocentral, and its peak can start at the N100 wave and overlap with the P200, which lasts about 100ms to 250 ms after the stimuli onset. MMNs were reported by several research methods, including MEG, fMRI, and PET (Näätänen, 1991). In the auditory modality, MMNs have often been obtained by presenting the participants with a series of identical stimuli (standard) in which the acoustically deviant stimulus is occasionally inserted. In a typical study by Sams et al. (1985), when participants were reading a book, a sequence of standard stimuli tones (e.g., 1000 Hz) were presented at the same time, which were occasionally inserted with the deviant target tones in the different frequency levels (e.g., 1032 Hz), MMNs were elicited by the participant's brain automatically in response to the deviant target tones (Fig.4.1)

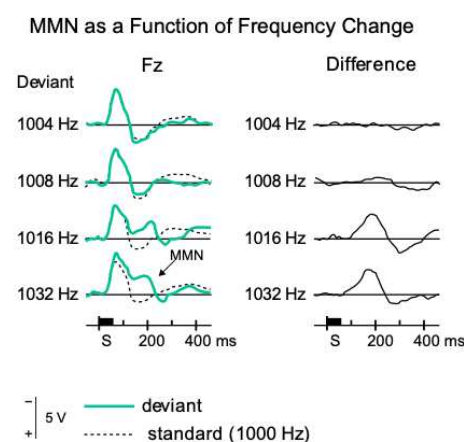


Fig.4.1. Figure by the research of Sams et al. (1985), in which participants are reading a book and listening to the tones in different F0. On the left, there are averaged ERPs from the frontal sites, standard stimuli are in 1000 Hz

(solid line), and deviant targets (broken line) are in different levels of frequencies in the middle. On the right, there are waveforms by subtracting the standard from the deviant for the different deviant frequency levels.

MMNs are evoked in the involuntary condition, which reflects the automatic brain process, which supports the hypothesis that its elicitation can be independent of the participant's attention. MMNs could be generated in response to the other phonetic changes, e.g., the intensity (Toivainen et al., 1998), duration (Näätänen et al., 2004), locus of origine (Paavilainen et al., 1989), and timbre (Toivainen et al., 1998). MMNs were revealed to be sensitive to the probability of deviant targets, the amplitude would be decreased by increasing the number of deviant targets (Sabri & Campbell, 2001; Sato et al., 2000).

4.2.3 LAN & ELAN

Left anterior negativity (LAN) is an earlier negative shift in the ERP waveform, which was elicited by syntactic anomalies, this negative shift appears over the anterior electrodes, and in some cases, it was found to be lateralized to the left side of the head. In previous research, LAN effects were reported at varying latencies following a stimulus, the early LAN (ELAN) effects were found as early as 100–300 ms from the stimulus onset with an anterior distribution, and these early effects were distinguished from the later LAN effects in the same latency window as the N400 (300–500 ms) (Munte et al., 1993; Neville et al., 1991). ELAN and LAN effects were found in word category violations. When the parser anticipates that the upcoming word will be of a specific grammatical category (e.g., noun, verb, etc.), but the upcoming word violates that expectation. E.g., “The young apprentice went to see the new *designing* in the museum,” in which a verb (*designing*) is in a noun position (*design*) (Friederici et al., 1996; Friederici et al., 2002; Hagoort et al., 2003).

4.2.4 P300

P300 was firstly reported by Sutton and his colleagues in 1965, it is a positive component in the 300-600ms time window, the peak latency of P300 is at 300ms - 350ms from the stimuli onset, the scalp distribution is over the midline electrodes sites (Fz, Cz, Pz), which may increase from the frontal to the parietal sites. P300 is also called P3 or late positive component (LPC). The increased amplitude of P3 is interpreted as the additional cost of resources in the cognitive system by neuronal activities, such as unexpected prosodic deviance in the incoming stimulus (Fig.4.2). In addition, both P3a and P3b are the subcomponents in the P300 family. P3a is related to the neural changes when the incoming stimuli replace the contents in the working memory (Desimone et al., 1995), and P3b is related to the memory storage operations (Knight, 1996; Squire & Kandel, 1999).

When a distracting stimulus modulates the frontal lobe attention, the P3a is produced; when attentional resources are allocated for the subsequent memory updating after stimulus evaluation, a P3b is produced to establish the connection with the storage areas in the associational cortex.

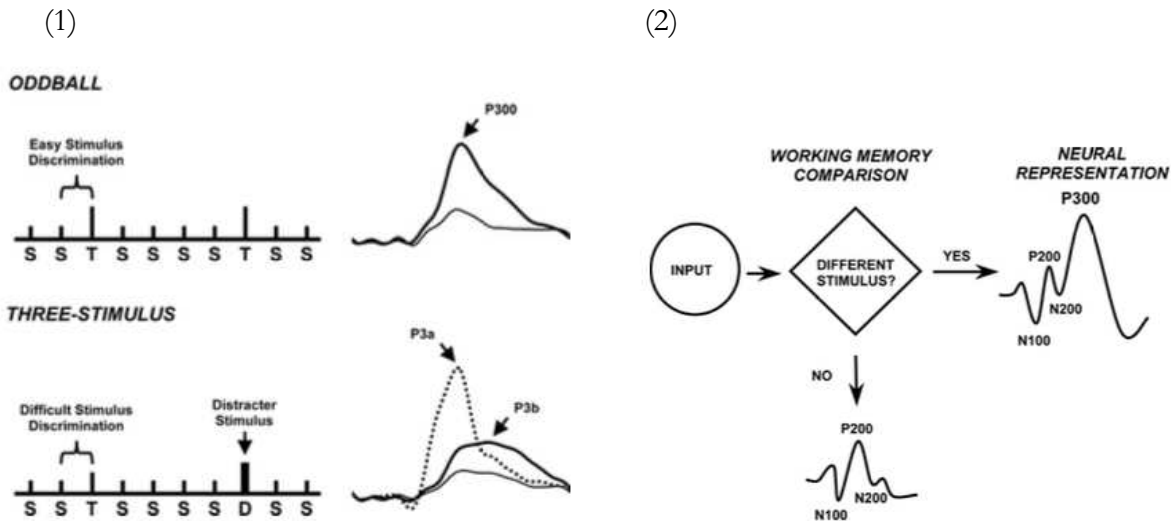


Fig.4.2. (1). P300 and its variants in typical oddball tasks. In the traditional oddball task, the infrequent target will elicit a P300 in the background of frequent standard stimuli. In the three-stimulus oddball task, there are frequent standard stimuli (S), an infrequent deviant target (T), and a distractor (D). P3a is elicited by the distracter, and P3b is elicited by the infrequent target (Polich, 2007). (2). Illustration of brain activities elicited by the P300 effect in an oddball task. Standard stimuli evoke the sensory potentials (N100, N200, P200), and infrequent targets elicit P300.

In the auditory modality, the P300 has been observed in the traditional oddball paradigm, in which infrequent targets are presented in the background of frequent standard stimuli (Polich, 2003). When an incoming target is the same as the standard stimulus in the background, only the sensory evoked potentials are elicited (N100, N200, P200). After the incoming stimuli input, the participant's cognitive system makes the comparison between the input and the frequent standard stimuli in the previous working memory. If the incoming stimuli are not the same as the standard stimulus (around 10% of the cases), then participants allocate more attention to process infrequent targets, P300 can be elicited by deviant targets (Polich, 2007; Rushby et al., 2005). The amplitude and peak latency of P300 is shown to be sensitive to the number of consecutive standard stimuli (Sommer et al., 1990), the target-to-target interval (Gonsalvez & Polich, 2002), and the category number of the target stimulus (Mertens & Polich, 1997).

4.2.5 N400

N400 is one of the most important language-related ERP components, it is a negative polarity shift in the waveform over the central-parietal region of the scalp (Kutas & Hillyard, 1983). Peak latency of N400 is usually between 380ms to 440ms after the stimulus onset. The increased

amplitude of the N400 component is commonly observed in response to a semantic incongruent sentence like “*He spread the warm bread with socks.*” used in the seminal study of Kutas and Hillyard (1980) compared with a semantically and syntactically acceptable sentence (e.g., *It was his first day at work.*) (Fig.4.3). The N400 effect has been replicated in different languages, sensorial modality, and populations (Kutas & Federmeier, 2011). Isel et al. (2007) have replicated the N400 in French sentences containing a subject-modifying relative clause (SRC). This is viewed as an indication of greater cognitive costs during lexical-semantic integration. In a typical experiment, the N400 effect is elicited by the presentation of the semantically manipulated critical stimulus to the participants, and it was found to be modality independent, which is observed in both the spoken and written language processing (Isel et al., 2007; Kutas & Federmeier, 2009).

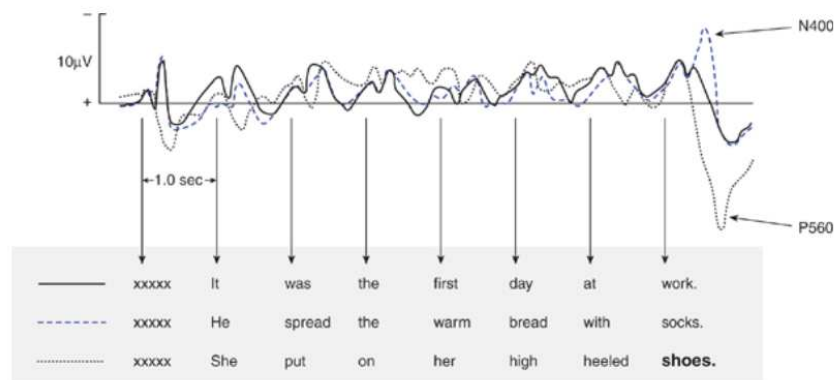


Fig.4.3. N400 in the visual modality by Kutas and Hillyard (1980). Sentences in three conditions are presented to the subject, and the final word in each sentence is the target stimuli. The solid line represents the control condition (correct), the dashed line represents the semantic anomalous condition, and the dotted line represents the semantically congruent but physically incongruent condition. Horizontal axis shows the time course in milliseconds, and the vertical axis shows the microvolts.

E.g., in Fig.4.3, in a research on sentence comprehension in English in the visual modality, the sentences in three manipulated are presented to participants one word at one fixation on the screen. ERPs are compared for the final word in the three conditions. In the sentence ‘He spread the warm bread with socks’ the so-called “semantic incongruent condition”, final stimuli ‘socks’ which is inappropriate with the sentence context has elicited the largest N400, which peaks at 400ms after the stimuli onset and maximal over centro-parietal electrode site. N400 component is thought to be sensitive to the manipulations of semantic properties and the other properties in the experiments, e.g., the lexical frequency (Barber et al., 2004), real/pseudoword (Leinonen et al., 2009), orthographical neighborhood (Holcomb et al., 2002), etc. And it is not only to be observed in semantic incongruities, but it was found to be sensitive to critical stimuli that are semantically correct but less expected in the sentence context (Grainger & Holcomb, 2009).

4.2.6 P600

The P600 effect is a late positive polarity in the ERP waveform over the posterior electrode site (Osterhout & Holcomb, 1992), the peak latency is usually observed at around 600ms after the stimuli onset, its onset is at around 500ms after the stimuli which might last for several hundred milliseconds (Friederici et al., 2002). It has been observed both in written and auditory experiments (Hagoort & Brown, 2000). The P600 has traditionally been associated with a syntactic violation (Hagoort et al., 1993). Generally, the P600 would be elicited when participants read sentences that contain a syntactic violation of word categories, such as in the seminal study of Kilborn and Friederici (1994) in German, when the noun in the prepositional sentence was deleted, a larger P600 effect was evoked in relation to the syntactically correct sentence. Note that some researchers found that the late positivity was preceded by an early left anterior negativity (ELAN) thought to reflect the detection by the language processing system of an unexpected syntactic word category. Typically, the Centro-posterior P600 effect was interpreted to reflect a process of syntactic repair (Osterhout & Holcomb, 1992). In some research, the P600-like component could also reflect the syntactic reanalysis process in the late integration stage (Kaan & Swaab, 2003). The P600 effect has been observed in response to both the written and auditory stimuli (Hagoort & Brown, 2000).

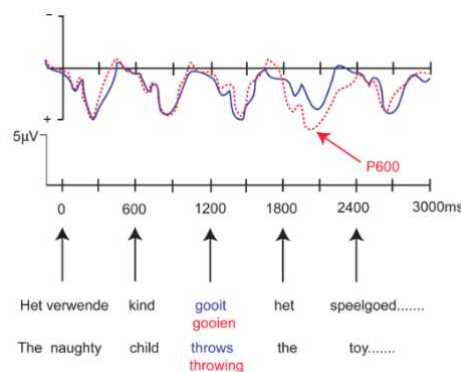


Fig.4.4. Elicitation of P600 effect in the research of Brown and Carr (1993). Subjects read individual sentences containing one of three different kinds of violations of the syntactic constraints of Dutch. P600 is elicited by the syntactic violation in the red line, and the syntactic correct condition is represented in the blue line.

In Fig.4.4, in contrast to the syntactic correct sentences, the stimuli word (throwing) in the sentences from the syntactic violated condition elicited a larger P600 effect. P600 was also shown to be sensitive to different kinds of syntactic anomalies in the experimental manipulation, e.g., the phrase structure (Friederici et al., 1996), gender violation (Coulson et al., 1998), and subadjacency violation (Neville et al., 1991). P600 has also been observed in using the violation of mathematical rules, abstract sequence, and music (Luck, 2005).

4.3 ERP evidence of rhythmic pattern processing in word recognition

4.3.1 Previous studies in Romance and Germanic languages

Previous studies concerning event-related potentials (ERPs) revealed that Romance and Germanic language listeners use stress information contained in the speech signal to realize word recognition. Several aspects of prosody pattern in online spoken-word processing were analyzed:

- a) the pitch and fundamental frequency (Friedrich et al., 2001; Friedrich, Kotz, Friederici, & Alter, 2004)
- b) the syllable duration (Magne et al., 2007)
- c) the sub-phonemic and supra-phonemic features (Ashby, 2010; Ashby & Martin, 2008; Ashby et al., 2009).
- d) the metric pattern (Böcker et al., 1999; Eckstein & Friederici, 2005; McCauley et al., 2013; Schmidt-Kassow & Kotz, 2009; Steinhauer et al., 1999)
- e) the predictability of the prosody pattern (Domahs et al., 2013; Henrich et al., 2014; Magne et al., 2010; Steinhauer et al., 1999).
- f) the detection of auditory changes by Näätänen and other colleagues (Näätänen et al., 2007).

The linguistic function of prosody in lexical processing was analyzed in several acoustic parameters at the word level, such as the F0 and duration. Friedrich et al. (2001) showed that P200 is sensitive to pitch contour processing, the initially unstressed words elicited a larger P200 effect than the initially stressed words. Based on the finding in 2001, Friedrich, Kotz, Friederici and Alter (2004) suggested that the pitch contour of words modulates the amplitudes of P350, which is an ERP component that links to the prosodic processing in word identification. By using the French trisyllabic spoken words, Magne et al. (2007) manipulated the syllable duration to examine the influence of rhythmic properties on semantic processing. French is described as a language with fixed accents at the end of rhythmic units, which are characterized by lengthening the last syllable of words (Delattre, 1966). Magne et al. collaborators, used the cross-splicing procedure to apply the duration of the last syllable (lengthening) to the penultimate syllable (non-accented and normal duration) within the trisyllabic words. This experimental manipulation aimed to create a ‘metrical violated condition’, in which the target words have misplaced stress accent and unnatural syllable duration. For example, the word *canapé* (sofa) is normally accented on the last syllable, while a metrically violated target word **canapé* is accented with incongruous lengthening penultimate syllable *na*. Semantic congruency of target words was also manipulated to create a semantic violated condition at the sentence level. Sentences with the targets placed at the last position are aurally

presented to participants. Results demonstrated an N400-like effect on the targets with misplaced stress accents, and the semantically incongruous words elicited a larger N400 than the semantically congruous words in metrical judgment tasks. Researchers interpreted this enhanced N400 effect as a reflection of disrupted access in response to the semantic integration caused by the abnormal metrical pattern. Grainger et al. (2006) have found that briefly presented homophone primes (e.g., brane–BRAIN) have yielded less-negative N250 amplitudes on the target words than the control primes (e.g., brant–BRAIN), this result has been attributed to the more effective application of the target phonology in lexical access with the homophone primes.

Prelexical processing of sub-phonemic and supra-phonemic features was demonstrated by Ashby and his colleagues. Ashby et al. (2009) analyzed the congruency of the consonant pattern by applying the masked priming paradigm. Pseudoword primes with a voiced or an unvoiced final consonant (e.g., fap, or faz) were followed by the words with either a congruent or an incongruent final consonant articulation (fap–fat [congruent] vs. fap–fad [incongruent]). Results showed that the early ERP components were modulated by the congruency of consonants contained in the prime–target, the congruent condition has yielded less-negative amplitudes within 250– 350 ms post-target. Ashby and Martin (2008) analyzed the processing of the sub-phonemic information in the visual word recognition in English, and smaller negative amplitudes were elicited within the 250-350ms post-target when the first syllables of the prime and target were matched (pi#### of pilot) than when they were mismatched (pil##### of pilot). And Ashby (2010) observed an even earlier effect of syllable congruency, the amplitude of N100 (in the 100-120ms time window) was reduced when the targets were preceded by syllable-congruent primes.

Event-related potentials have been applied to explore the deviant prosody pattern in lexical processing. The metrical function of prosody is firstly examined by Böcker et al. (1999) through the rhythmic properties (metrical stress) of spoken words in Dutch. Participants either listened passively or discriminately to sequences of four disyllabic words, which ends with a word that had either the same or different stress pattern (strong-initial / weak-initial) as the rest of the list. Results showed that the weak-initial words, which were much less common in the Dutch lexicon (about 12%), could elicit several larger negative components than the strong initial words. An enhanced left-frontal distributed negativity N325 at around 325ms has been reported when the weak-initial words are followed by a sequence of strong-initial words. The N325 was also larger for participants that showed good performance in both the metrical discrimination task and the passive listening task. Authors suggested that N325 is a correlate of metrical stress perception, which manifested the extraction of metrical stress from the acoustic signals.

Steinhauer et al. (1999) observed a P600 in these well-formed sentences with incongruous prosodic phrasing, the authors interpreted this effect as a reflection of the cost of both syntactic and prosodic revisions. Eckstein and Friederici (2005) have manipulated the critical noun in two different positions in the German phrase (both penultimate and final positions), and the prosodic congruency of the targets is controlled to examine whether the two levels (prosodic and syntactic) are interactive or independent from each other in the language processing. Researchers observed the N400 in syntactically correct phrases, the prosodic incongruity enhanced this effect, which provides evidence for the reflection of lexical integration cost to prosodically unexpected targets. P600 is observed in syntactically correct sentences that contained a prosodic mismatch at the final word, which is interpreted in this study as a reflection of intonation contour violation. And there is an over-additivity to the P600 effect in the combined prosodic-syntactic violation condition in relation to the pure syntactic/prosodic condition.

Schmidt-Kassow and Kotz (2009) explored the perceptual regularity of disyllabic trochee pattern words in German. In trochaic sentences, the words are disyllabic with the first syllable stressed. Target words in the sentences were manipulated with metrical and/or syntactic violations, both the metrical homogeneity task and grammaticality judgment task are applied to participants. E.g., ‘Gina hatte Norbert gestern abend *kussen*/**be’lohn*en sollen’ (Gina should have *kissed*/*rewarded* Norbert yesterday evening), the target word was manipulated with the stress pattern deviations. Results showed that the brain is sensitive to stress violation when participants focus on the metrical structure in the metric task, the subtle metric deviation of the targets elicits a P600, which revealed that metric and syntactic cues interact in the late integrational stage indexed by the P600. However, this effect has not been observed in the grammatic task, when participants focus their attention in the syntactic structure.

McCauley et al. (2013) observed an N400-like effect in response to the rhythmic violation during the lexical-semantic integration during silent reading. A picture/word matching task was used to examine the role of compound/phrasal stress distinction in online language processing. E.g., the picture (content: a green-colored house) is matched with the compound/phrasal pairs *greenhouse* (incongruent) and *green house* (congruent). A centro-parietal negativity N400 was found in response to the violated stress pattern, which resulted from the incongruity between the content of the image and the semantic representation of the word. Unexpected phrasal stress can elicit a P600 effect in the semantic incongruent condition, researcher interpreted the P600 as a reflection of the syntactic integration difficulty resulting from the unexpected stress pattern.

The influence of rhythmic expectancy was investigated in some electroencephalography (EEG) studies. Magne et al. (2010) examined the metrically unexpected English words (e.g., the

last target with contrastive stress pattern compared to the other words in the session) in visually presented experiments. Results indicated that the regular stress pattern can elicit expectancy about the stress pattern of incoming words. An N400 effect was observed in response to the processing of the words with an unexpected stress pattern, which was interpreted to be elicited by unexpected stress patterns in semantic processing. This latter result is consistent with the previous study to support the idea that incorrect rhythmic patterns may hinder semantic processing. Domahs et al. (2013) have conducted an event-related potential (ERP) study about the predictability of prosody patterns in Turkish. For this purpose, the stress pattern of trisyllabic words was manipulated either with the standard final stress pattern (e.g., mi'sır, "*corn*") or with an exceptional nonfinal stress pattern (e.g., 'mısır, "*Egypt*"). Results revealed that the asymmetrical ERP responses were found for different stress violations: the final stress produced an N400 effect whereas the stress violations with nonfinal stress produced a P300 effect. Henrich et al. (2014) have investigated how rhythmic predictability and violation influence lexical processing. The alternation of stressed and unstressed syllables was manipulated for the minimal pairs of English noun-noun compounds, for example, champagne cocktail (stress on the final position) and chàmpagne cocktail (stress shifted on the initial position). Results showed that rhythmic regularity was relevant for language processing: the deviations contained in the rhythmic pattern evoked different ERP components, and an N400 was found for the items with a deviant stress pattern, which reflects a higher cost in lexical processing due to the prosody deviation.

Some studies investigated automatic auditory discrimination, and Mismatch negativity (MMNs) is thought to reflect the processing of auditory changes. This kind of experiment often used both two categories of stimulus (deviant and standard), and MMNs were viewed as an index to estimate the detection of auditory changes in the incoming input. The response and its magnetic counterpart, i.e., the MMNs was stronger when the acoustic difference between the deviants and standards crosses a language-specific category boundary than when it does not. At the word level, the effect of MMNs was observed for the distinctive changes in the vowels (Näätänen & Alho, 1997; Silva et al., 2017; Winkler et al., 1999), in the consonants (Dehaene-Lambertz, 1997; Sharma & Dorman, 1999, 2000; Zhang et al., 2005), in the stress pattern (Rodrigues Silva & Rothe-Neves, 2020), in the pitch contour and emotional prosody (Leitman et al., 2009), and in the decrement of duration and intensity of the stimuli (Kaukoranta et al., 1989; Näätänen et al., 1989; Paavilainen et al., 1991). Moreover, at the sentence level, MMNs were found to be elicited by the omission of an element of a compound word stimulus (Nordby et al., 1994; Winkler et al., 1993), and the omission of stimulus in a stimulus sequence (Rüsseler et al., 2001; Yabe et al., 1997).

4.3.2 Previous studies in Chinese Mandarin

The rhythmic pattern has been shown to constrain the semantic integration in Mandarin word processing, a set of EEG studies investigated the neural mechanism of prosody processing in several linguistic parameters, including:

- a). the processing of lexical tones (Luo et al., 2006; Ren et al., 2009; Tsang et al., 2011; Zhang et al., 2012)
- b). the neutral tone in the word recognition (Isel & Shen, 2020; Luo et al., 2016; Yan et al., 2014)
- c). unexpected or violated rhythmic patterns (Wang et al., 2005)
- d). the number of syllables of compound words (Luo & Zhou, 2010)
- e). the congruency of lexical tones and accentuation (Li et al., 2008)

The primary experimental paradigm adopted by the researchers is the passive auditory oddball paradigm. In this kind of experiments, participants are watching a book or a silent movie while simultaneously they are presented auditorily with two categories of stimuli, both the frequent standard stimulus and infrequent deviant stimulus. Event-related potentials (ERPs) are recorded during the experiments to investigate the neural responses to standard and deviant stimuli. Tsang et al. (2011) investigated the pre-attentive processing of lexical tones in Cantonese with a passive oddball paradigm. Standard and deviant stimulus in the experiments differed in their pitch levels and pitch contours, and ERP were recorded when native listeners were watching a silent movie. Results revealed that tones were processed pre-attentively, MMN and P3a were elicited by all types of deviant tones, and the modification in pitch level and pitch contour were respectively in relation to the amplitude and latency of MMN and P3a. Luo et al. (2006) also investigated the pre-attentive processing of Mandarin lexical tones with an auditory oddball paradigm. Native participants were presented with the stimulus in the consonant-vowel structure, which was either manipulated in the lexical tone patterns or the initial consonants. Results indicated that a larger MMN was elicited by the deviant stimulus at the pre-attentive stage. Similar research has been conducted by Ren et al. (2009), which analyzed the pre-attentive processing of Mandarin lexical tones and intonation. Ren et al. (2009) reported a larger amplitude of MMNs associated with deviant lexical tones compared to deviant intonation. Additionally, the N2b and P3 are known to reflect the cognitive procedure which takes place in the attentive processing stage during the auditory stimulus perception in an explicit oddball paradigm. N2b is increased negativity that follows MMNs in time and peaks at about 200–300ms from stimulus onset. Zhang et al. (2012) showed that N2b and P3 are sensitive to the detection of deviant Mandarin tones in the auditory modality, and the larger amplitudes of

these two components were elicited by the across-category stimulus (e.g., two tones from the same category of Tone 2) than within-category stimulus (e.g., two tones respectively from the categories of Tone 2 and Tone 4).

Some studies investigated the processing of neutral tone in word recognition. Isel et al. (2020) examined the ERP in the processing of neutral tone in two-character Chinese compounds. Sixty pairs of full tone and neutral tone compounds were selected for both the explicit stress-discrimination task and the implicit passive listening task, and the modified oddball paradigm was applied to 18 native listeners. Results revealed the perception of neutral tones by native Mandarin listeners. Consistent with the research of metric pattern of Dutch compounds by Böcker et al. (1999), the larger amplitudes of N200, P200 and N325 were observed for the infrequent neutral tone compounds (Strong-weak) in contrast to the frequent full tone compounds (Strong-strong). Moreover, Luo et al. (2016) analyzed the application of phonological knowledge during the reading of Chinese scripts, especially the processing of neutral tones during silent reading. Both full tone and neutral tone two-character compounds were selected, their second characters possess the same syllable structure with either a full tone or a neutral tone. Native readers read the sentences with the two types of target compounds embedded within them, and the eye-tracking techniques are used to record the eye movements of participants in the reading procedure. Results indicated that the shorter viewing durations, fewer refixations, and fewer regressive saccades on the target words when it contains a neutral tone. Similar result was observed by Yan et al. (2014), the shorter viewing durations are observed for neutral tone compounds than for full tone compounds, these results are only found in the northern-dialects speakers and not in the southern dialect speakers of Chinese Mandarin. The latter result suggests that the neutral tone's perception is dialect-specific. However, both two researchers didn't attribute the results to the easier processing of neutral tone compound. In the same study, Luo et al. (2016) found a more-negative amplitude of N400 during the readings of neutral tone compounds in relation to the readings of full tone compounds. Yan et al. (2014) found that readers processed an upcoming parafoveal word less effectively when a neutral tone compound was fixated than in the fixation of a full tone compound. Hence, the consumption of cognitive resources in the processing of neutral tones is still not clear according to our knowledge.

Mandarin readers seem to use sub-phonemic or supra-phonemic word features in the lexical access. Wang et al. (2005) used an auditory oddball paradigm to study whether the prosodic salience (e.g., stressed syllables in disyllabic speech sound) can serve as the cues to capture listener's attention in speech processing. Disyllabic standard and deviant input, which differed in prosodic and temporal status, are presented to participants. Two types of stimuli with prosodic salience in different positions (***BA**ga and ba**GA***) are chosen as standards. According to the temporal position

and prosodic status, voiced consonants were replaced by unvoiced consonants (e.g., /b/ → /p/). The four types of deviants are manipulated and produced: (1) stressed initial: ***PAga***; (2) unstressed-noninitial: ***BAka***; (3) stressed-noninitial: ***baKA***; (4) unstressed-initial: ***paGA***. Participants were tested while watching a silent movie. The behavioral EEG results showed that prosodic salient deviance capture attention in unattended sounds, and MMNs were elicited by deviants in all four conditions in the experiments, suggesting discrimination between the standard and deviant input. P3a was observed when the deviance was placed on a stressed syllable (e.g., ***PAga*** and ***baKA***). The listener's sensitivity of P3a to the syllabic prosodic manipulation can be viewed as strong evidence for the hypothesis that prosodic salience induces the attentional resources during the involuntary speech sound perception. Stressed syllables work as attentional cues to facilitate speech perception. The results suggest that the cognitive system of listeners might give priority to prosodically salient information irrespective of its temporal position in the speech input. This finding proposes that prosodic salience can capture listeners' attention, and segmentation is achieved through increased attention to stressed syllables rather than word onsets.

Luo and Zhou (2010) investigated to what extent the prosodic constraint in Mandarin plays a role during silent reading. The number of syllables and semantic congruency of verbs within the compound combinations were manipulated in the Mandarin lexical combination, e.g., the word in the verb-noun combination 种蒜 (plant + garlic) is a normal [1+1] rhythmic pattern, compared to 种植蒜 (plant + garlic), which is an abnormal [2+1] rhythmic pattern. Critical words were embedded in the same position of phase, and both EEG signals and behavioral results of listeners were recorded. Researchers found that the abnormal [2+1] rhythmic pattern can evoke a posterior positivity P300-like effect in the 300- to 600ms window and a frontocentral N400-like effect in semantically congruent sentences, this P300-like effect was interpreted to reflect the detection of rhythmic violation between [1+1] and [2+1]. The observation of an N400-like effect suggests that a mismatch between the incoming prosodic information and stored rhythmic pattern could affect the recognition of the current word. Results offered us evidence that abnormal rhythmic patterns may evoke an N400-like effect, and the additional semantic incongruency may enhance this effect. Li et al. (2008) conducted event-related brain potentials (ERP) study to investigate the congruency of pitch accent and lexical tone in spoken discourse comprehension in Chinese Mandarin. Results showed that both inconsistent pitch accent and inconsistent lexical tone yielded an N400 effect, and the negativities evoked by the inconsistent pitch accent had presented the same topographical distribution as the negativities evoked by inconsistent lexical tone violation.

Chapter 5

Acoustic-phonetic correlates and tonal coarticulation of the Mandarin two-character neutral tone compounds

5 Acoustic-phonetic correlates and tonal coarticulation of the Mandarin two-character neutral tone compounds

5.1 Experiment design

In Mandarin, there are four lexical tones, Tone 1/2/3/4, they are the four canonical tone patterns in Chinese, which are also called ‘full tone’, in contrast to ‘neutral tone’ (Tone 0). Because one constituent (one character in writing form) could carry more than one tone, it’s possible that a constituent possesses both a full tone form and a neutral tone form, which might be semantically different. Within the Mandarin two-character compounds, the syllable of the second (rightmost) constituent could be tone-neutralized, which is in combination with the first full tone syllable and constitutes a neutral tone compound in a trochaic prosody pattern [Strong-weak]. In contrast, a full tone compound that has two full tone syllables constitutes a spondee prosody pattern [Strong-strong]. In Mandarin, it exists a basic distinction between the two prosody patterns, the former one with a full tone syllable carried by the second constituent, and the latter one with a neutral tone syllable carried by the second constituent. The phonological realization of the second neutral tone is quite different in the acoustic cues in relation to the full tones. Gandour (1983) claimed that pitch contour and pitch height are the two dimensions of F0 for tonal perception in Mandarin.

Based on previous studies, the main acoustic correlated of neutral tone are duration, pitch (F0), and intensity (Lin & Yan, 1990; Liu & Samuel, 2004; Yip, 2002). Moreover, because there are four lexical full tones (Tone 1/2/3/4), the phonological realization of a neutral tone syllable within compounds varies with the neighboring syllables, and the contextual tonal variation exists due to the coarticulation effect. Combined with different preceding and following tones, the acoustic cues and pitch value of the second neutral tones could be variant in the four tonal contexts. With respect to the role of prosody structure in continuous speech recognition, according to Grosjean and Gee (1987), the alternation of strong and weak syllables within words plays a role in the auditory lexical access in some Germanic languages, such as in English. Considering the trochaic prosody pattern of two-character neutral tone compounds, it is crucial to understand the phonetic features of the second neutral tone, however, fewer study has been constructed to explore the acoustic correlates of neutral tone within multi-morphemic words. The current analysis of acoustic features of neutral tone compounds is conducted to fill this gap in the literature.

In the present dissertation, the main goal of Chapter 6 is to describe the prosodic features of the Mandarin two-character neutral tone compounds. Accordingly, two acoustic analyses were run. In sub-chapter 6.2, we investigate the acoustic features of the second neutral tones in contrast with their full tone counterparts within disyllabic neutral tone compounds. In sub-chapter 6.3, we

provide the detailed tonal variation information of neutral tone compounds in the four full tone contexts, both the carry-over effects and the anticipatory effects contained in the two neighboring tones will be analyzed. For this purpose, a multi-speaker acoustic production study was designed to realize the analysis with native speakers of Mandarin. Firstly, we recorded the production in a sound-proof chamber at the MoDyCo Laboratory. And we examined the recordings with Praat and finally the results were analyzed in RStudio.

Methodology

Participants Six speakers (3 males and 3 females) were recruited for the recordings, they are college students at Paris Universities, their ages range from 26 to 32 (mean 28.28, ± 2.21), all of them are native Mandarin speakers who were born and grown up in Beijing of China, Mandarin is their mother language, none of them had lived in France more than three years (ranging from 0.5 to 3 years, mean 1.92, ± 0.23). They were paid 10€ / hour for participation.

Linguistic materials and recording procedure As for the two sub-chapters, in total, 360 two-character compound words, and 20 monomorphemic characters in Chinese Mandarin were selected as experimental materials, including 160 compounds for the analysis in sub-chapter 6.2 and 200 compounds and 20 characters the analysis in sub-chapter 6.3. During the recording, target words were embedded in the carrier phrase and recorded by the speakers, six recordings were produced respectively by the participants. Before the production task, all the speakers received instructions. The recordings were carried out in the MoDyCo laboratory (CNRS) at Paris Nanterre University. They were recorded individually in a sound-proof phonetic chamber, using an attached microphone, which was placed at a distance of approximately 5 centimeters from speaker's mouth. Speech samples were recorded digitally at 44,100 Hz, 16-bit mono.

Example: 我说了招呼这个词两遍。

[I said the word salutation twice.] (salutation is the target compound)

Acoustic measurements and analysis Selected words were embedded into the same last position of the carrier phase and produced by the six speakers. In each part of the analysis, only the acoustic correlates of target syllables were extracted and analyzed in the results. For the six recordings, the target syllables were firstly annotated manually by Praat (Boersma, 2006). Secondly, three Praat scripts were applied to extract the acoustic variation of target syllables, including their syllable duration, pitch (F0), intensity, and transition pitch (F0) between the two syllables within the disyllabic compounds. The acoustic measurements extracted were subject to be examined by RStudio for the results. Finally, the averaged F0 curve of the target compounds was represented through graphical comparisons in point-to-point tracing for the selected compounds. As for the

Praat annotation (Boersma, 2006), the first syllable, second syllable, and the pause between them for each critical compound were manually marked, which yields four markers, the first marker and the second marker at the beginning and the end of the first constituent, the third marker and the fourth marker at the beginning and the end of the second constituent. As for the pitch (F0), each syllable was divided into ten segments normalized in time, and another Praat script was applied to extract the averaged F0 (Hz) from them. We have analyzed the pitch range (Hz), pitch level (Hz), and pitch transition (Hz) for each target syllable. In our study, pitch range (Hz) was calculated by subtracting the minimal pitch value from the maximal pitch value, pitch level (Hz) was calculated by averaging the maximal and minimal pitch value, and pitch transition (Hz) was analyzed by measuring the F0 differences between the offset of the first syllable and the onset of the second syllable, the measurements were subject for the Repeated ANOVAs and paired t-tests in RStudio.

5.2 Acoustic correlates of lexical neutral tone within Mandarin two-character compounds in the four full tone contexts

Within the two-character neutral tone compounds, the second constituent carries a neutral tone syllable, and the first constituent carries a full tone syllable, i.e., ‘*tou*(T2) *fa*(T0), 头发 hair’. Neutral tone is weak and toneless, in relation to the full tones syllables which are viewed as strong and stressed (Duanmu, 2007). In the auditory modality, the ideogram and context are necessary to differentiate the homophonic morphemes in the auditory word recognition (Zhou & Marslen-Wilson, 1994). Such as the pair of compounds 嘴巴/*zui3ba1*/ and 下巴/*xia4ba0*/, their second constituents are 巴/*ba1*/ and 巴/*ba0*/, which have the same character in writing form and possess an identical syllable structure, they carry either a Tone 1 or a Tone 0 within the two compounds, /*ba1*/ is the full tone counterpart (in full tone 1) of /*ba0*/. As for /*xia4ba0*/, the first constituent carries a Tone 4, the second constituent carries a Tone 0. Through the combination of a leftmost full tone and a rightmost neutral tone, a prosody pattern [Strong-weak] is formed, compared with the full tone compound /*xia4ba0*/ [Strong-strong]. Fig.5.2.1 presents a pair of compounds 下场/*xia4chang3*/ (translation in English) and 下场/*xia4chang0*/ (translation in English) pronounced in the citation form, their second syllables are either in Tone 3 or Tone 0, the contrasting pitch between the strong full tone and weak neutral tone is signaled mostly by duration and F0. Within the neutral tone compound, /*chang0*/ has a shorter duration and mid-fall in F0 contour, compared with a longer duration of its full tone counterpart /*chang3*/. During auditory recognition, Mandarin listeners are sensitive to contrastive suprasegmental information contained in the syllables (Isel & Shen, 2020). When listeners hear the phonological form ‘/xiachang/’, their language processing

system activates both a neutral tone form ‘/xia4chang0/’ and a full tone form ‘/xia4chang3/’, the correct representation is activated by the neutral tone /chang0/.

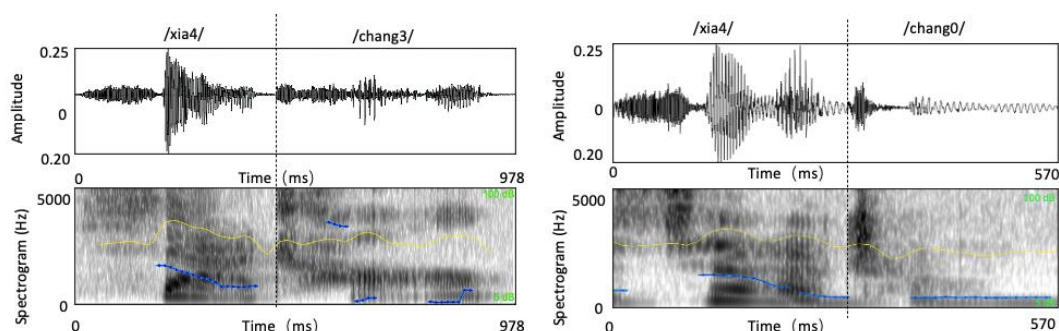


Fig.5.2.1. Left: Full tone compound xia(T4) chang(T3) in citation form. Right: Neutral tone compound xia(T4) chang(T0) in citation form. The waveform is represented on the top part, and the spectrogram (gray scale), the F0 (blue line), and the intensity (yellow line) are represented on the bottom part.

As for neutral tone morphemes, although most of them are historically derived from their full tone counterparts, however, in modern Chinese Mandarin, they are now phonologically and semantically distinct from their full tone counterparts today (Cao, 1992; Shen, 1992, 1993). E.g., in neutral tone compound 下场 /xia4chang0/ (end of life), the second constituent 场 /chang0/ has a neutral tone, its intrinsic/original full tone is Tone3, that is the second constituent of the full tone compound 广场 /guang3chang3/ (field/place). Accordingly, the second full tone constituent 场 /chang3/ is namely the ‘full tone counterpart’ of the neutral tone constituent 场 /chang0/. As for two-character neutral tone compounds, the intrinsic tone pattern of the second neutral tone constituent can be one of the four full tones, T1, T2, T3, or T4 (Table.5.2.1).

Table.5.2.1. Examples of neutral tone compounds, in which their second neutral tone constituents bear intrinsic full tone patterns as Tone 1/2/3/4.

Neutral tone compound	2 nd constituent (C2)	Full tone counterpart of C2	Intrinsic /original tone pattern of C2
亲家 /qin1jia0/ Family of marriage	家/jia0/ Family/home	家/jia1/ Family/home	Tone 1
打量 /da3liang0/ Look over	量/liang0/ Measure/estimate	量/liang2/ Measure/estimate	Tone 2
下场 /xia4chang0/ End of life	场/chang0/ Field/courtyard	场/chang3/ Field/courtyard	Tone 3
大意 /da4yi0/ careless	意/yi0/ Meaning/intention	意/yi4/ Meaning/intention	Tone 4

The opposition between the initial full tone and the second neutral tone within disyllabic neutral tone compounds provides us a good case to analyse the contextual tonal variation between the two constituents. In the current study, we aimed to explore the acoustic features of the second neutral tones in relation to their counterparts in four full tones (Tone1/2/3/4) within neutral tone compounds. Our goal is two-fold. Firstly, we aim to provide a detailed description of the second neutral tone within two-character compounds, we assumed that their acoustic features are much reduced, their duration of neutral tone is significantly shorter, and their intensity is much weaker than the full tone counterpart in the four full tone contexts, and the pitch contour of neutral tone has a smaller pitch range and lower pitch level than full tones. Secondly, we aim to investigate the differences between neutral tone and their full tone counterparts in the phonological form.

5.2.1 Linguistic materials

Materials 160 two-character compounds were selected from the Contemporary Chinese dictionary (Zhao, 2015), including 80 neutral tone compounds and 80 full tone compounds, and 80 minimal pairs were constructed (see Appendix 1). The frequency and stroke number of selected compounds were calculated based on the Corpus of Contemporary Mandarin which is created by the China State language commission (<http://corpus.zhonghuayuwen.org/Resources.aspx>).¹ Here, the 80 minimal pairs of compounds were selected based on the intrinsic tone pattern of the second neutral tones, one neutral tone compound [Strong-weak] is paired with one full tone compound [Strong-strong], and their second constituents (C2) possesses the identical syllable structure with either a neutral tone (T0) or a full tone (T1/2/3/4). Thus, four conditions were constructed based on the tone patterns of paired C2s (C1C2(T1) vs.C1C2(T0); C1C2(T2) vs.C1C2(T0); C1C2(T3) vs.C1C2(T0); C1C2(T4) vs.C1C2(T0)), and 20 minimal pairs of compounds were in each condition (Table.5.2.2). Two-way ANOVAs analysis was applied for the lexical frequency and stroke number of selected items, the prosodic pattern of the paired compounds (neutral tone & full tone), and the tone pattern of the second constituents (Tone 1/2/3/4) are the within-subject factors. Results revealed that the selected compounds were matched in the lexical frequency (Mean [neutral tone] =284.02 occurrences per ten million, mean [full tone] =497.86 occurrences per ten million, $t = -1.525$, $df = 106.83$, $p > .05$) in the stroke number (Mean [neutral tone] =16.42, Mean [full tone] = 15.21, $t = 1.65$, $df = 155.51$, $p > .05$). All the 160 selected compounds are embedded in the same position of the carrier phrase, in total 960 phrases are recorded by six participants.

¹ Corpus of contemporary Mandarin is an on-line word-text corpus constructed by China State language commission since 1998, it comprises commonly used written and spoken Chinese from a range of sources, including approximately 10 million words (20 million in characters).

Example: 我说了下巴这个词两遍。

[I said the word chin twice.]

Table.5.2.2. Examples of selected compounds in the four conditions. Each condition consists of 20 pairs of compounds, each pair consists of a neutral tone compound (SW: Strong-weak) and a full tone compound (SS: Strong-Strong), the second constituents have an identical structure and carry either a neutral tone or a full tone.

Compound	Condition 1 (T0 vs. T1)	Condition 2 (T0 vs. T2)	Condition 3 (T0 vs. T3)	Condition 4 (T0 vs. T4)
C1C2 (Tone 0) [Strong-weak]	xia4 <u>ba</u> 0 下巴 <n.> Jaw	mei2 <u>mao</u> 0 眉毛 <n.> Eyebrow	jie4 <u>zhi</u> 0 戒指<n.> Ring	jiao4 <u>huan</u> 0 叫唤<v.> Cry out
C1C2 (Full tone) [Strong-Strong]	zui3 <u>ba</u> 1 嘴巴<n.> Mouth	yu3 <u>mao</u> 2 羽毛<n.> Feather	shou3 <u>zhi</u> 3 手指<n.> Finger	zhao1 <u>huan</u> 4 召唤<v.> Call

5.2.2 Results

In this analysis, target syllables are the second constituents (C2) within the two-character compounds, after the annotation in Praat, we applied three Praat scripts respectively to extract the syllable duration (s), intensity (dB), pitch range (Hz), pitch level (Hz) and pitch transition (Hz) of each target syllable (the second constituent of a compound) from the recordings, and by using the RStudio, the Repeated ANOVAs analysis and paired t-tests are performed separately for the three categories of acoustic cues (duration, F0 and intensity).

Duration

Repeated two-way ANOVAs were performed to analyze the effect of prosody pattern and tone pattern on the duration of paired C2s. Prosody pattern of compounds (two levels: NT: neutral tone, FT: full tone) and Tone pattern of the paired C2s (four levels: Tone1/2/3 4 vs. Tone0) were considered as the within-subject factors, the duration of C2 was the dependent variable. Results of ANOVAs revealed a main effect of Prosody pattern [$F(1,959) = 228.36, p < .001, MSE = 0.005, \eta^2_p = 0.19$] (Fig.5.2.2), indicating that on average, the duration of C2 was shorter in neutral tone (mean: 0.22s, SD = 0.05) than in full tone (mean: 0.29s, SD = 0.08). A main effect of Tone pattern [$F(3,957) = 9.49, p < .001, MSE = 0.006, \eta^2_p = 0.02$] was revealed, suggesting that the duration of neutral tone C2 was significantly shorter than full tone C2 across the four full tones. Finally, the interaction between the Prosody pattern and Tone pattern [$F(3,957) = 3.072, p < .05, MSE = 0.004, \eta^2_p < .001$] was found. The mean duration of neutral tone C2 and full tone C2 is respectfully 0.26s

and 0.29s in Tone 1 ($d = 0.03s$), 0.26s and 0.31s in Tone 2 ($d = 0.05s$), 0.27s and 0.30s in Tone 3 ($d = 0.03s$) and 0.26s and 0.29s in Tone 4 ($d = 0.03s$). Complementary post hoc analyses revealed that as for the durational difference of the paired C2 (neutral tone C2 vs. full tone C2) in Tone 1, the effect of the prosody pattern is significantly larger than in Tone 3 ($p < .01$) and in Tone 4 ($p < .05$), as for the durational difference of paired C2 (neutral tone C2 vs. full tone C2) in Tone 2, the main effect of prosody pattern is significantly larger than in Tone 3 ($p < .001$) and Tone 4 ($p < .001$). In addition, by taking the intrinsic tone patterns of neutral tone C2s as a factor (4 levels: intrinsic tone patterns, Tone 1/2/3/4), a one-way ANOVAs was run for the duration of neutral tones, results revealed no significant difference in the mean duration of the second neutral tones considering their four intrinsic tones ($p > .05$) (Fig.5.2.3).

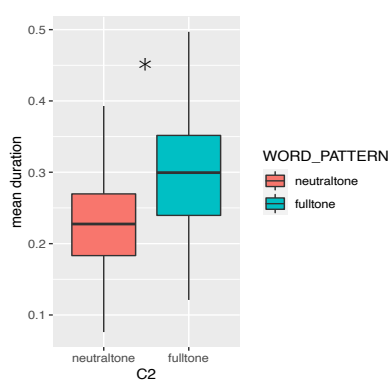


Fig.5.2.2. Duration(s) of paired second constituents (C2) in neutral tone (red) & full tone (green).

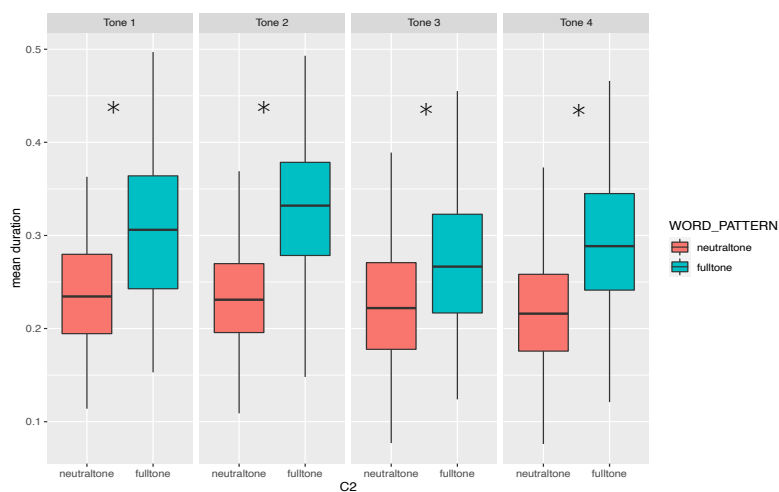


Fig.5.2.3. Duration(s) of paired C2s in neutral tone (red) and full tone (green) in four intrinsic tone patterns.

Pitch F0

Three parameters were calculated to analyze the fundamental frequency (F0) of the second constituent, including the pitch range, pitch level, and pitch transition (between the offset of first full tone and the onset of second neutral tone). Two-way ANOVAs were performed respectively

for pitch range and pitch level of paired C2s. The main effect of Prosody pattern (2 levels: neutral tone, full tone) was revealed [$F(1,96) = 25.26, p < .001, MSE = 2613.760, \eta^2_p = 0.03$], indicating that the averaged pitch range of neutral tone C2 (mean: 50.06Hz, $SD = 45.75$) was significantly narrower than the full tone C2 (mean: 66.66Hz, $SD = 56.07$). Moreover, a significant main effect of Prosody pattern was revealed for the pitch level [$F(1,959) = 6.507, p < .05, MSE = 2563.083, \eta^2_p < 0.001$], indicating that the averaged pitch level of neutral tone C2 was lower (mean: 183.45Hz, $SD = 49.73$) than the full tone C2 (mean: 191.79Hz, $SD = 51.58$) (Fig.5.2.4).

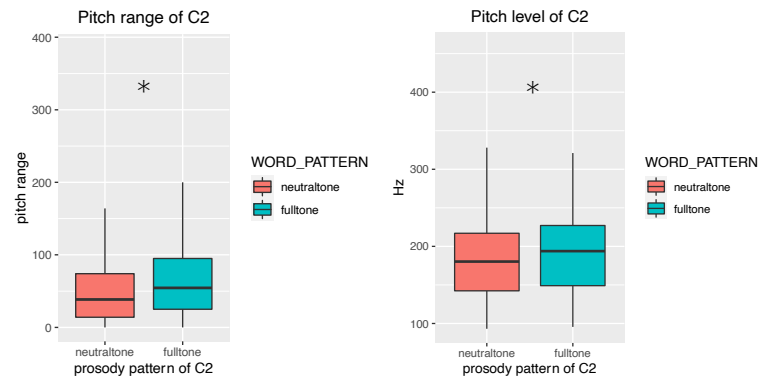


Fig.5.2.4. Left: pitch range (Hz) of paired C2s in neutral tone (red) & full tone (green); Right: pitch level (Hz) of paired C2s in neutral tone (red) & full tone (green).

The mean pitch range of neutral tone C2 and full tone C2 is respectively 46.25Hz and 45Hz in Tone 1 ($d = 1.25$ Hz), 50.85Hz and 57.25Hz in Tone 2 ($d = 6.4$ Hz), 54.70Hz and 74.02Hz in Tone 3 ($d = 19.95$ Hz), 48.42Hz and 92.91Hz in Tone 4 ($d = 44.49$ Hz). The Repeated two-way ANOVA showed a significant main effect of Prosody pattern [$F(1,959) = 26.98, p < .001, MSE = 2613.76, \eta^2_p = 0.03$] as well as a main effect of Tone pattern [$F(3,957) = 13.12, p < .001, MSE = 2582.12, \eta^2_p = 0.04$]. The Prosody pattern \times Tone pattern interaction also reached the significance level [$F(3,96) = 10.65, p < .001, MSE = 2431.617, \eta^2_p = 0.03$], indicating that the mean pitch range of neutral tone C2 was significantly smaller than their full tone counterparts in the four full tone contexts. Further post hoc tests revealed that as for the pitch range difference between the paired C2s (neutral tone C2 vs. full tone C2), the effect of the Prosody pattern in Tone 3 was significantly larger than in Tone 1 ($p < .001$). As for the difference in pitch range between the paired C2s (neutral tone vs. full tone), the effect of Prosody pattern in Tone 4 was significantly larger than in Tone 1 ($p < .001$) and Tone 2 ($p < .01$). In addition, by taking the intrinsic tone patterns of neutral tone C2s as a factor (4 levels: intrinsic tone patterns, Tone 1/2/3/4), a one-way ANOVAs was run for the pitch range of neutral tones, results revealed no significant difference in the mean pitch range of the second neutral tones considering their four intrinsic tones ($p > .05$) (Fig.5.2.5).

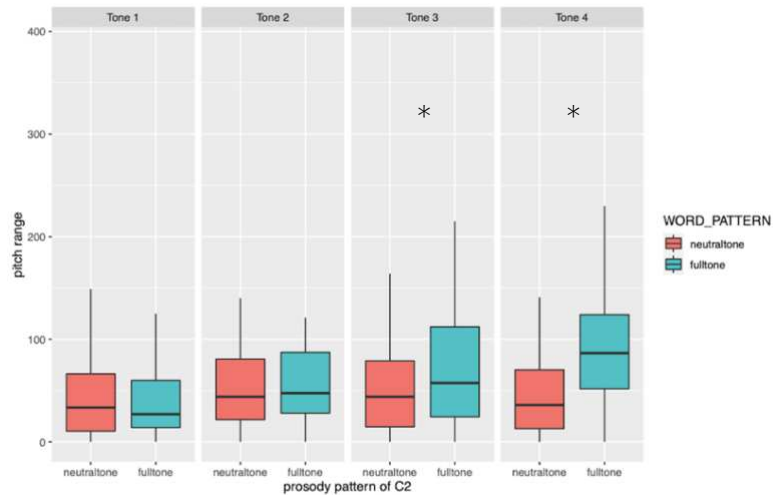


Fig.5.2.5. Pitch range (Hz) of paired C2s in neutral tone (red) & full tone (green) in four intrinsic tone patterns.

The mean pitch level of neutral tone C2 and full tone C2 is respectively 183.60Hz and 196.72Hz in Tone 1 ($d = 13.12\text{Hz}$), 183.01Hz and 172.26Hz in Tone 2 ($d = 10.75\text{Hz}$), 188.21Hz and 189.49Hz in Tone 3 ($d = 1.28\text{Hz}$), 178.97Hz and 208.69Hz in Tone 4 ($d = 29.72\text{Hz}$). Two-way ANOVAs showed a main effect of Prosody pattern [$F(1,959) = 6.708, p < .01, MSE = 2563.083, \eta^2_p < 0.001$] and a main effect of Tone pattern [$F(3,957) = 4.696, p < .01, MSE = 2543.932, \eta^2_p = 0.01$], and the interaction between Prosody pattern and Tone pattern reached the significance level [$F(3,957) = 7.179, p < .001, MSE = 2470.632, \eta^2_p = 0.02$], suggesting that the pitch level of neutral tone C2 was significantly lower than full tone C2 across the four full tones. Further post hoc tests revealed that as for the difference in pitch level between paired C2s (neutral tone C2 vs. full tone C2), the effect of Prosody pattern in Tone 4 was significantly larger than in Tone 2 ($p < .01$), and the effect of Prosody pattern in Tone 1 was significantly larger than in Tone 2 ($p < .05$). In addition, by taking the intrinsic tone patterns of neutral tone C2s as a factor (4 levels: intrinsic tone patterns, Tone 1/2/3/4), a one-way ANOVAs was run for the pitch level of neutral tones, results revealed no significant differences in the mean pitch level of the second neutral tones considering their four intrinsic tones ($p > .05$) (Fig.5.2.6).

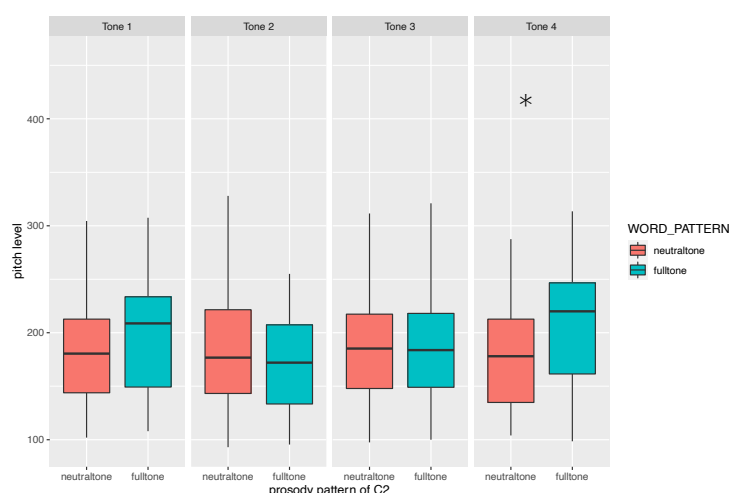


Fig.5.2.6 Pitch level (Hz) of paired C2s in neutral tone (red) & full tone (green) in four intrinsic tone patterns.

To analyze the pitch contours of two-character neutral tone and full tone compounds in the four full tone contexts, averaged pitch contour of the minimal pairs in the four categories was calculated (Fig.5.2.7 and Fig.5.2.10), each syllable was divided into ten segments and normalized in time, the averaged pitch value of the first segment and the last segment was defined as the pitch onset and offset.

C2(Tone 1) vs. C2(Tone 0)

Paired t-tests were performed respectively for pitch range and pitch level of the paired C2 (Tone 1 vs. Tone 0) as a factor. Results revealed no significant difference in pitch range and pitch level between them (T1 vs. T0) ($p > .05$), and paired t-tests were run for the pitch transition between the offset of full tone C1 and the onset of neutral tone C2 in the context of Tone 1, no significant differences were revealed for pitch transition ($p > .05$), the onset pitch value of second neutral tone is revealed to be not significantly different from the offset pitch value of first full tone in Tone 1.

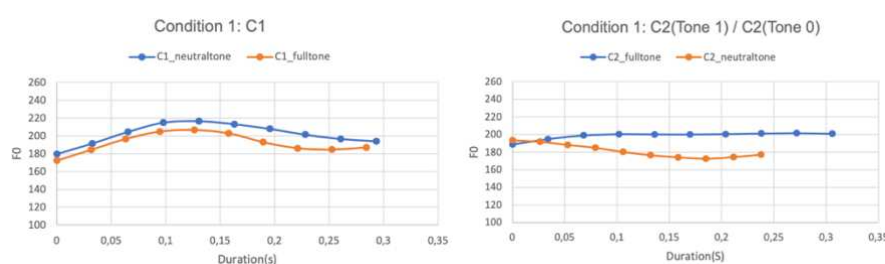


Fig.5.2.7. Averaged F0 curve of compound pairs in Condition 1(C1C2(Tone 1) vs. C1C2 (Tone 0)). Left: pitch contour of C1 respectively in neutral tone (blue) and full tone (orange); Right: pitch contour of C2 respectively in neutral tone (blue) and full tone (orange).

C2(Tone 2) vs. C2(Tone 0)

Again here, paired t-tests were performed respectively for the pitch range and pitch level of the second constituents with Tone pattern (Tone 2 vs. Tone 0) as a factor. Results indicated no significant difference in pitch range and pitch level of the paired C2s (T2 vs. T0) ($p > .05$). As for the pitch transition between the full tone C1 offset and the neutral tone C2 onset in the context of Tone 2, no significance was revealed ($p > .10$), the onset pitch value of second neutral tone is revealed to be not significantly different from the offset value of first full tone in Tone 2.

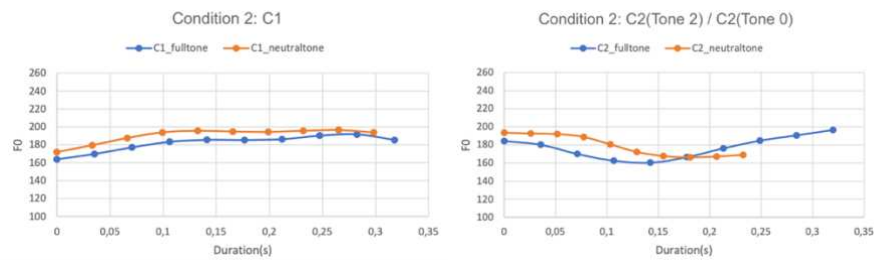


Fig.5.2.8. Averaged F0 curve of compound pairs in Condition 2 (C1C2(Tone 2) vs. C1C2 (Tone 0)). Left: pitch contour of C1 respectively in neutral tone (blue) and full tone (orange). Right: pitch contour of C2 respectively in neutral tone (blue) and full tone (orange).

C2(Tone 3) vs. C2(Tone 0)

Paired t-tests were run on the pitch range and pitch level of the second constituents, using Tone pattern (Tone 3 vs. Tone 0) as a factor. A significant difference was revealed for pitch range between the paired C2s (T3 vs. T0) [$t = -2.562, p = .011$], but not for pitch level ($p > .05$). As for the pitch transition between the full tone C1 offset and the neutral tone C2 onset in the context of Tone 3, results failed to show a significant effect ($p > .10$), the pitch onset of the second neutral tone is revealed to be not significantly different from the offset value of first full tone in Tone 3.

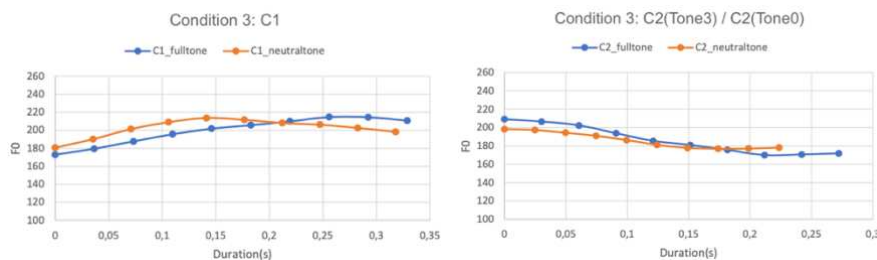


Fig.5.2.9. Averaged F0 curve of compound pairs in Condition 3 (C1C2(Tone 3) vs. C1C2 (Tone 0)). Left: pitch contour of C1 respectively in neutral tone (blue) and full tone (orange). Right: pitch contour of C2 respectively in neutral tone (blue) and full tone (orange).

C2(Tone 4) vs. C2(Tone 0)

Paired t-tests were run on both pitch range and pitch level of the second constituent, and Tone pattern (Tone 4 vs. Tone 0) as a factor. A significant effect was observed for pitch range [t

= -6.496, $p < .001$] and pitch level between the paired C2s (T4 vs. T0) [$t = -4.478$, $p < .001$]. As for the pitch transition between the full tone C1 offset and the neutral tone C2 onset in the context of Tone 4, a significance is revealed between the C1 offset (mean = 211.51 Hz) and the C2 onset (mean = 183.40 Hz) [$t = -3.46$, $p < .01$], the pitch onset of second neutral tone is revealed to be significantly lower than the offset of first full tone in Tone 4.

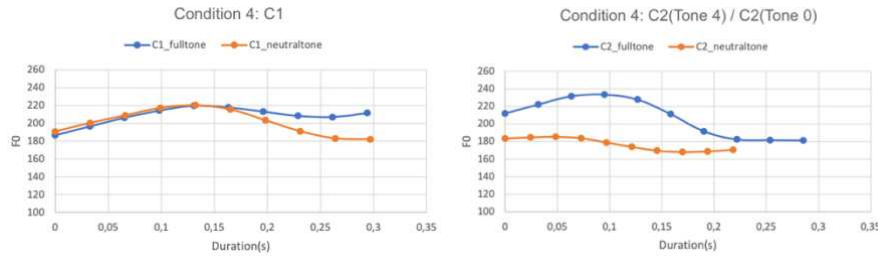


Fig.5.2.10. Averaged F0 curve of compound pairs in Condition 4 (C1C2(Tone 4) vs. C1C2 (Tone 0)). Left: pitch contour of C1 respectively in neutral tone (blue) and full tone (orange). Right: pitch contour of C2 respectively in neutral tone (blue) and full tone (orange).

Intensity

Repeated two-way ANOVAs were performed for the effect of the Prosody pattern and Tone pattern on the intensity of the paired second constituents. Prosody pattern (Strong-weak & Strong-strong) and Tone pattern (Tone 1/2/3/4) were considered as the within-subject factors, and the intensity of C2s was the dependent variable. Results of ANOVAs analysis revealed a main effect of Prosody pattern [$F(1,959) = 55.34$, $p < .001$, $MSE = 42.782$, $\eta^2_p = 0.05$], indicating that the mean intensity of neutral tone C2 is significantly lower (mean = 51.97dB, $SD = 6.91$) than full tone C2 (mean = 53.47dB, $SD = 6.71$) (Fig.5.2.11). A main effect of the Tone pattern reached the significance level [$F(3,957) = 29.8$, $p < .001$, $MSE = 41.437$, $\eta^2_p = 0.08$], revealing that on average, the intensity of neutral tone C2 was significantly lower than the full tone C2 across four full tone contexts. Results indicated an interaction between Prosody pattern and Tone pattern [$F(3,957) = 6.61$, $p < .001$, $MSE = 38.425$, $\eta^2_p = 0.02$], the mean intensity of neutral tone C2 and full tone C2 are respectfully 52.83dB and 57.62dB in Tone 1 ($d = 4.97$ dB), 52.89dB and 55.73dB in Tone 2 ($d = 2.84$ dB), 50.23dB and 50.36dB in Tone 3 ($d = 0.13$ dB), and 51.95dB and 56.10dB in Tone 4 ($d = 4.15$ dB)(Fig.5.2.12). Post hoc tests showed that as for the paired C2s (neutral tone C2 vs. full tone C2) in Tone 3, the effect of the Prosody pattern is much smaller than in Tone 1 ($p < .001$), Tone 2 ($p < .001$) and Tone 4 ($p < .001$). Paired t-tests were performed for each comparison across the four full tone patterns (C2(T0) vs. C2(T1); C2(T0) vs. C2(T2); C2(T0) vs. C2(T3); C2(T0) vs. C2(T4)), the results failed to reveal the significant difference between the neutral tone C2 and full tone C2 in the Tone 3 [$t = 0.8783$, $p > .05$]. In Tone 1/2/4, the intensity of the second neutral tones was

as higher as their full tone counterparts, in Tone 3, the mean intensity of neutral tone C2 and full tone C2 is respectively 50.23dB & 50.36dB, revealing almost no difference ($p > .05$). In addition, by taking the intrinsic tone patterns of neutral tone C2s as a factor (4 levels: intrinsic tone patterns, Tone 1/2/3/4), one-way ANOVAs was run for the duration of neutral tones, results revealed the significant difference in the intensity of neutral tone C2s in context of their intrinsic tone patterns [$F(3,477) = 3.91, p < .001, MSE = 42.782, \eta^2_p = 0.02$].

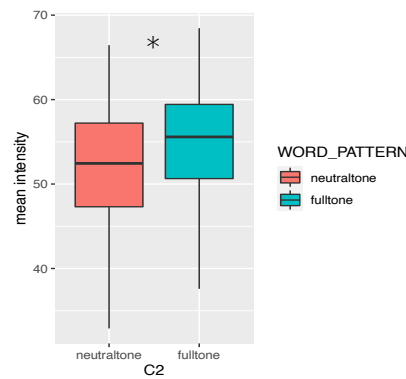


Fig.5.2.11. Mean intensity (dB) of paired C2s in neutral tone (red) and full tone (green).

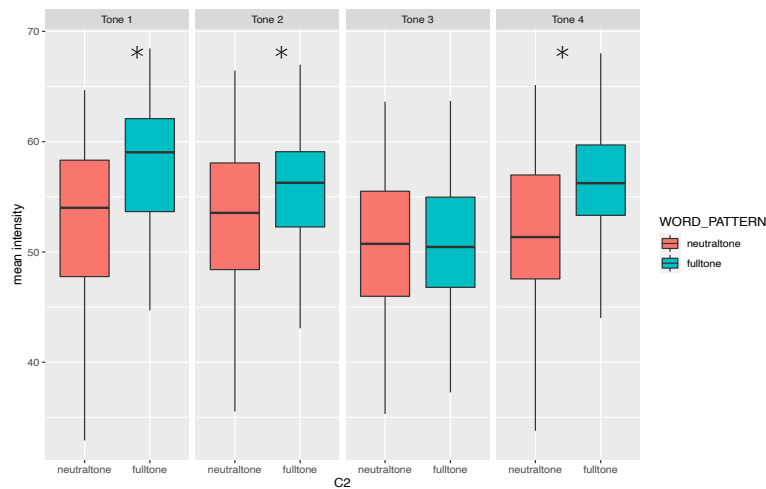


Fig.5.2.12. Intensity(dB) of paired C2s in neutral tone(red) & full tone(green) in four intrinsic tone patterns.

5.2.3 Discussion

To investigate the prosodic features of the second neutral tones within the two-character compounds, we designed this acoustic study by comparing eighty minimal pairs of compounds which carry the contrastive prosody patterns on the second constituents. Acoustic correlates of the second syllable are measured and analyzed by Praat and RStudio across four full tone contexts, including syllable duration, pitch (F0) and intensity. In accordance with the previous research, our investigation confirms that within compounds, neutral tone is phonologically weaker and lower in

contrast to full tone. An exception is observed in Tone 3, on average, the intensity of neutral tone and full tone indicates no significant difference in the third tone. As for pitch (F0), no significant difference was revealed in pitch range between the second neutral tones and their counterpart full tones, whereas the pitch level of the neutral tones was significantly lower than full tones in Tone 3 and Tone 4. Consistent with the previous studies, the pitch onset of the second neutral tones is revealed to be dependent on the pitch offset value of the preceding first full tones.

By analyzing the eighty minimal pairs of compounds in the four categories based on the tone patterns of the second constituents, we found that the duration of neutral tone is about 70%-80% of the full tone counterpart. And as for the duration of neutral tones, no significant difference was revealed considering their four intrinsic full tone patterns. This result slightly diverges from the findings of Lin and Yan (1990), which claimed that the duration of the neutral tone of disyllabic compounds is on average about 50% of full tones. Furthermore, a significant difference in intensity was observed between the second neutral tones and full tones in the tonal contexts of Tone 1/2/4. In Tone 3, the mean intensity of the second neutral tones was as higher as full tones. This could be viewed as consistent with Wai-Sum and Eric (2014), and Tang (2014), which have claimed that neutral tone is not definitely weaker than full tone. In our result, it is noteworthy that in Tone 3, we found no difference in intensity between neutral tones and their full tone counterparts.

As for the pitch target of the second neutral tones, considering their four intrinsic full tone patterns, in general, we found no significant differences in their pitch range and pitch level. More precisely, the pitch range of the second neutral tones and their full tone counterparts revealed no significant difference in the contexts of Tone 1/2/3, except in Tone 4. This suggests that the pitch range of the second neutral tone within two-character compounds is not definitely narrower than full tone as our assumption. What's more, the pitch level of the second neutral tone is significantly lower than full tone in the contexts of Tone 3 and Tone 4. However, in Tone 1 and Tone 2, the pitch level of neutral tone is as higher as full tone. According to Xu (1997), considering the acoustic features of the four lexical full tones, Tone 1 and Tone 2 possess a relatively higher pitch level in relation to Tone 3 and Tone 4. Our results indicate that when their intrinsic tone patterns have a lower pitch level (Tone 3 & Tone 4), neutral tones have a significantly lower pitch level, in relation to the intrinsic tone patterns possessing a high pitch level (Tone 1 & Tone 2), the neutral tone's pitch level is as high as their full tone counterparts. As for the pitch transition between the first full tone and the second neutral tone, no significant result was revealed in Tone 1/2/3, except in Tone 4. This confirms the previous findings that the pitch onset of neutral tone approximates the pitch offset of the preceding full tone (Cao, 1986; Lee & Zee, 2008; Li, 2003; Wai-Sum & Eric, 2014; Yang, 1991), and the pitch contour of neutral tones could be dependent and co-varying with

the preceding full tone. In our results, within the two-character neutral tone compounds, the pitch offset value of the first full tone determines the pitch onset value of the second neutral tone, except in the tonal context of Tone 4. In a previous study, Chen and Xu (2006a) explained the articulatory mechanism of neutral tone, the weaker articulatory strength and slower speed make it have a large variability in the surface pitch contour. Compared with the strong articulatory strength and the distinctive phonological form of full tones, neutral tones receive more impact from adjacent tones, especially from the preceding strong full tones.

5.2.4 Conclusion

In our results, the common features between the neutral tone and full tone are mainly revealed in the intensity and pitch level. When the intrinsic tone pattern of neutral tones is Tone 3, the intensity is as high as their full tone counterparts. The pitch onset of neutral tone is revealed to be determined by the pitch offset of the preceding full tone, which is consistent with previous studies. When its intrinsic tone pattern has a higher pitch level (Tone 1 and Tone 2), the pitch level of neutral tone is as high as its full tone counterpart, this can be evidence of the common acoustic features shared by the neutral tone and their full tone counterpart. Taken together, this acoustic analysis reinforces the idea that within two-character compounds, prosodic features of the second neutral tone compounds significantly differ from the second full tones, we found that the second neutral tone presents a consistency in their phonetic realization across the four intrinsic full tone patterns, our study contributes to the future study as to explore the acoustic correlates of the second neutral tone under the influence of preceding tonal context of two-character compounds, which could help us to better understand the tonal variation.

5.3 Contextual tonal variation of neutral tone within the Mandarin two-character compounds

In tone language, information contained in the suprasegmental feature is used to express distinct semantic meanings. As we know that Mandarin has four lexical full tones (Tone 1/2/3/4), in continuous speech, the acoustic realization of syllables varies with the neighboring syllables, the contextual tonal variation exists due to the coarticulation effects. Acoustic correlates such as pitch (F0), intensity and duration can be varying in different tonal contexts. Previous studies proposed that Mandarin tonal variation occurs in both two directions between the successive syllables, which are the carry-over effect and the anticipatory effect (Shen, 1990; Zhang & Kawanami, 1999). The carry-over effect is found to be assimilatory, which indicates that the pitch onset of a syllable is

influenced by the preceding tonal context, it can be assimilated to the offset of its preceding tone. While the anticipatory effect is dissimilatory in the reverse direction, indicating that the low pitch offset of a tone sometimes raises the maximum to the onset pitch value of the following tone.

Concerning the Mandarin tonal variation, two main propositions exist in previous studies. In the early years, tonal variation was claimed to exist only in the vowel segments, coarticulation between two adjacent tones occurs only in the left-to-right direction, the preceding tone was not affected by the following tone (Howie, 1974). During the years of 1990s, this concept was revised by acoustic studies. In a continuous speech, the anticipatory adjustment effect from the consonant segment to its preceding syllable was examined and identified by Shen (1990) and Shen and Lin (1991). In the late 1990s, by examining the tonal coarticulation effect depending on the different phonological contexts, Zhang and Kawanami (1999) proposed that Mandarin tonal coarticulation might be unstable and flexible between two neighboring tones, which could be uni-directional, bi-directional or in both two directions. Xu (1997) analyzed the coarticulation between neighboring tones within Mandarin compounds, results provided the evidence of tonal variation in the full tone compounds. These studies had been so far focused on the four canonical full tones, and less study has examined the tonal variation within two-character compounds which possess a second neutral tone. We raised our questions from the previous study: what is the magnitude and extent of neutral tone in the coarticulation effects? Are the acoustic correlates of neutral tones within compounds coextensive with preceding tones? An acoustic analysis based on the production of native speakers was constructed to answer the questions. Tonal variation contained in the trochaic prosody pattern of neutral tone compounds was less discussed in the previous study, and the contrastive metric pattern possesses by the first full tone and second neutral tone provides a good case to analyze the coarticulation between syllables, we thus constructed this phonetic analysis. The goal of this study is two-fold. Firstly, we aim to analyze the assimilatory effects from the preceding full tone to the second neutral tone. In sub-chapter 6.3.1, we explored how the second neutral tone phonetically varies regarding the four preceding tones. Secondly, we examined the dissimilatory effects from the second neutral tone to the preceding tones. In sub-chapter 6.3.2, two-character compounds which have the same first constituents and second constituents in four full tones and neutral tones were selected, and the anticipatory effect from the second (rightmost) neutral tone to its preceding first full tone is analyzed. In our hypothesis, we assumed that the tonal variation induced by both carry-over and anticipatory effect will be observed between the two constituents of neutral tone compounds. When preceded by the first constituents in four full tones, the duration, pitch contour and intensity of the second neutral tone will be variant, and the first full tone will be phonetically affected by the following neutral tone in the second position, which will be phonetically reduced.

Instead of the coarticulation effect in uni-direction, the tonal variation would occur in bi-directions within two-character neutral tone compounds.

5.3.1 Carry-over effect

Within two-character compounds, the two constituents can interact with each other in the phonological form. This analysis concerned the carry-over effect from the first full tone to the second neutral tone. Considering the evidence of the assimilatory effect from the first syllable to the second syllable within the two-character full tone compounds (Xu, 1997), we assumed that the acoustic features of the second neutral tones are coextensive with the preceding tonal contexts.

5.3.1.1 Linguistic materials

Materials 80 neutral tone compound words [Strong-weak] were selected based on the tone patterns of their first constituents (Tone 1/2/3/4) as Condition 1/2/3/4, 20 compounds were in each condition (see Appendix 2-1). Furthermore, another 20 two-character full tone compounds [Strong-strong] were selected as the control group in the Condition 5 (Table.5.3.1). Repeated ANOVAs analysis were applied for the frequency of occurrences and stroke number for selected compounds in five conditions, they were matched in frequency [Mean [frequency] = 262.64 per ten million; $F(4,96) = 1.111, p > .05, MSE = 4820.9, \eta^2_p = 0.04$] and in stroke number [Mean [stroke number] = 17.59; $F(4,96) = 1.166, p > .05, MSE = 22.821, \eta^2_p = 0.05$] in the Corpus of Contemporary Mandarin. All the target compounds were embedded in the same position in the carrier phrase, in total 600 phrases were produced by six participants, and the acoustic properties of the second constituents (C2) in the five conditions were examined.

Table 5.3.1. Examples of selected compounds in the five conditions. In Condition 1/2/3/4, each consists of 20 neutral tone compounds (SW: Strong-weak), the first constituents are respectively in Tone 1/2/3/4, the second constituents (C2) are neutral tones, there were 20 full tone compounds (FT: Strong-strong) as the control group.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5
Prosody	C1(T1)+C2(T0)	C1(T2)+C2(T0)	C1(T3)+C2(T0)	C1(T4)+C2(T0)	C1(FT)+C2(FT)
	Strong-Weak	Strong-Weak	Strong-Weak	Strong-Weak	Strong-Strong
Compound	招呼<v.> zhao1 hu 0 salute	粮食<n.> liang2 shi0 food	寡妇<n.> gua3 fu0 widow	厉害<adj.> li4 hai0 bravo	制服<v.> zhi4 fu2 overpower

5.3.1.2 Results

In this analysis, the target syllable is the second constituent (C2) within critical compounds. After the annotation in Praat, we used the Praat scripts to extract the acoustic measurements which were subject to the next analyses in RStudio. Repeated ANOVAs analysis and paired t-tests were performed separately for the three categories of acoustic correlates (duration, F0, and intensity). As for Condition 1-4, the one-way ANOVAs were performed respectively for their second neutral tones, the Tone pattern (Tone 1/2/3/4) of the first constituent (C1) is a within-group factor, and the measurement of the C2 is the dependent variable. Moreover, the paired t-test was conducted respectively between the second neutral tones in the four tonal contexts with the full tone C2.

Duration

Repeated ANOVAs analyses were performed for the duration of the second neutral tone (C2) under the influences of the preceding tones. Tone pattern (preceding tones: Tone 1/2/3/4) of C1 was the within-group factor, the duration of C2 was the dependent variable. Results revealed a main effect of the Tone pattern on the duration of C2 (neutral tone) [$F(3,477) = 11.15, p < .01, MSE = 0.003, \eta^2_p = 0.07$] (Fig.5.3.1), and the mean duration of neutral tone C2 was respectively 0.226s, 0.211s, 0.243s and 0.206s when preceded by Tone 1/2/3/4. Tukey HSD tests revealed the detailed effect of preceding tones. When preceded by Tone 4, C2 had the shortest duration; When preceded by Tone 3, C2 had the longest duration and significantly longer than preceded by Tone 2 ($p < .01$) and Tone 4 ($p < .01$). As for the full tone compounds [C1(Full tone) C2 (Full tone)] in the control group, the mean duration of C2 (full tone) was 0.305s. And to analyze the differences between the C2 (neutral tone) and C2 (full tone), paired t-tests were conducted respectively between the C2 (neutral tone) in the four tonal contexts with C2 (Full tone) in the control group. Results revealed the significant results respectively between C2 (neutral tone) preceded by Tone 1/2/3/4 and C2 (full tone) ($p < .01$), indicating that C2 (neutral tone) was significantly shorter than C2 (full tone) across the four preceding tonal contexts.

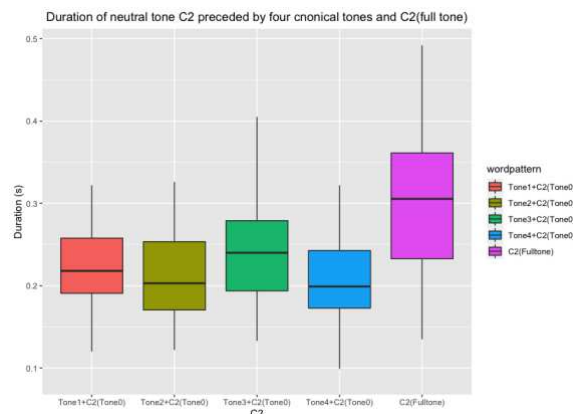


Fig.5.3.1. Duration(s) of C2 (neutral tone) when preceded by four full tones and C2 (full tone) in the control group.

Five colors are applied, red: Tone 1, olive: Tone 2, green: Tone 3, blue: Tone 4, and pink: control group.

Pitch F0

Repeated ANOVAs analyses were performed to analyze the influences from the preceding full tone contexts (Tone 1/2/3/4) to the pitch range and pitch level of the second neutral tones, with the tone patterns of C1 (Tone 1/2/3/4) as a within-group factor, pitch range (Hz) and pitch level (Hz) of C2 (neutral tone) as the dependent variable.

Firstly, results revealed a significant effect of the tone patterns of C1 on the pitch range of C2 (neutral tone) [$F(3,477) = 15.87, p < .01, MSE = 1777.236, \eta^2_p = 0.09$], the mean pitch range of C2 (neutral tone) was respectively 59.7 Hz, 48.7 Hz, 34.9 Hz, and 24.7 Hz when preceded by Tone 1/2/3/4, and 61.0 Hz for C2 (full tone) in the control group. Post hoc tests revealed the detailed effects of four preceding tones on the pitch range of neutral tone C2, the pitch range of C2 (neutral tone) was the largest when preceded by Tone 1, which was significantly larger than preceded by Tone 3 ($p < .01$) and Tone 4 ($p < .01$) (Fig.5.3.2(1)).

Secondly, a main effect of Tone pattern of C1 is revealed on the pitch level of C2 (neutral tone) [$F(3,477) = 20.3, p < .01, MSE = 2098.622, \eta^2_p = 0.11$]. Post hoc tests indicated that the mean pitch level of C2 (neutral tone) was respectively 201.2 Hz, 193.1 Hz, 161.4 Hz, and 169.4 Hz when preceded by Tone 1/2/3/4, and 198.3 Hz for C2 (full tone) in the control group. Post hoc tests revealed the detailed effects of preceding full tones on the pitch level of neutral tone C2, the pitch level of C2 (neutral tone) was the highest when preceded by Tone 1, which was significantly higher than preceded by Tone 3 ($p < 0.01$) and Tone 4 ($p < 0.01$) (Fig.5.3.2(2)).

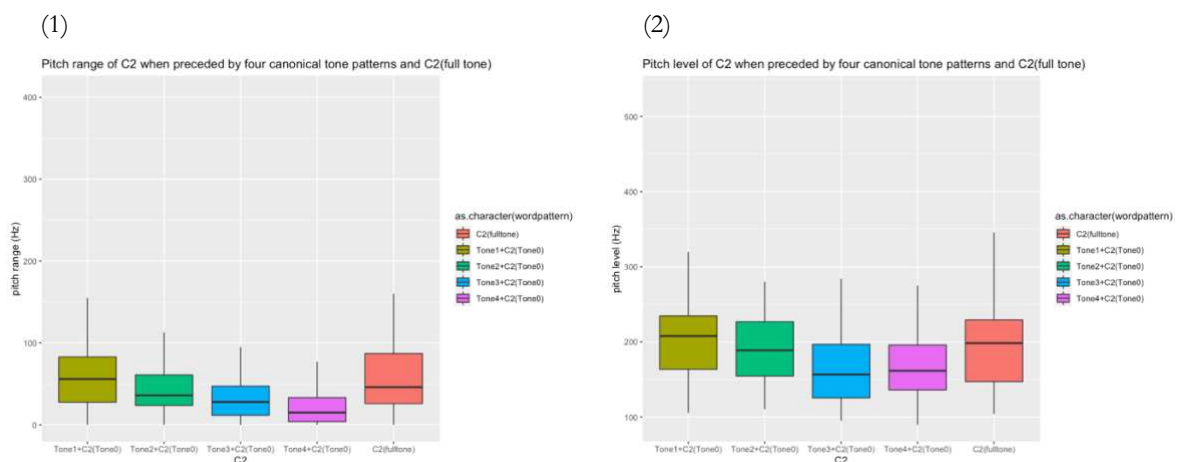


Fig.5.3.2. (1) Pitch range (Hz) of C2 (neutral tone) when preceded by four full tones and C2 (full tone) in the control group. (2) Pitch level (Hz) of C2 (neutral tone) when preceded by four full tones and C2 (full tone) in the control group. Five colors are applied, olive: Tone 1, green: Tone 2, blue: Tone 3, and pink: Tone 4, red: control group.

To present the pitch contours of compounds in the five conditions, the pitch value was measured respectively for C1(T1)C2(T0), C1(T2)C2(T0), C1(T3)C2(T0), C1(T4)C2(T0), C1(Full tone)C2(Full tone). For this purpose, we extracted the averaged F0 from ten points of the two constituents within the compound (Fig.5.3.3). Paired t-tests were applied respectively for the pitch transition between the offset of C1 (full tone) and the onset of C2 (neutral tone) in the four tonal contexts (Tone 1/2/3/4) within two-character compounds. And in this study, C2 is the target, by performing the paired t-tests, we compared the pitch level (Hz) and pitch range (Hz) between the C2 (neutral tone) when preceded by Tone 1/2/3/4 with C2 (Full tone) in the control group.

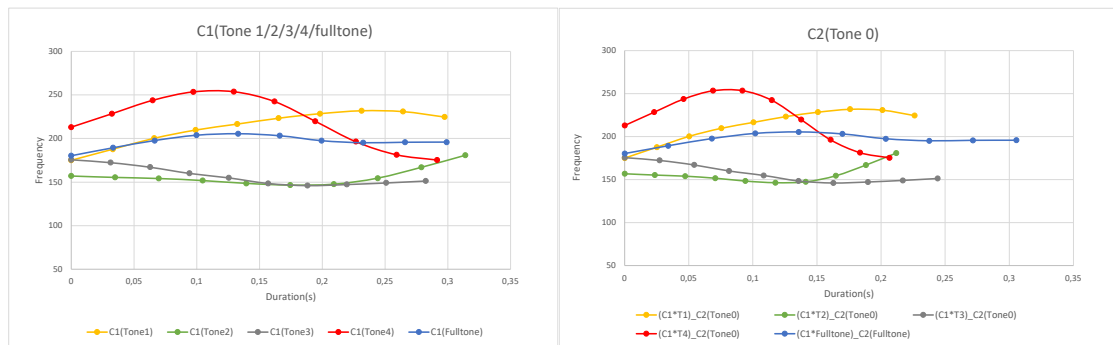


Fig.5.3.3. Averaged F0 curve of compounds in five conditions. The Pitch contours of C1 in Conditions (1/2/3/4) and the control group (full tone) are presented in the left figure, pitch contours of C2 (neutral tone) in Conditions (1/2/3/4) and control group (full tone) are represented in the right figure. Five colors are applied, yellow: Tone 1, green: Tone 2, grey: Tone 3, red: Tone 4, and blue: control group.

C1 (Tone 1) + C2 (Tone 0)

Paired t-test revealed no significance in the pitch transition ($p > .05$), the onset pitch value of the second neutral tone was revealed to be not significantly different from the offset pitch value of C1 (Tone 1). And when preceded by Tone 1, the pitch level and pitch range of C2 (neutral tone) revealed no significant difference in relation to C2 (full tone) in the control group (Fig.5.3.4).

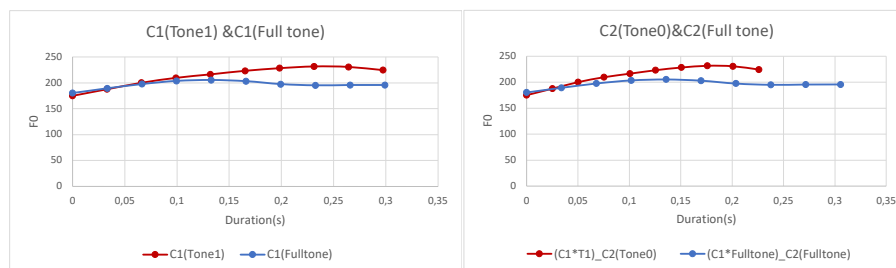


Fig.5.3.4. Averaged F0 curve of neutral tone compounds in Condition 1 and full tone compounds in the control group (C1(Tone 1) C2 (Tone 0) vs. C1(full tone) C2(full tone)). Left: pitch contour of C1 respectively in Tone 1 (red) and full tone (blue); Right: pitch contour of C2 respectively in neutral tone (red) and full tone (blue).

C1 (Tone 2) + C2 (Tone 0)

Paired t-test showed no significance in the pitch transition ($p > .05$), the onset pitch value of the second neutral tone was revealed to be not significantly different from the offset pitch value of C1(Tone 2). And when preceded by Tone 2, the pitch level and pitch range of C2 (neutral tone) revealed no significant difference in relation to C2 (full tone) in the control group (Fig.5.3.5).

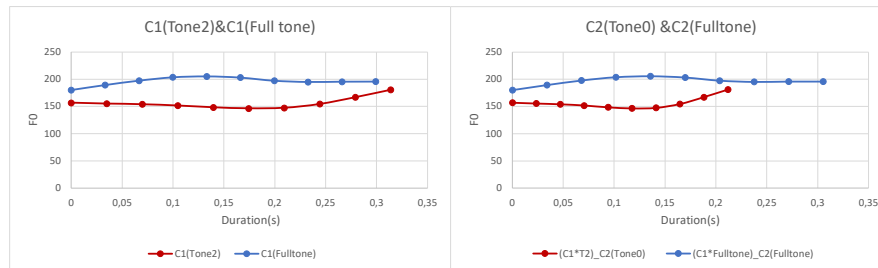


Fig.5.3.5. Averaged F0 curve of neutral tone compounds in Condition 2 and full tone compounds in the control group (C1(Tone 2) C2 (Tone 0) vs. C1(full tone) C2(full tone)). Left: pitch contour of C1 respectively in Tone 2 (red) and full tone (blue); Right: pitch contour of C2 respectively in neutral tone (red) and full tone (blue).

C1 (Tone 3) + C2 (Tone 0)

Paired t-test revealed no significance in the pitch transition ($p > .05$), the onset pitch value of the second neutral tone was revealed to be not significantly different from the offset pitch value of C1(Tone 3). When preceded by Tone 3, a significant difference in the pitch range was found between C2 (neutral tone) and C2 (control group) [$t = -0.708$, $p < .01$], a significant result in pitch level was found between the C2 (neutral tone) and C2 (full tone) in the control group [$t = -5.197$, $p < .01$] (Fig.5.3.6)

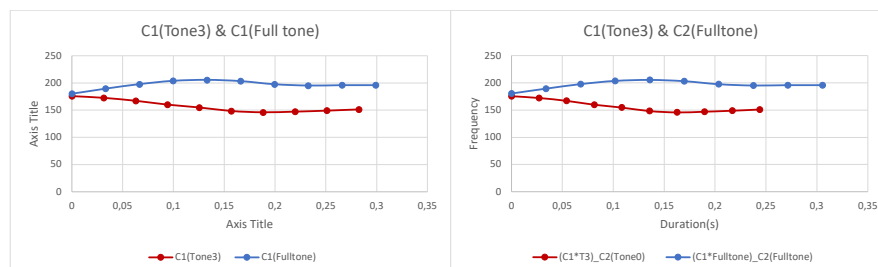


Fig.5.3.6. Averaged F0 curve of neutral tone compounds in Condition 3 and full tone compounds in the control group (C1(Tone 3) C2 (Tone 0) vs. C1(full tone) C2(full tone)). Left: pitch contour of C1 respectively in Tone 3 (red) and full tone (blue); Right: pitch contour of C2 respectively in neutral tone (red) and full tone (blue).

C1 (Tone 4) + C2 (Tone 0)

Paired t-test showed no significance in the pitch transition, the onset pitch value of the second neutral tone was revealed to be not significantly different from the pitch offset of C1(Tone 3) ($p > .05$). When preceded by Tone 4, a significant result was found in the pitch range between

C2 (neutral tone) and C2 (full tone) [$t = -5.729, p < .01$], a significant result in pitch level was found between C2 (neutral tone) and C2 (full tone) in the control group [$t = -4.049, p < .01$] (Fig.5.3.7).

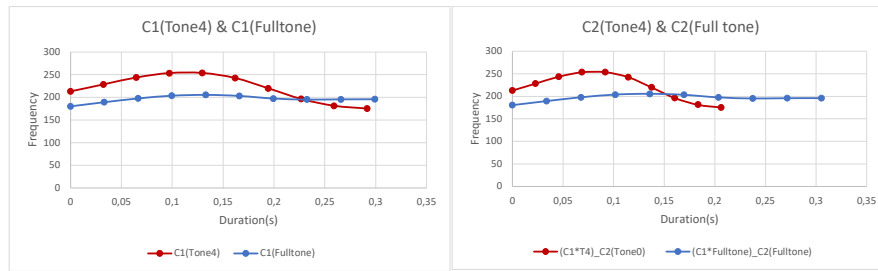


Fig.5.3.7. Averaged F0 curve of neutral tone compounds in Condition 4 and full tone compounds in the control group (C1(Tone 4) C2 (Tone 0) vs. C1(full tone) C2(full tone)). Left: pitch contour of C1 respectively in Tone 4 (red) and full tone (blue); Right: pitch contour of C2 respectively in neutral tone (red) and full tone (blue).

Intensity

Repeated ANOVAs were conducted for the intensity of the second neutral tone under the influence of preceding tonal contexts (Tone 1/2/3/4) (Fig.5.3.8). Tone pattern (Tone 1/2/3/4) of C1 was the within-group factor, the intensity of C2 was the dependent variable. Results revealed a significant effect of tone pattern on the intensity of the neutral tone C2 [$F(3,477) = 11.79, p < .01, MSE = 84.627, \eta^2_p = 0.07$] (Fig.5.3.8), and the mean intensity of neutral tone C2 was respectively 53.36Hz, 56.81Hz, 51.42Hz and 50.22Hz when preceded by Tone 1/2/3/4, the mean intensity of C2 (full tone) was 57.79Hz in the control group. Post hoc tests revealed a main effect of preceding tonal contexts on the intensity of C2 (neutral tone), it was the highest when preceded by Tone 2, which was significantly higher than in the context of Tone 1 ($p < .05$), Tone 3 ($p < .01$) and Tone 4 ($p < .01$). To calculate the intensity difference between C2 (neutral tone) and C2 (full tone), paired t-tests were conducted. Results revealed a significant difference respectively between the intensity of C2 (neutral tone) when preceded by Tone 1/3/4 with the C2 (full tone) ($p < .01$), except in the context of Tone 2, no significance was revealed.

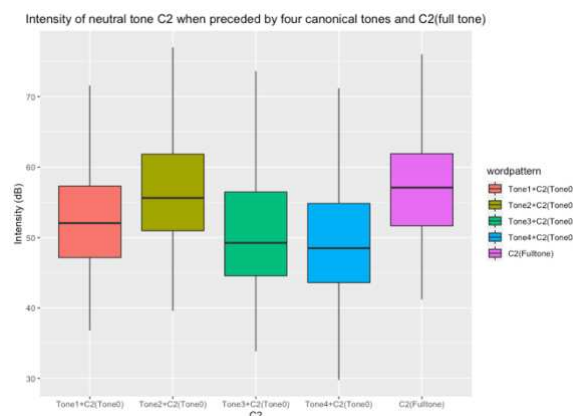


Fig.5.3.8. Intensity (dB) of C2 (neutral tone) when preceded by four full tones and C2 (full tone) in the control group. Five colors are applied, red: Tone 1, olive: Tone 2, green: Tone 3, blue: Tone 4, and pink: control group.

5.3.1.3 Discussion

This study investigated the carry-over effects from preceding tonal contexts to the second neutral tones within two-character compounds. Results indicated that the duration of the second neutral tone presents a significant difference when preceded by four full tones, it's shortest when preceded by Tone 4 and longest when preceded by Tone 3. Considering the acoustic production study of full tones by Xu (1997) and Lin (1988), the mean duration of the third tone (Tone 3) was claimed to be the longest, and Tone 4 was the shortest, our results revealed a significant carry-over effect from the preceding tones to the second neutral tones in the syllable duration. Furthermore, the intensity of the second neutral tone presented a significant difference when preceded by four full tones, and the intensity of the second neutral tone constituent was significantly higher when preceded by Tone 2 than the others. A main effect of preceding tones was revealed on both the pitch range and pitch level of the second neutral tones, which was significantly narrower and lower when preceded by Tone 3 and Tone 4. When preceded by Tone 1 and Tone 2, the pitch range and pitch level of the second neutral tones were phonetically close to the preceding tones. According to Xu (1997), the syllables in Tone 1 and Tone 2 possess relatively higher pitch levels than in Tone 3 and Tone 4, our results revealed the spreading influences from the preceding tones to the second neutral tones in the pitch within two-character compounds.

5.3.2 Anticipatory effect

As for the coarticulation effect within two constituents of compound words, not only the second constituent is influenced by the preceding tone (such as the phenomena of tone sandhi of Tone 3), the first constituent can be acoustically varied according to the following tones. In the case of two-character full tone compounds, the anticipatory effect was revealed to be much less important than the carryover effect (Xu, 2005), and the pitch onset of the second full tone was revealed to have little influence on the preceding tone (Xu, 1997). However, previous studies have not been concerned with the anticipatory effect from the neutral tone to its preceding full tone. According to the Target Approximation theory (Xu, 2005), we assumed an anticipatory effect from the second neutral tone to preceding tones, the global pitch contour of neutral tone compounds will be characterized by the syllable-synchronization. The goal of Analysis 2 was to examine the prosodic features of the first constituents (C1) of neutral tone compounds, in order to investigate the anticipatory effect from the second constituent to the preceding tones.

5.3.2.1 Linguistic materials

Materials 100 two-character compounds and 20 monosyllabic characters were selected as materials (see Appendix 2-2). We aim to analyze the anticipatory effect of the tone patterns of the second constituents to the first constituents. 20 neutral tone compounds were firstly selected, and the first constituents (C1) were extracted and then paired respectively with the second constituents (C2) in the four full tones (Tone 1/2/3/4). Thus, five conditions were thus constructed, including Condition 1: C1+C2(T1), Condition 2: C1+C2(T2), Condition 3: C1+C2(T3), Condition 4: C1+C2(T4) and Condition 5: C1+C2(T0) (Table.5.3.2). In Condition 6, the 20 first constituents (C1) were produced in isolation as the control group (Mean [frequency] = 504 occurrences per ten million, Mean [stroke number] = 7.15 \pm 4.8). Repeated ANOVAs were applied to the frequency of occurrence and stroke-number of selected compounds in Condition 1/2/3/4/5, and the selected compounds were matched in frequency [Mean [frequency] = 502.63 per ten million; $F(4,96) = 1.638, p > .05, MSE = 17.17, \eta^2_p = 0.06$] and in the stroke-number [Mean [stroke number] = 15.32; $F(4,96) = 0.576, p > .05, MSE = 10752, \eta^2_p = 0.02$]. All the 100 selected compounds and 20 monosyllabic characters were embedded in the carrier phrase, and acoustic properties of target syllables (C1) in different following tones were analyzed, in total 720 phrases were produced by the six participants.

Table 5.3.2. Examples of selected items in the six conditions. In Condition 1/2/3/4/5, each consists of 20 two-character compounds. In this example, in Condition 1/2/3/4/5, the compounds have the same first constituent (C1) ‘窗’ /chuang1/(window), their second constituents (C2) are respectively in Tone 1/2/3/4/0. In Condition 6, the monosyllabic C1 will be produced as the control condition.

Condition 1 C1+C2(T1)	Condition 2 C1+C2(T2)	Condition 3 C1+C2(T3)	Condition 4 C1+C2(T4)	Condition 5 C1+C2(T0)	Condition 6 C1 isolated
窗花<n.>	窗前<n.>	窗口<n.>	窗外<n.>	窗户<n.>	窗<n.>
chuang1 hua 1	chuang1 qian2	Chuang1 kou3	Chuang1 wai4	Chuang 1 hu 0	Chuang1
decoration for window	in front of window	window/wicket	outside of window	window	window

5.3.2.2 Results

In this analysis, target syllables are the first constituents (C1) within compounds, after the annotation in Praat, we applied three Praat scripts to extract the acoustic measurements, including duration, F0, and intensity of the first constituent (C1). As for Condition 1-5, one-way ANOVAs were applied respectively for the acoustic measurements (duration, intensity, and F0) of the first constituents, Tone pattern (Tone 1/2/3 4/0) of the second constituent (C2) was taken as a within-

group factor, the measurement of C1 was the dependent variable, and the 20 first constituents (C1) were produced in Condition 6 as the control condition. In addition, to examine the phonetic differences between the C1 produced in compounds and in isolation, paired t-tests were conducted respectively between C1 followed by four full tones and isolated C1.

Duration

Repeated ANOVAs were conducted to analyze the duration of the first constituent (C1) under the influences of the second constituent (C2), Tone pattern of C2 (Tone 1/2/3/4/0) was a within-group factor, and the duration of C1 was the dependent variable. The mean duration of C1 was respectively 0.289s, 0.297s, 0.296s, 0.300s, and 0.295s when followed by Tone 1/2/3/4/0, and the mean duration of C1 in isolation was 0.460s. Results revealed no effect of Tone pattern on the duration of C1 [$F(4,596) = 0.443, p > .05$] (Fig.5.3.9). To analyze the anticipatory effect from the second neutral tone to the preceding full tone, paired t-tests were performed respectively between C1 followed by neutral tone and by one of the four full tones (C1+ T0 vs. C1+ T1; C1+ T0 vs. C1+ T2; C1+ T0 vs. C1+ T3; C1+T0 vs. C1+T4), no significant result was found ($p > .05$). The first constituents present no differences in the duration when followed by neutral tone or by full tones. To investigate the differences between C1 within compounds and in isolation, we added the isolated C1 (Condition 6) into ANOVAs, Post hoc tests revealed significant results respectively between C1 within compounds when followed by one of the five tones (Tone 1/2/3/4/0) and isolated C1 ($p < .01$), C1 was significantly longer when in isolation than within compounds.

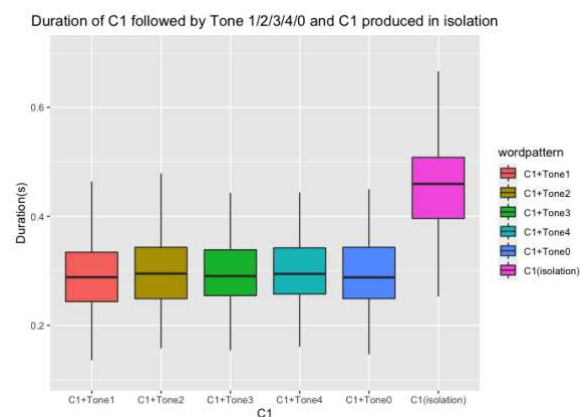


Fig.5.3.9. Duration(s) of C1 followed by C2 in neutral tone and four full tones and C1 produced in isolation.

Five colors are applied, red: Tone 1, olive: Tone 2, green: Tone 3, blue: Tone 4, and pink: control group.

Pitch F0

The pitch range and pitch value of C1 were calculated to analyze the anticipatory effect of the second constituents. Repeated ANOVAs were performed respectively for them, Tone pattern

of C2 (Tone 1/2/3/4/0) was applied as a within-group factor, and the pitch range and pitch level of C1 were dependent variables.

Firstly, results revealed a main effect of Tone pattern of C2 (Tone 1/2/3/4/0) on the pitch range of C1 [$F(4,596) = 2.844, p < .05, MSE = 1826.171, \eta^2_p = 0.02$], the mean pitch range of C1 was respectively 193.16Hz, 193.43Hz, 205.73Hz, 203.06Hz and 203.41Hz when followed by Tone 1/2/3/4/0, and the mean pitch range was 222.13 Hz when C1 was produced in isolation. Within compounds, the mean pitch range of C1 was largest when followed by Tone 3 and smallest when followed by Tone 1. Overall, the pitch range of C1 was revealed to diverge when followed by Tone 1/2/3/4/0 (Fig.5.3.10(1)).

Secondly, results reveal no main effects of the Tone pattern of C2 on the pitch level of C1 [$F(4,596) = 1.623, p > .05$], the mean pitch level of the C1 was respectively 61.75Hz, 69.85Hz, 78.43Hz, 69.53Hz and 76.45Hz when followed by Tone 1/2/3/4/0, and the mean pitch level was 123.05 Hz when C1 produced in isolation. Overall, the pitch level of C1 revealed a tendency to converge when followed by the five tones, no significance was revealed ($p < .05$) (Fig.5.3.10(2)).

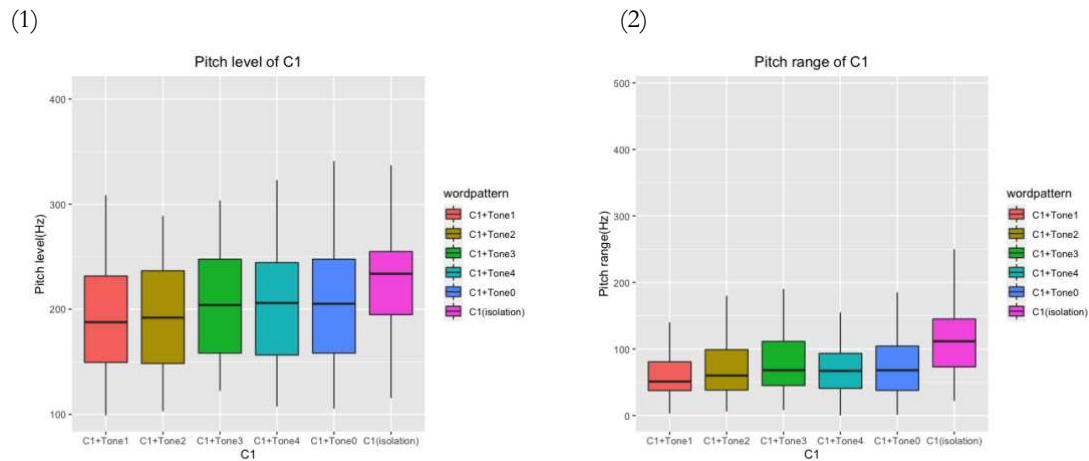


Fig.5.3.10. (1) Pitch level (Hz) of C1 followed by neutral tone and four full tones and C1 produced in isolation. (2) Pitch range (Hz) of C1 followed by neutral tone and four full tones and C1 produced in isolation. Five colors are applied, red: Tone 1, olive: Tone 2, green: Tone 3, blue: Tone 4, and pink: control group.

To present the pitch contours of compounds in the five conditions, including C1C2(T1), C1C2(T2), C1C2(T3), C1C2(T4) and C1C2(T0), we extracted the averaged F0 from ten points of the two constituents for selected compounds (Fig.5.3.11). To calculate the F0 difference between C1 offset and C2 onset, five paired t-tests were applied respectively to analyze the pitch transition between C1 and C2 in the five tonal contexts (Tone 1/2/3/4/0), and no significance was found between C1 pitch offset and C2 pitch onset, the pitch onset of the second constituent was revealed to be not significantly different from the pitch offset of the first constituent in any combination.

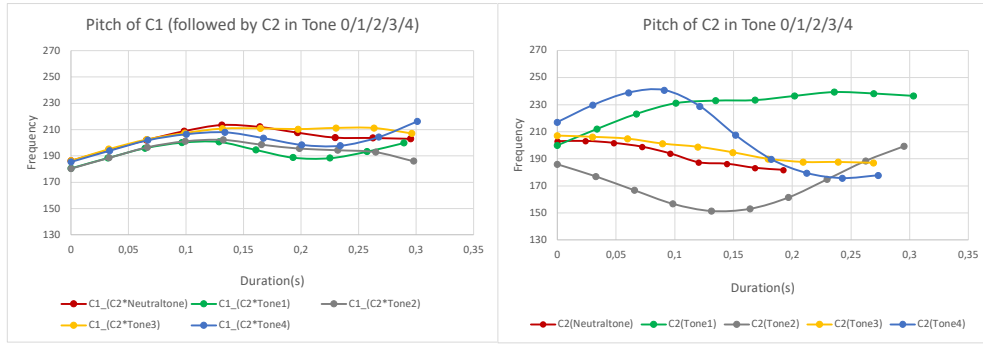


Fig.5.3.11. Left: Pitch contours of C1 followed by C2 in neutral tone (red) and C2 in the four full tones. Right: Pitch contours of C2 in neutral tone (red) and C2 in the four full tones. Five colors are applied for the five tone patterns, red: neutral tone, green: Tone 1, gray: Tone 2, yellow: Tone 3, and blue: Tone 4.

Intensity

Repeated ANOVAs were conducted to analyze the intensity of the first constituents under the influences of the following tones. Tone pattern of C2 (Tone 1/2/3/4/0) was the within-group factor, and the intensity of C1 was the dependent variable. Results revealed no main effect of Tone pattern on the intensity of C1 [$F(4,596) = 1.438, p > .05$] (Fig.5.3.12), the mean intensity of C1 was respectively 60.72dB, 61.12dB, 61.92dB, 60.34dB, 60.17 dB when followed by Tone 1/2/3/4/0, and the mean intensity of C1 in isolation was 65.83dB. On average, the intensity of C1 was highest when followed by Tone 3 and lowest when followed by Tone 2. To examine the anticipatory effect from the second neutral tone to the first full tone, paired t-tests were conducted respectively for the differences between C1 followed by a neutral tone and by four full tones (C1+ T0 vs. C1+ T1; C1+ T0 vs. C1+ T2; C1+ T0 vs. C1+ T3; C1+T0 vs. C1+T4), the significant result was revealed in the comparison between the C1 followed by Tone 0 and Tone 3 ($t = -2.21, df = 233.45, p < .05$), and the intensity of C1 was significantly higher when followed by Tone 3 than by Tone 0. And no significance was revealed in the other comparisons ($p > .05$). To check the differences between C1 in compounds and C1 in isolation, we added the isolated C1 (Condition 6) into ANOVAs, results revealed a significant difference respectively between C1 followed by any of the five tones (Tone 1/2/3/4/0) and C1 in isolation ($p < .01$), revealing that the intensity of isolatedC1 was significantly higher than C1 within compounds.

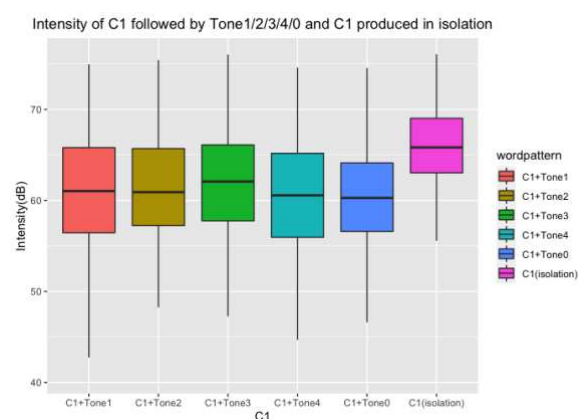


Fig.5.3.12. Intensity (dB) of C1 followed by neutral tone and four full tones and C1 produced in isolation. Five colors are applied, red: Tone 1, olive: Tone 2, green: Tone 3, blue: Tone 4, and pink: control group.

5.3.2.3 Discussion

This analysis was conducted for the anticipatory effect from the second constituent to the first constituent within compounds. Concerning the duration and intensity of the first constituents, results revealed no significant anticipatory effect from the following tones to the first constituents. Within the two-character compounds, the pitch contours of the first constituents revealed barely any significant difference when followed by a neutral tone or by any of four full tones. Concerning the anticipatory effect within the two-character compounds, we only observed the influence of the following tones on the pitch range of the first constituents.

5.3.3 Conclusion

Two acoustic analyses were performed based on the production of six native participants to explore the contextual tonal variation of neutral tone compounds. The goal was two-folded, the first was to explore whether the preceding full tones produce the carry-over effect on the second neutral tones, and the second was to analyze the anticipatory effect from the second tones to the preceding tones, especially the dissimulatory effect from the right to left direction. By applying the comparative method, our analysis examined the acoustic correlates of two neighboring tones in different combinations, including the duration, pitch (pitch range, pitch level, pitch contour), and intensity of target syllables.

The results of the first analysis indicate the carry-over effect from the leftmost full tone to the rightmost neutral tone, which is discernible in two-character neutral tone compounds. Results reveal the substantial effect from the preceding tones to neutral tones across the three categories of acoustic correlates. Within two-character compounds, we observed that the pitch offset of the first constituent was close to the pitch onset of the second constituent, and the pitch targets of the

second neutral tones were mainly affected by the preceding tones. By analyzing the pitch transition from preceding full tones to the second neutral tone, the pitch offset value of the first constituent didn't differ from the pitch onset of the second neutral tone, which is consistent with the previous study that the pitch contour of a neutral tone was associated with its preceding tone (Cao, 1986; Shen, 1992; Shi, 1988; Wang, 1996; Yip, 1980). In the second analysis, barely any anticipatory effect was revealed from the rightmost neutral tone to the preceding tone within neutral tone compound. Consistent with Xu (2005), few anticipatory effects were revealed between the two constituents of full tone compounds. More precisely, in our study, we analyzed the anticipatory effect from the rightmost neutral tone to the leftmost full tone. The phonetic features of the first full tone followed by a neutral tone revealed similar phonetic features in relation to the case when followed by full tones, neutral tone reveals barely any anticipatory effect on its preceding full tone.

In our results, a carry-over effect was revealed from the first full tone to the second neutral tone, and barely any anticipatory effect in the reverse direction was revealed within two-character neutral tone compounds. This can be accounted for by the articulatory mechanism and constraints of Mandarin neutral tone. Cheng (1973) proposed that the acoustic realization of a neutral tone is due to the loss of stress, whereas Lin (1962) and Chen (1984) have suggested that the lack of stress doesn't always trigger the atonicity of a neutral tone, and a neutral tone is not definitely weaker in the stress level. So far, most studies explained the surface pitch contour of neutral tone by the context-dependency theory, which claimed that the influence from the preceding tone determines the pitch realization of neutral tone. Sproat and Shih (1987) and Li (2003) suggested that when a neutral tone is in the medial position within a sentence or a word, its surface pitch contour can be derived from the interpolation between the preceding full tone and the post-lexical boundary tone. However, Xu (2005) and Chen and Xu (2006b) accounted for the phonological realization of the neutral tone in a different way, in their research, the PENTA model and the Target Approximation Theory (TA theory) were constructed, which proposed that each syllable possess an independent tonal target, including the neutral tone. Due to these articulatory constraints of neutral tones, such as the weaker articulatory strength and the lower articulatory speed, the underlying tonal target of neutral tones can't be fully produced in the tonal realization procedure, which led to the reduction of their vowel identity and the loss of syllable duration, that makes the neutral tone sound 'weaker' in our impression. Chen and Xu (2006b) have mentioned the greater variability contained in the surface pitch contour of the neutral tones, their pitch target was suggested to be largely affected by the surrounding tones, which can covary with adjacent full tones. In our research, the phonetic features of the second neutral tones were revealed to be coextensive with the acoustic features of the preceding full tones, few spreading influences were observed in the reverse direction. Thus,

the contextual tonal variation within neutral tone compounds could be described as in the direction from the leftmost strong full tone to the rightmost weaker neutral tone.

Chapter 6.

EEG-Experiment 1.

6 Role of frequency effect and prosody pattern in the two-character neutral tone compounds processing

6.1 Introduction

The frequency effect has been largely investigated in the compound word's recognition in Mandarin. During the lexical processing, the lack of constituent frequency effect combined with a reliable whole-word frequency effect could be viewed as evidence of the holistic processing of compound words. Some studies found that Mandarin compounds are processed as a whole unit at the recognition level (Cui et al., 2021; Hung et al., 1999; Janssen et al., 2008; Taft, 2003; Zhang & Peng, 1992). It was shown that whole-word frequency modulates the lexical decision time, which is separated from the constituent (character or morpheme) frequency. And in the other case, the morpheme frequency effect combined with whole-word frequency effect could provide evidence of the dual-route processing. For example, by varying the whole-word frequency and morpheme frequency, Zhou and Marslen-Wilson (1994) observed a robust word frequency effect in disyllabic compound representation, but the authors failed to report morpheme or syllable frequency effect. Cui et al. (2021) used eye-tracking technique and found that in Mandarin, frequent two-character compounds were processed holistically, whereas morpheme frequency affected the identification of infrequent compounds.

Psycho- and neurolinguistic evidence had brought support to the activation of individual constituents during the processing of compound words. The importance of the first constituents (leftmost) has even been revealed in right-headed languages, such as Finnish and Polish (Kehayia et al., 1999; Mäkisalo et al., 1999). Researchers have examined the morphological processing of the Mandarin disyllabic compounds by manipulating the morpheme frequency in the word recognition task. Results indicated that the frequency of individual characters plays a crucial role in the word recognition of polymorphemic compounds (Taft, 1994; Zhou et al., 1995). Morpheme frequency is important in the processing of infrequent disyllabic compounds, evidence was revealed not only for the initial morpheme (Yan et al., 2006) but also for the second morpheme (Cui et al., 2017). Zhang and Peng (1992) found that as for the Mandarin disyllabic coordinative compound words (bi-headed), e.g., 浅薄 /qian3bo2/ ('shallow-thin' → 'superficial'), the frequency effect of both two constituents could affect word processing. Regarding the whole word frequency effect in lexical recognition, previous study revealed that compared with the frequent compounds, the infrequent compounds were revealed to be processed through the morphological decomposition route. Zhou and Marslen-Wilson (1994) have found that a robust word frequency was observed in Mandarin disyllabic compound processing, frequent compounds were revealed to be identified faster than

infrequent compounds. Wang et al. (2017) examined the frequency effect in the online processing of Mandarin disyllabic compounds. Their findings were consistent with Zhou and Marslen-Wilson (1994), as they also reported the robust word frequency effect, the effect of morpheme frequency was revealed for both first constituent and second constituent. More specifically, as for frequent and infrequent compounds, results showed that the compound's recognition involved a sub-lexical decomposition into their constituent lexemes.

Auditory lexical access requires the phonologically legal match between the acoustic input information and the representation stored in the mental lexicon (Gaskell & Marslen-Wilson, 1996). When we want to study the process of multimorphemic words, like compounds, the first question is to understand whether a compound is represented as a whole unit (i.e., a unified representation) or by its constituent morphemes in the lexicon (i.e., a decomposed representation). Generally, the form priming paradigm can be used to study this procedure, in which a prime word is followed by the same word serving as a target, like book^{prime} – BOOK^{target}, and results in the repeated condition will be compared with the results in the non-repeated condition, e.g., table – BOOK. It was used as a useful paradigm to examine the nature of lexical representation as well as their access (Marslen-Wilson & Zwitserlood, 1989; Vroomen & De Gelder, 1997; Zwitserlood, 1996). Frauenfelder and Peeters (1990) used the form-priming in their experiment, the primes and the targets were related in their forms (ex. 'car' - 'carpet') in the lexical decision task. In this case, these authors found a facilitatory priming effect reflected by a shorter lexical decision time in the repeated condition in comparison with non-repeated condition. The observation of the facilitation of /car/ in /carpet/ suggests that embedded word (we note that the embedded segment doesn't have morphological relation with carrier word /carpet/) was transitorily activated. In addition, by using a cross-modal priming paradigm, Shillcock (1990) found a facilitatory priming effect on the visual target 'RIB', when auditorily primed by the semantically related word 'trombone' ('rib' is primed by 'bone').

In Mandarin, previous electroencephalographic (EEG) studies examined the incongruent metrical pattern in the lexical processing, such as in the incongruent pitch accents and lexical tones (Li et al., 2008), the unnatural rhythmic patterns of compound words (Luo & Zhou, 2010) and the irregular syllabic duration (Wang et al., 2005). For example, Wang et al. (2005) investigated whether the prosodic salience (e.g., stressed syllables within disyllabic sounds) could serve as cues to capture listener's attention. The disyllabic standard (prosodic salience in different positions, **B**Ag**a**, **b**aG**A**) and deviant input (stressed initial: **P**Ag**a**, unstressed-noninitial: **B**Ak**a**, stressed-noninitial: **b**aK**A**, unstressed-initial: **p**aG**A**), which differed in their prosodic status (voiced consonants vs. unvoiced consonants), were presented to the participants by using the auditory oddball paradigm. Results showed that prosodic salient deviance capture attention in the unattended sounds. The MMN was

elicited by the deviants in the four conditions, which revealed the discrimination between standard and deviant input. The P3a was typically related to the orientation of involuntary attention in the listening, and P3a was observed when deviance is on a stressed syllable (e.g., *PAga* and *baKA*). Stressed syllables (prosodic salience) can capture the listeners' attention, which works as attentional cue to facilitate speech processing. Luo and Zhou (2010) examined to what extent the prosodic constraints in Mandarin play a role in the lexical processing, the number of syllables and semantic congruency of verbs in the compounds were manipulated, e.g., the verb-noun combination 种蒜 (plant + garlic) is a normal [1+1] rhythmic pattern, compared to 种植蒜 (plant + garlic), an abnormal [2+1] rhythmic pattern. The experimental manipulations yielded an increased N400-like effect in the incongruent condition compared to the congruent ones, and N400 effect suggests the higher cognitive cost in the representation of incongruent prosody structures, and incongruent prosody patterns can hinder the lexical /semantic representation in online processing.

Moreover, some studies revealed that the additional cognitive resources might be required in the processing of neutral tone syllables in the auditory modality. E.g., in contrary to the full tone compound, a more negative amplitude of the N400 component in the neutral tone compound's processing was observed by Luo et al. (2016). Consistent findings were found by Yan et al. (2014) in oculomotor research, the readers processed the upcoming parafoveal word less effectively when neutral tone compound was fixated than full tone compound. Isel and Shen (2020) have found the perception of neutral tones by native Mandarin listeners, by using the auditory oddball paradigm, standard stimuli (full tone compounds, prosody: Strong-Strong) and deviant stimuli (neutral tone compounds, prosody: strong-weak) were presented auditorily to the native listeners. Researchers found that the neutral tones can serve as cues to capture listener's attention in speech processing. Concerning the trochaic rhythmic pattern (the first strong full tone followed by the second weak neutral tone) carried by disyllabic neutral tone compounds, depending on the previous research, neutral tones could consume more resources in the recognition, which might assist the activation of the individual constituent in the processing of two-character compounds.

6.2 Goal of research and hypotheses

Two-character neutral tone compounds consist of a full tone syllable at the first (leftmost) position and a weak neutral tone syllable at the second (rightmost) position, resulting in a trochaic prosody pattern [Strong-weak]. E.g., in the compound 下场 /xia4chang0/ (bad end), the first constituent 'xia4' (descend) is in Tone 4, the second constituent 'chang0' (field) is in a neutral tone, it is a non-transparent two-character compound with a prosody pattern [Strong-weak]. Because in

Mandarin, the neutral tone constituent within a compound doesn't exist alone in the speech, it occurs in the non-initial position of compounds, the lexical representation of a second neutral tone constituent within a neutral tone compound cannot be realized separately in auditory modality, its semantic activation must be achieved through the combination with the first full tone constituent. For example, listeners cannot activate the lexical representation of the syllable 'chang (T0)' in the listening, its semantic processing must be achieved by combining it with the first tone 'xia (T4)' in the phonological context. In the previous studies, only a handful of works have focused on the acoustic-syntactic decomposition of two-character neutral tone compounds, less is known about its lexical processing/representation mechanism. As raised in the literature, two questions will be investigated in this study. The first question is whether the two-character neutral tone compounds would be recognized as a single unit, via its individual constituents, or by a combination of the two processing ways, and whether the first constituent would be decomposed and activated in auditory lexical access. The second question is about the role of prosody pattern in lexical representation, and whether lexical processing can be modulated by the second neutral tone within two-character compounds. In other words, we want to study whether the neutral tone carried by the rightmost character in the neutral tone compounds works as a decomposition trigger in auditory processing. In addition, our goal was also to investigate whether the neutral tone could interact with the lexical frequency in compound word processing.

Because the electrical responses measured by event-related potential (ERP) at the surface of the scalp are considered to provide the specific information of temporal dynamics in the lexical recognition, in the current study, continuous EEG signals were recorded when participants were performing the lexical decision task on Mandarin characters and pseudo-characters (equivalent to word/pseudoword). We used the cross-modal (auditory-visual) repetition priming paradigm, with a decision task on the visual targets. In each trial, participants firstly listened to an auditory prime (a spoken two-character Mandarin compound, e.g., 巴掌/ba1zhang0/ hand), which was followed by a visual target, i.e., a Mandarin character or pseudo-character, and the participant's task was to decide whether the visual target presenting at the center of the screen was a correct character or not. Three variables were manipulated to the auditory primes (spoken compounds), including the frequency of the whole compound, the frequency of first constituent (C1), and the prosody pattern of compounds. Therefore, a 2x2x2 experimental design is constructed, three variables are involved, and each variable consists of two levels, including the **frequency of the whole-compound** (two levels: frequent, rare), the **frequency of the first constituent** (two levels: high-frequency, low-frequency), and the **prosody pattern of compounds** (two levels: [Strong-strong], [Strong-weak]).

In this study, we want to analyze the participants' behavioral results, and the ERP results which correspond to the segments between the onset and offset of the visual target's presentation. Because N400 is an ERP component that is typically related to the semantic integration in both the lexical representation (Brown & Hagoort, 1993) and lexical access (Van Petten & Kutas, 1990), thus the mean amplitude of the N400 effect will be the main indicator to be analyzed in the present study. The rationale of our experiment is that if the frequency of whole-compound and the neutral tone on the second constituent would work as a determining cue in morphological decomposition, the rare neutral tone compounds should be decomposed via the individual constituent in auditory processing, and this decomposition should be revealed by the repetition priming effect on the first characters. Moreover, if the compounds were decomposed in the processing, we should observe that the first constituent's frequency works with whole-compound's frequency and modulate the priming effects. Here, considering the three variables manipulated in this experiment, we want to investigate their cumulative effects by measuring the priming effects which should be elicited by one or more than one variable. For example, we want to compare these two priming effects, when their auditory primes were respectively the rare full tone compounds containing the high-frequency first constituents or the rare full tone compounds containing the low-frequency first constituents, through this comparison, we assumed that a smaller priming effect should be observed when the auditory primes were the rare full tone compounds containing the low-frequency first constituents, and this result should be represented by an increased N400 effect in the ERP results.

We predicted that a priming effect of the first constituent within two-character compounds would be observed in the behavioral result of the cross-modal priming experiment. Moreover, the priming effect is assumed to be modulated by the whole-compound frequency and prosody pattern of the spoken compounds (auditory prime), which should be smaller for the rare compounds than the frequent ones, and this should be revealed by the increased N400 effect. We assumed that the neutral tones contained in the trochaic prosody pattern of two-character neutral tone compounds should modulate the priming effect of the first character, and this effect should be smaller for the neutral tone compounds than the full tone compounds, which should be reflected by a significantly larger N400 effect. Furthermore, the frequency of first constituent should be involved in auditory processing of two-character compounds, we predicted that the frequency of the first constituents should be observed to modulate the repetition priming effects of the first constituents (characters).

6.3 Methodology

Participant 30 native speakers of Mandarin joined the experiment (18 females and 12 males; age: 22 to 35, mean: 28.28 ± 2.36), they were born and grew up in the China mainland, and

speak Mandarin as their mother language. And they are college students in the universities of Paris when participated in the experiment. None of them have lived in France for more than three years (0.5 to 3 years, mean 2.15 ± 0.8), and they speak Mandarin frequently in daily life (2 to 5 hours per day, mean 3.85 ± 0.7). All the participants are right-handed with the normal or corrected-to-normal vision without a history of neurological, psychiatric, or auditory impairment, and they were paid 15 euros for their participation. Participants were clearly instructed on the nature of the tasks, they gave their informed content before the experiments by completing an information sheet. All the research protocol was approved by the local ethical standards of Paris Nanterre University.

Linguistic material The minimal pairs of two-character Mandarin compounds constituted the experimental materials, each minimal pair consists of a two-character neutral tone compound [Strong-weak] and a two-character full tone compound [Strong-strong]. Their second constituents (C2) possess the identical syllabic structure but carry either a neutral tone or a full tone. E.g., 合伙/he2 huo3/, ‘collaborate’ and 家伙 /jia1 huo0/ ‘fellow’. In this experiment, we applied the auditory-visual cross-modal repetition priming paradigm with a lexical decision task on the visual target. Concretely, participants were firstly auditorily presented with a prime (a spoken compound), then they were visually presented with a target (a Chinese character), and the visual target would either repeat (priming condition) or not (non-priming condition) the first constituent of auditory prime. Firstly, 160 two-character neutral tone compounds [Strong-weak] were selected, which have a first full tone constituent and a second neutral tone constituent, and we didn’t choose the neutral tone compounds which have lexical suffixes in the second position (e.g., 梳子 /shu1zi0/, comb). Then, as for all the 160 neutral tone compounds, each was paired with one two-character full tone compound, in each minimal pair, their second constituents possess the same syllable structure but either in a full tone or in a neutral tone. To select the two-character compounds (auditory primes), three variables were manipulated, including the **whole-compound frequency** (frequent, rare), the **first constituent frequency** (high-frequency, low-frequency), and the **prosody pattern** (neutral tone compound [Strong-weak], full tone compound [Strong-strong]). This 2x2x2 experimentation design yields eight conditions (Table 1): (1) **SS & F & H-C1**, frequent full tone compound [Strong-strong] with a high-frequency first constituent; (2) **SS & F & L-C1**, frequent full tone compound [Strong-strong] with a low-frequency first constituent; (3) **SS & R & H-C1**, rare full tone compound [Strong-strong] with a high-frequency first constituent; (4) **SS & R & L-C1**, rare full tone compound [Strong-strong] with a low-frequency first constituent; (5) **SW & F & H-C1**, frequent neutral tone compound with a high-frequency first constituent [Strong-weak]; (6) **SW & F & L-C1**, frequent neutral tone compound [Strong-weak] with a low-frequency first constituent; (7) **SW & R & H-C1**, rare neutral tone compound

[Strong-weak] with a high-frequency first constituent; (8) **SW & R & L-C1**, rare neutral tone compound [Strong-weak] with a low-frequency first constituent. In each of the eight conditions, 20 pairs of compounds (20 full tone compounds and 20 neutral tone compounds) were selected, thus 320 two-character compounds were selected (Table.6.1). Secondly, to construct the cross-modal priming experiment (‘auditory prime – visual target’), 320 characters were selected as visual targets to be paired with primes. 40 pairs of ‘prime-target’ existed in each of the eight conditions, among which 10 pairs in the **repeated group**, 10 pairs in the **non-repeated group**, and 20 pairs in the **pseudo-character group** (Fig.6.1). All the 320 characters were hence attributed into one of the three groups: 1). **Repeated**: 80 characters are the repetition of the first constituent of auditory prime, 2). **Non-repeated**: 80 characters are not the repetition of the first constituent of auditory prime, 3). **Pseudo-character**: 160 pseudo-characters. Finally, as to manipulate the proportion of the **Repeated group** into 20%, we added another 10 pairs of ‘auditory prime & visual target’ in each of the eight conditions, these 80 pairs were added as the fillers into the **Pseudo-character** condition, selected compounds in the filler pairs were matched with the frequency. In total, 400 two-character compounds were selected as the auditory primes, which were matched with their lexical frequency (mean frequency of selected compounds is listed in Table.6.2), and 400 Chinese characters were selected as visual targets, which were matched with their stroke numbers (mean 17.59, \pm 8.59) (see Appendix 3-1/3-2/3-3).

Table 6.1. Example of selected auditory primes in the eight conditions. The second constituents in each minimal pair of compounds have the same syllabic structure with either neutral tone or full tone, i.e., compound 合伙 he(T2) huo (T3) (collaborate) is in the Condition 1: SS & F & H-1; and 家伙 jia(T1) huo(T0) (fellow) is in the Condition 5: SW & F & H-1, their second constituents have the same syllable ‘huo’ in either Tone 3 or Tone 0.

Condition	Prosody pattern	Compound frequency	First constituent frequency	Example compound	Prime-Target pair		
					Repetition of C1		Pseudo-character
					Yes	no	
1. SS & F & H-1	[Strong-strong]	Frequent	High frequency	合伙 collaborer [heT2 huoT3]	10	10	20
2. SS & F & L-1			Low frequency	估量 estimator [guT1 liangT0]	10	10	20
3. SS & R & H-1		rare	High frequency	义务 obligation [teT4 wuT0]	10	10	20
4. SS & R & L-1			Low frequency	泡沫 écume [paoT4 moT4]	10	10	20
5. SW & F & H-1	[Strong-weak]	Frequent	High frequency	家伙 camarade [jiaT1 huoT0]	10	10	20

6. SW & F & L-1			Low frequency	掂量 soupeser [dianT1liangT0]	10	10	20
7. SW & R & H-1		rare	High frequency	特务 espion [teT4 wuT0]	10	10	20
8. SW & R & L-1			Low frequency	唾沫 crachat [tuoT4 moT0]	10	10	20

	Neutral tone compound [Strong-Weak]	Full tone compound [Strong-Strong]	
1.Related group	compound – character (C1) Ex. 客气 ke4qi0 – 客	compound – character (C1) Ex. 脾气 pi4qi4 – 脾	80 pairs
2.Non-related group	compound - character (not C1) Ex. 巴掌 ba1zhang0 – 客	compound - character (not C1) Ex. 手掌 shou34zhang0 – 脾	80 pairs
3.Pseudo-character group	compound – pseudo-character Ex. 含糊 han2hu0 – X	compound – pseudo-character Ex. 黏糊 nian2hu0 – X	160 pairs
4.Fillers pairs	compound – pseudo-character Ex. 机灵 ji1ling0 – X	compound – pseudo-character Ex. 精灵 jing1ling0 – X	80 pairs

Fig.6.1. Illustration for the four types of ‘auditory prime- visual target’ pairs. X is for the pseudo-character.

400 compounds were selected from the contemporary Chinese dictionary 5th edition (Lu & Ding, 2008), their lexical frequency was calculated by using RStudio based on the Corpus of contemporary Mandarin (Corpus with word class labeling) by China State language commission (<http://corpus.zhonghuayuwen.org/Resources.aspx>).² As for the lexical properties which might interact with lexical frequency to affect word recognition, such as semantic transparency and word concreteness, a lexical pretest was made for the selected compounds to avoid the influences of the two lexical properties (semantic transparency and word concreteness) on the lexical decision of participants. 16 native speakers of Mandarin who were students at the Universities of Paris were recruited for the pretest, and they did not join the EEG experiments. These 16 students rated the semantic transparency and the concreteness of selected compounds. Semantic relativeness of both two constituents is rated respectively with the whole compound, and the two rates were averaged to yield the semantic transparency of compounds, a 5-point Likert scale was applied in the test (1 for barely related and 5 for highly related). And word concreteness was also rated for each selected compound, and the 5-point Likert scale was used (1 for barely concrete and 5 for highly concrete). Table.6.2 presents the information of selected compounds in eight conditions, including the mean occurrence times of selected compounds in the corpus, the evaluation of semantic transparency and word concreteness for selected materials.

² Corpus of contemporary Mandarin is an on-line word-text corpus, which comprises commonly used written and spoken Chinese from a wide range of sources, including approximately 10 million words (20 million in characters).

Table 6.2. Information for selected compounds in the eight conditions. Compound frequency and first constituent frequency are calculated by counting the occurrence times per 10 million in the corpus. And the semantic transparency and word concreteness were evaluated by a 5-point scale (1 for barely related and 5 for highly related)

Condition	Compound Frequency	First constituent frequency	Semantic transparency	Word concreteness
1. SS & F & H-C1	1511,65	5048	2.89	3.08
2. SS & F & L-C1	521,05	91,3	2.9	2.97
3. SS & R & H-C1	97,55	3428,4	3.02	2.8
4. SS & R & L-C1	83,75	118,65	3.18	3.14
5. SW & F & H-C1	1157,7	6612,15	3.17	3.2
6. SW & F & L-C1	721,8	129,76	2.8	2.86
7. SW & R & H-C1	85,76	4769,85	2.98	3.1
8. SW & R & L-c1	96,84	97,73	3.1	3.05

6.4 ERP recordings and analyses

Experimental procedure The prime compounds were recorded by a native Beijing female speaker in a soundproof room using a digital audiotape (sampling at 44.1kHz). In this experiment, the cross-modal repetition priming paradigm was combined with a decision task, the decision task was performed by 30 participants, they listened firstly to an auditory prime (a spoken compound) and determined whether a visual target (a Chinese character) is a real/correct character or not. In each trial, a fixation cross was firstly presented in the middle of the screen for 1000ms, and a blank screen was presented for 500ms; Then an image of a ‘loudspeaker’ in the middle of the screen was presented for 1000ms with the auditory prime words playing at the same time in their earphones. After the auditory prime, a blank screen was presented for 500ms, and the visual target (a Chinese character) was presented on the screen for 500ms, then a “?” appeared directly in the middle of the next screen to cue the participants to answer the question, if the character on the screen was a real character or not, participants were asked to give the answer by pressing the ‘yes’ or ‘no’ button. When the character was a correct Chinese character, they were asked to press ‘yes’; when it was a pseudo-character, they must press ‘no’, and the “?” disappeared when the participants finished the answer. The intertrial interval (ITI) was 1500ms, the SOA was 1500ms, and the ISI was 500ms. 20 practice trials proceeded before the formal tests (Fig.6.2).

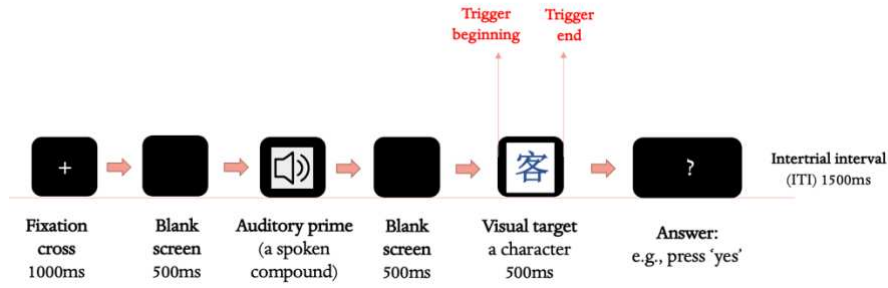


Fig.6.2. Illustration of the trial scheme. The auditory prime is neutral tone compound ‘客气, /ke4qi0/, courteous’, the visual target is the first character of the auditory prime ‘客’, and participants will press ‘yes’ to answer when ‘?’ appears on the computer screen.

EEG recordings and preprocessing The EEG signals were recorded from the surface of scalp by using the 64-channel Biosemi ActiveTwo AD box with a sampling rate of 512 Hz, and an online bandpass filter was applied with the lowpass cutoff set at 100Hz and the highpass cutoff at 0.05Hz. To measure eye movements from two eyes, 4 electrodes were used for electrooculogram (EOG), and the measurements were obtained from below versus above the left eye (vertical EOG) and the left versus right lateral orbital rim (horizontal EOG). Moreover, two additional electrodes were placed over the right and left mastoids. After the recordings, we used the Python-MNE for the preprocessing of EEG data. Firstly, through a visual inspection, bad channels which contained EMG, or the other artifacts (not eye movements or blinks) were manually removed. We used the algebraic average of left and right mastoids (M1 & M2) as a reference to perform the interpolation of bad channels, then each electrode was re-referenced offline to the algebraic average of the left and right mastoids (M1 & M2). Secondly, all the remaining data (except M1 & M2) were offline bandpass-filtered from 0.05 to 30 Hz. Then, the data were corrected for the eye-blink artifacts by applying the ICA to semi-automatically identify and subtract the oculomotor artifacts. Finally, the cleaned EEG data were segmented into the epochs spanning -200ms to 500 ms corresponding to the onset and offset of visual target, finally, a visual inspection was realized to reject manually the epochs which contained the excessive artifacts. For all the remaining epochs, a baseline correction was applied to the pre-stimulus time window from -200ms to 0ms. To analyze the EEG results, we only selected the epochs in the ‘Repetition’ and ‘Non-repetition’ groups, and the epochs in the ‘Pseudo-character’ group and fillers groups were excluded. All the selected epochs were averaged within and across participants to obtain the grand averages in the eight conditions.

ERP data analyses Preprocessed ERP data were then analyzed with Repeated ANOVAs analysis in the RStudio. We extracted the mean amplitude in the time windows of 350-500ms for the N400 component, and the mean amplitudes were extracted from 9 fronto-central and centro-parietal electrodes (F3, Fz, F4/C3, Cz, C4/P3, Pz, P4). Firstly, the amplitudes at the three midline

electrodes (Fz/Cz/Pz) were entered the ANOVAs tests. Secondly, in the lateral sites, we selected three regions of interest (anterior, central, and posterior), and six lateral electrodes were selected for the three regions (anterior: F3, F4; central: C3, C4; posterior: P3, P4), the mean amplitudes of lateral electrodes were entered the ANOVAs by measuring the mean amplitude of three regions of interests, including the average of F3 and F4 for anterior region, the average of C3 and C4 for central region, and the average of P3 and P4 for posterior region. To adjust for the violation of sphericity (Greenhouse & Geisser, 1959), the Greenhouse-Geisser correction was applied when there was more than one degree of freedom in the numerator. To examine the influences of the three variables on the priming effects of visual targets, three groups of four-way ANOVAs were created, each group consisted of two four-way ANOVAs respectively for the midline and lateral electrodes. In each four-way ANOVAs, when a significant main effect or interaction was observed, Post-hoc Tukey tests were performed to interpret it. Results of the three groups of ANOVAs are presented respectively in Table 6.3, Table 6.4, and Table. 6.5.

ANOVAs-1: Compound frequency * Prosody pattern * Repetition * Electrodes

ANOVAs1-1: Compound frequency (Frequent & Rare) * Prosody pattern (Strong-Strong & Strong-Weak) * Repetition (Repeated & Non-repeated) * Midline electrodes (Fz/Cz/Pz)

ANOVAs1-2: Compound frequency (Frequent & Rare) * Prosody pattern (Strong-Strong & Strong-Weak) * Repetition (Repeated & Non-repeated) * Lateral electrodes (3 regions: anterior: F3&F4; central: C3&C4; posterior: P3&P4).

ANOVAs-2: Compound frequency * C1 Frequency * Repetition * Electrodes

ANOVAs2-1: Compound frequency (Frequent & Rare) * C1 Frequency (high-frequency & low-frequency) * Repetition (Repeated & Non-repeated) * Midline electrodes (Fz/Cz/Pz)

ANOVAs2-2: Compound frequency (Frequent & Rare) * C1 Frequency (high-frequency & low-frequency) * Repetition (Repeated & Non-repeated) * Lateral electrodes (3 regions: anterior: F3&F4; central: C3&C4; posterior: P3&P4).

ANOVAs-3: C1 Frequency * Prosody pattern * Repetition * Electrodes

ANOVAs3-1: C1 Frequency (high-frequency & low-frequency) * Prosody pattern (Strong-Strong & Strong-Weak) * Repetition (Repeated & Non-repeated) * Midline electrodes (Fz/Cz/Pz)

ANOVAs3-2: C1 Frequency (high-frequency & low-frequency) * Prosody pattern (Strong-Strong & Strong-Weak) * Repetition (Repeated & Non-repeated) * Lateral electrodes (3 regions: anterior: F3&F4; central: C3&C4; posterior: P3&P4).

6.5 Results

Behavioral results We calculated the reaction time for the Repeated (priming condition) and the Non-repeated (non-priming condition) groups, and the RTs of ‘Pseudo-character’ group and fillers group were excluded in the analysis of behavioral results. Overall, the accuracy rate was 98.75%, more specifically, the accuracy rate in the priming condition and non-priming condition was 99.02% and 98.48%, respectively. A paired t-test for the accuracy rates between the priming condition and non-priming condition was realized, and no significant result was found ($p > .05$). Reactions times which were greater than two SDs beyond the global mean were discarded (10.8%), a paired t-test for RTs in the priming condition and non-priming condition revealed a significance ($t = -2.3376$, $df = 393.39$, $p = 0.01991$, $p < .05$), indicating a significantly shorter decision time in the priming condition (mean 0.28 ± 0.19) than in the non-priming condition (mean: 0.33 ± 0.26). Moreover, a four-way repeated ANOVAs was performed for the reaction times in the two priming conditions among the three variables (Prosody pattern, Compound frequency, and C1 frequency), results revealed that neither main effect nor interaction effect reached the significance across the three variables. Fig.6.3 presents the boxplot of the RTs in the two priming conditions.

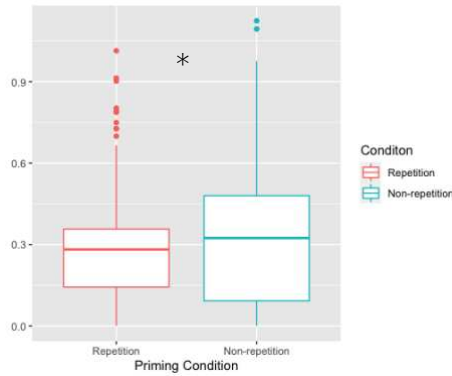


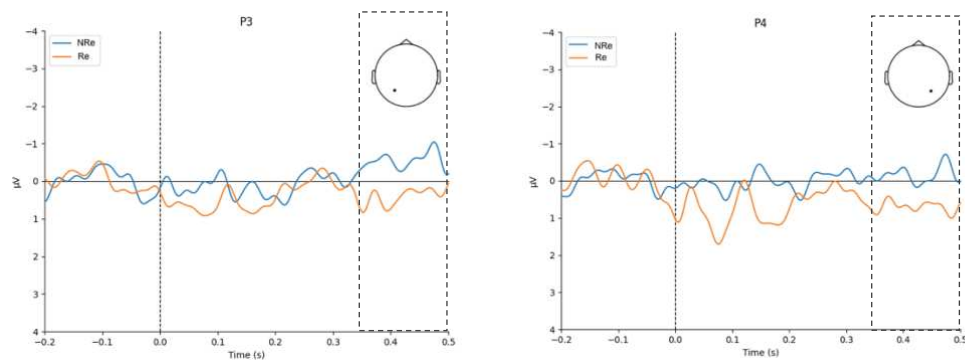
Fig.6.3. Reaction times in two priming conditions, red is C1 repeated, green is C1 non-repeated ($p < .05$).

ERP results The first group of four-way ANOVAs was conducted: **Compound frequency** [Frequent & Rare] * **Prosody pattern** [Strong-strong & Strong-weak] * **Repetition** [Repeated & Non-repeated] * **Electrodes** (midline & lateral) (Table.6.3). At the lateral electrodes, the significant interaction was revealed between the **Repetition** and **Electrodes** ($F(2,72) = 7.454$, $p = 0.001$, $MSE = 0.380$, $\eta^2p = 0.13$), revealing that the difference of N400's mean amplitudes between the priming and non-priming condition was significantly larger in the posterior region than in the central region ($p < .05$) (Fig.6.4).

Table 6.3. Results of four-way ANOVAs-1 computed on midline and lateral electrodes.
Variables: Compound frequency * Prosody pattern * Repetition * Electrodes

Factors	df	Site	Time window N400		Site	Time window N400	
		Midline	F	P	Lateral	F	P
Compound frequency	1	(Fz, Cz, Pz)	0.116	0.736	(F3/F4, C3/C4, P3/P4)	3.116	0.081
Prosody	1		0.059	0.810		0.286	0.594
Repetition	1		0.906	0.350		0.260	0.611
Electrodes	2		4.894	0.016*		4.183	0.019*
Compound frequency *Prosody	1		0.041	0.841		0.964	0.329
Compound frequency*Repetition	1		3.206	0.086		0.297	0.587
Prosody*Repetition	1		0.403	0.531		0.002	0.962
Compound frequency*Electrodes	2		0.835	0.446		1.783	0.175
Prosody*Electrodes	2		0.399	0.675		0.782	0.461
Repetition*Electrodes	2		1.772	0.191		7.454	0.001**
Compound frequency*Prosody *Repetition	1		0.153	0.699		0.197	0.658
Compound frequency*Prosody *Electrodes	2		0.170	0.844		1.744	0.182
Compound frequency*Repetition *Electrodes	2		0.828	0.449		1.782	0.175
Prosody*Repetition*Electrodes	2		0.398	0.676		0.361	0.698
Compound frequency*Prosody *Repetition*Electrodes	2		0.097	0.907		0.045	0.956
Residuals	72						

(a)



(b)

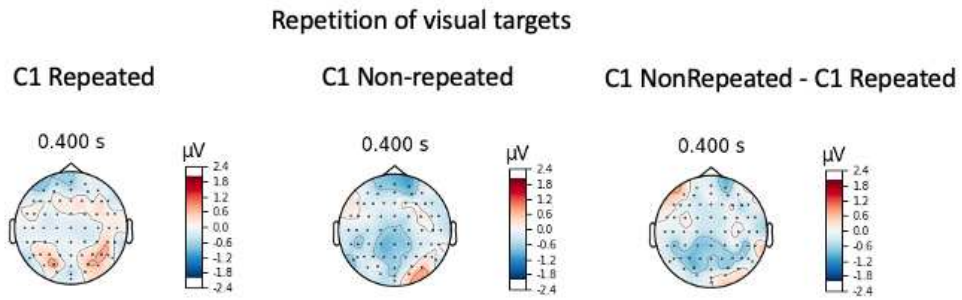


Fig.6.4. (a) Grand averages of ERPs were measured at the P3 and P4 electrodes and time-locked to the onset of visual targets in the two priming conditions (blue: non-repeated; orange: repeated). The time window of N400 is indicated by dashed lines. (b). Topographic maps for the significant priming effect in the timing of 400ms post-stimuli, $p < .05$.

The second group of four-way ANOVAs was performed: **Compound frequency** [Frequent & Rare] * **C1 Frequency** [High-frequency & Low-frequency] * **Repetition** [Repeated & Non-repeated] * **Electrode** (midline & lateral) (Table.6.4). Results of ANOVAs revealed the significant interaction between **Repetition** and **Electrode**. At the midline electrodes, a significant interaction was revealed among **C1 Frequency**, **Repetition** and **Electrodes** ($F(2,72) = 2.746$, $p = 0.044$, $MSE = 0.15$, $\eta^2p = 0.08$). Post hoc tests revealed that when auditory primes were compounds containing low-frequency first constituents, the significantly more negative amplitude of N400 was elicited by the non-priming condition in relation to the priming condition at the central electrodes ($p < .05$). Moreover, an interaction was revealed among the **Compound frequency** and **Repetition** ($F(1,72) = 3.248$, $p = 0.0207$, $MSE = 0.267$, $\eta^2p = 0.07$), when auditory primes were rare compounds, the significantly more negative amplitude of N400 was elicited by non-priming condition in relation to priming condition ($p = 0.05$). And interaction between **C1 Frequency** and **Repetition** was found in the ANOVAs ($F(1,72) = 4.071$, $p = 0.0117$, $MSE = 0.259$, $\eta^2p = 0.08$), when auditory primes have low-frequency first constituents, a significantly more negative amplitude of N400 was elicited by the non-priming condition in relation to the priming condition ($p < .05$) (Fig.6.5 and Fig.6.6).

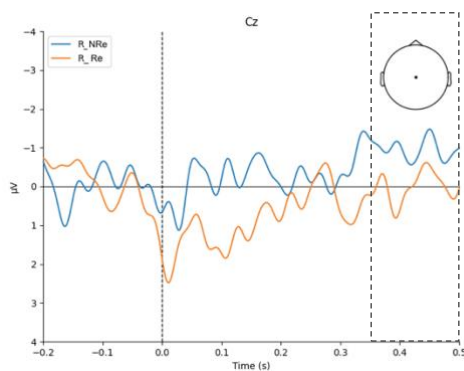
Table 6.4. Results of four-way ANOVAs-2 computed on midline and lateral electrodes.

Variables: Compound frequency * C1 Frequency * Repetition * Electrodes

Factors	df	Site	Time window		Site	Time window	
		Midline	N400		Lateral	N400	
			F	p		F	p
Compound frequency	1	(Fz, Cz, Pz)	0.221	0.642	(F3/F4, C3/C4, P3/P4)	3.174	0.079
C1 Frequency	1		1.430	0.243		0.177	0.675

Repetition	1	1.732	0.200	0.265	0.608
Electrodes	2	9.351	<0.001***	4.260	0.017*
Compound frequency*C1Frequency	1	0.335	0.567	0.375	0.542
Compound frequency*Repetition	1	6.127	0.0207*	0.302	0.584
C1 Frequency *Repetition	1	7.446	0.0117*	0.213	0.645
Compound frequency *Electrodes	2	1.595	0.223	1.815	0.170
C1Frequency *Electrodes	2	0.566	0.573	2.509	0.088
Repetition*Electrodes	2	3.387	0.050	7.591	0.001**
Compound frequency*C1Frequency *Repetition	1	0.669	0.421	0.001	0.979
Compound frequency*C1Frequency *Electrodes	2	2.008	0.156	1.413	0.250
Compound frequency*Repetition *Electrodes	2	1.581	0.226	1.814	0.170
C1Frequency*Repetition*Electrodes	2	3.565	0.044*	0.074	0.928
Compound frequency*C1Frequency *Repetition*Electrodes	2	2.511	0.102	0.009	0.990
Residuals	72				

(a)



(b)

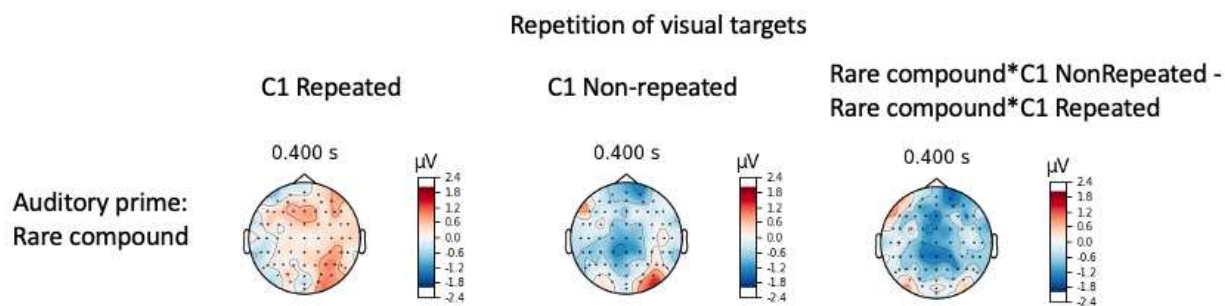
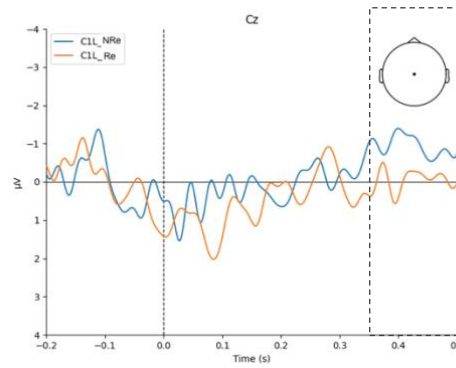


Fig.6.5. (a) Grand averages of ERPs were measured at the Cz electrode and time-locked to the onset of visual targets in the two priming conditions, when auditorily preceded by rare compounds (blue: non-repeated; orange: repeated).. The time window of N400 is indicated by dashed lines. (b) Topographic maps of compound frequency effect in the timing of 400ms post-stimuli, $p < .05$.

(a)



(b)

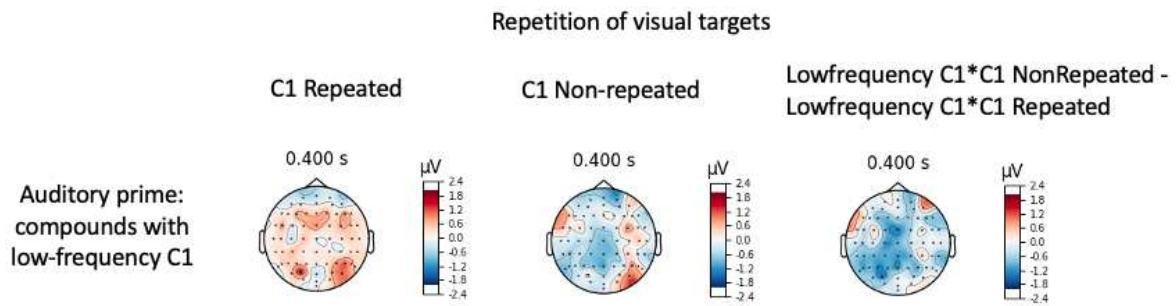


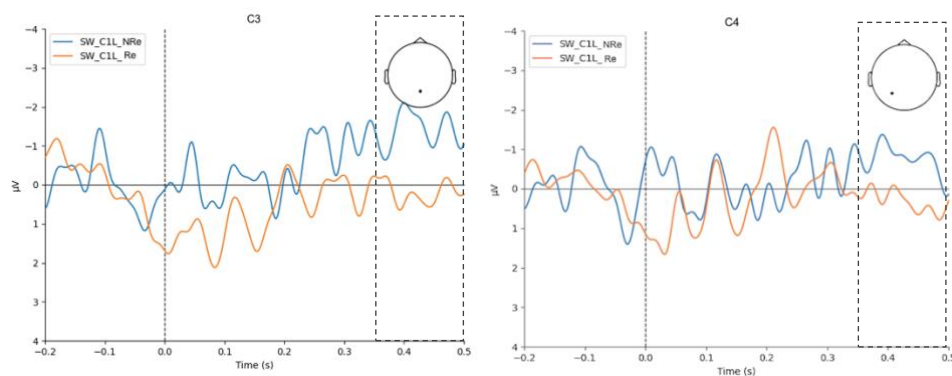
Fig.6.6. (a) Grand averages of ERPs were measured at the Cz electrode and time-locked to the onset of visual targets in the two priming conditions, when auditorily preceded by compounds with low-frequency C1 (blue: non-repeated; orange: repeated).. The time window of N400 is indicated by dashed lines. (b) Topographic maps of the first constituent frequency effect in the timing of 400ms post-stimuli, $p < .05$.

The third group of four-way ANOVAs was created: **Compound frequency** [Frequent & Rare] * **C1 Frequency** [high-frequency & low-frequency] * **Repetition** [Repeated & Non-repeated] * **Electrodes** (midline & lateral) (Table.6.5). In addition to the second four-way ANOVAs, at the lateral electrodes, the interactions between **Repetition** and **Electrodes** and between **C1 Frequency** and **Repetition** were significant. Moreover, at the lateral electrodes, Repeated ANOVAs revealed an interaction among the **Prosody pattern**, **C1 Frequency** and **Repetition** ($F(1,72) = 4.311$, $p = 0.041$, $MSE = 0.408$, $\eta^2_p = 0.01$), when auditory primes were neutral tone compounds with low-frequency first constituents, a significantly more negative amplitude of N400 was elicited by the non-priming condition in relation to the priming condition ($p < .05$) (Fig.6.7).

Table 6.5. Results of four-way ANOVAs-3 computed on midline and lateral electrodes
Variables: Prosody pattern * C1 Frequency * Repetition * Electrodes

Factors	df	Site	Time window		Site	Time window	
		midline	N400		Lateral	N400	
		(Fz, Cz, Pz)	F	P	(F3/F4, C3/C4, P3/P4)	F	P
Prosody	1		0.096	0.760		0.289	0.592
C1Frequency	1		1.210	0.282		0.175	0.677
Repetition	1		0.146	0.238		0.262	0.610
Electrodes	2		7.909	0.002**		4.216	0.018*
Prosody*C1Frequency	1		2.619	0.118		0.045	0.833
Prosody*Repetition	1		0.652	0.427		0.002	0.962
C1Frequency*Repetition	1		6.298	0.019*		0.211	0.647
Prosody*Electrodes	2		0.645	0.533		0.788	0.458
C1Frequency*Electrodes	2		0.479	0.625		2.484	0.090
Repetition*Electrodes	2		2.865	0.076		7.513	0.001**
Prosody*C1Frequency*Repetition	1		0.127	0.725		4.311	0.041*
Prosody*C1Frequency*Electrodes	2		1.355	0.277		0.761	0.470
Prosody*C1Frequency*Electrodes	2		0.643	0.534		0.364	0.696
C1Frequency*Repetition*Electrodes	2		3.016	0.067		0.073	0.929
Prosody*C1Frequency*Repetition*Electrodes	2		3.379	0.0509		2.299	0.107
Residuals	72						

(a)



(b)

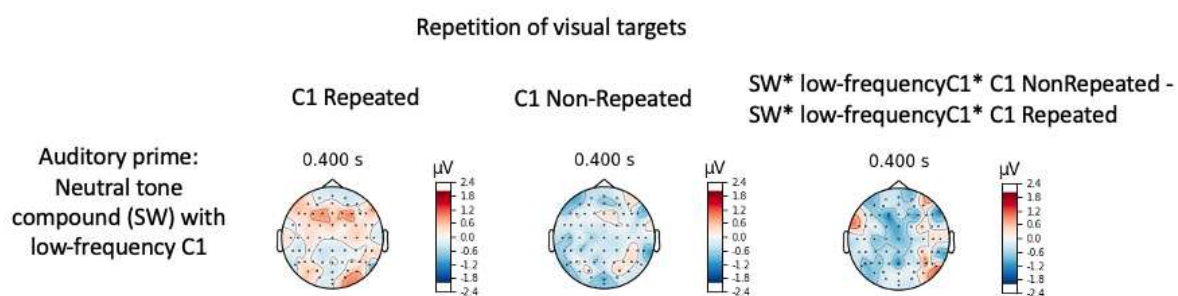


Fig.6.7. (a) Grand averages of ERPs were measured at the C3 and C4 electrodes and time-locked to the onset of visual targets in the two priming conditions, when auditorily preceded by neutral tone compound with low-frequency C1. The time window of N400 is indicated by dashed lines. (b) Topographic maps of the significant effect in the central region in the timing of 400ms post-stimuli, $p < .05$.

6.6 Discussion

The aim of this study was to investigate the priming effect of the first constituents within two-character Mandarin compounds, and whether this priming effect could be modulated by the compound's frequency, the compound's prosody pattern and the first constituent's frequency. We used a cross-modal repetition priming paradigm, in which a spoken compound was presented as an auditory prime, and it was immediately followed by a visual target (a Chinese character, either the first constituent of spoken compound, or a correct character but not the first constituent, or a pseudo-character). Participants must perform a lexical decision task to answer whether the visual target was a correct character or not. The continuous EEG was recorded. As expected, behavioral results revealed a priming effect of the first constituents, a significantly shorter response time was observed for the repeated visual targets compared with the non-repeated visual targets. Consistent findings were revealed in the ERP results, non-repeated visual targets elicited a significantly larger N400 effect than repeated visual targets, the so-called N400 repetition priming effect (increased N400 effect in the non-priming condition). The main question is to investigate whether prosodic information of neutral tones carried by the second constituents and/or the lexical frequency (both the whole compound frequency and the first constituent frequency) works as cues to trigger the morphological decomposition in the lexical access, and we also want to investigate whether there are cumulative effects among the three variables.

The first major finding of our study is the observation of the whole-compound frequency effect and first constituent frequency effect in Mandarin two-character compound's processing, which replicates the previous studies in other languages (Fiorentino & Poeppel, 2007; Hyönä & Pollatsek, 1998; Juhasz et al., 2003; Shoolman & Andrews, 2003). Through the analysis of the mean amplitudes of N400, the frequency effect of the whole-compound and first constituent were found

respectively to modulate the identification of visual targets. These effects were straightforwardly observed in the time window of 350-500ms after the stimuli onset, which presented a clear central spatial distribution.

We found an interaction between the priming conditions of visual targets and compound's frequency of auditory primes, an increased N400 effect was observed in the non-priming condition when auditory primes were rare compounds. Consistent with Chen (2003) and Juhasz et al. (2003), which had proposed that the role of whole-word frequency is important in the recognition of two-character compounds. Our results contribute more to cross-modal priming experiments, whole-compound frequency effect involves in the lexical representation of Mandarin compounds, when the compound's frequency is low, the first constituent might be decomposed, and more cognitive resources will be required in the access of rare compounds in the case of the activation of individual constituent. Moreover, we found an interaction between the priming conditions of visual targets and the frequency of first constituent of auditory primes, an increased N400 effect was observed in the non-priming condition when auditory primes were the compounds with low-frequency first constituents. Our observation of the lexical frequency effects of the first constituents is in line with the morpheme's frequency effect in the previous studies (Yan et al., 2006). In our study, the priming effect of the first constituents, combined with the effect of first constituent's frequency, provided evidence of the involvement of the first constituent in the lexical access of two-character compounds. When the first constituent is low-frequency, there will be a sub-lexical decomposition in the auditory processing, which would cost more resources in the representation procedure. Our results support the assumption that two-character compounds are decomposed and processed via their individual constituents, and consistent with previous study of morphological decomposition in the complex compounds (Andrews et al., 2004; Juhasz et al., 2003). Furthermore, according to the results in the eye-movement research, in order to recognize compound words, readers tend to decompose the whole compound into individual lexemes, a longer fixation would be generated by the infrequent lexemes (Inhoff et al., 2008; Juhasz et al., 2003; Kuperman et al., 2009). Additionally, it is noteworthy that no interaction was revealed respectively between the frequent compounds or high-frequency first constituents and the priming conditions of visual targets. Our results revealed that the frequent compounds and compounds with high-frequency first constituents are turned to be represented in their whole form in the mental storage, while rare compounds and compounds with low-frequency first constituents are turned to be represented through individual constituents. These results seem to be consistent with the investigation of lexical frequency in previous studies (Taft, 1994; Yan et al., 2006; Zhou & Marslen-Wilson, 1995). For example, Zhu and Taft (1994) revealed a robust whole-word frequency effect in the behavioral studies of compound recognition,

the high-frequency compounds were identified faster than low-frequency compounds. In line with the eye-tracking study of Cui et al. (2021), frequent Chinese compound was demonstrated to be processed holistically, whereas the representation of individual character was found to be involved in the processing of infrequent compounds.

The second major finding shed new light on the role of neutral tone in Mandarin speech processing. Moreover, we observed an interaction between the priming conditions of visual targets and the prosody pattern of auditory primes, an increased N400 effect has been observed in the non-priming condition when auditory primes were neutral tone compounds with low-frequency first constituents, and this N400 effect was distributed in the central region. Our results reveal that neutral tone is involved in the lexical representation of two-character compounds, which confirms the role of prosodic information in auditory lexical access. Our findings are consistent with the behavioral results in the previous studies, in which the prosody structure of lexical tones has been indicated to modulate participant's performance in the cross-modal priming experiments (Cutler & Van Donselaar, 2001; Lee & Nusbaum, 1993), some studies in Mandarin and English found that listeners make use of the tonal information in the lexical identification (Connine et al., 1987; Fox & Unkefer, 1985).

Moreover, our results suggest that when the first constituents within two-character neutral tone compounds are low-frequency, the first constituent should be decomposed and activated in the lexical representation, and the compound should be processed via their individual constituent. Because in the auditory modality, the recognition of the second neutral tone constituent must be realized through the combination with the preceding first full tone, global processing of the whole compounds should take place after the activation of the first constituents, which made the lexical representation consume more resources than compounds with two full tones, as reflected by the increased N400 effect in the non-priming condition, which is typically associated with the semantic integration. Our result is in line with Yan et al. (2014), which attributed a more resource-consumed processing of Mandarin neutral tone in lexical access. In their eye-movement study, when neutral tone compound was fixated, the readers processed an upcoming parafoveal word less effectively, compared with the fixations on the full tone compounds.

Depending on these results, we considered that two processing routes should exist in the auditory processing of two-character compounds, including a holistic processing route to activate the whole form of compounds and a decomposition route to represent the compounds via their individual constituents, and the activation of the two routes might be assisted by different cues, such as the frequency of compounds, the frequency of first constituent, and the prosodic pattern of the compounds. Based on the eye-movement study of Cui et al. (2021), which found that the

frequent compounds are turned to be processed holistically, and infrequent compounds are turned to be processed via the individual constituent, we considered that neutral tone compounds with low-frequency first constituents tend to be processed via a decomposition route, the frequency of first constituents combined with the prosody of the second neutral tone plays a crucial role in the decomposition, which may trigger a decomposition route in the auditory processing. The neutral tone is revealed to be involved in the phonological encoding of individual constituents, due to the dependence from the second neutral tone to preceding first full tone in the semantic representation in the auditory modality, once the first full tone is decomposed, an additional holistic processing should occur, which might consume more in listener's cognitive system.

6.7 Conclusion

In line with the previous studies, this study confirmed the whole-compound frequency effect and the individual constituent frequency effect in the auditory processing of two-character Mandarin compounds. Our research reveals that both the holistic processing and decomposition processing routes exist in the auditory processing of Mandarin two-character compounds, and both the whole compound's frequency and the first constituent's frequency might determine the selection of the processing routes. Moreover, our results confirmed the crucial role of prosody in the processing of neutral tone compounds, neutral tone is considered to be involved in the lexical representation. These findings in Mandarin not only support the PAP model of spoken compound recognition proposed by Isel et al. (2003), but also enrich the functional description for Mandarin compounds by Isel and Shen (2020). Importantly, the results suggest that different cues can jointly contribute to assisting the spoken word processing system to select the most efficient processing strategy and recognize the compound words. It seems that the frequency of first constituent, the frequency of compounds, and the prosody pattern (the presence of a neutral tone on the second constituent) are the clues that could be in coalition to accelerate the recognition of compounds, notably by activating or not the morphological decomposition route.

Chapter 7.

EEG-Experiment 2.

7 Influence of syllabic lengthening in Mandarin spoken neutral tone compounds processing

7.1 Introduction

Within the neutral tone compounds, the alternation of the first full tone and second neutral tone creates an opposition in rhythmic structure and a trochaic prosody pattern. In previous works, the behavioral study demonstrated that to analyze the rhythmic structure of words, listeners extract the suprasegmental information in the lexical processing. Consequently, researchers have found that the prosodic information could facilitate the lexical recognition of spoken words, at least in languages with lexical accents, like English and German (Cutler & Van Donselaar, 2001; Friedrich, Kotz, Friederici, & Gunter, 2004). When listeners tried to process the ambiguous bi-morphemic sequences, the subphonemic durational differences across multiple segments can work as a word recognition cue (Spinelli et al., 2003; Tabossi et al., 2000). For instance, authors have reported that, in the context of liaison in French, French listeners used the durational differences contained in the last consonant *r* of the word *dernier* to distinguish the final liaison *dernier oignon* from the initial consonant *roignon*. Some consistent results have also been found by eye-tracking research (Salverda et al., 2003; Shatzman & McQueen, 2006).

According to previous studies, metrical structure plays a crucial role in lexical access, even some subtle violations would consume more resources in the processing procedure (Bohn et al., 2013; Henrich et al., 2014). Researchers have applied the event-related potential (ERP) to examine the processing consequences of violated or non-violated prosody patterns in lexical access. Some ERP studies showed that prosodic pattern influences the auditory processing of words (Böcker et al., 1999; Friedrich, Kotz, Friederici, & Alter, 2004; Knaus et al., 2007; Steinhauer et al., 1999). Moreover, deviant stimuli (e.g., accented/stressed/lengthening syllable within a multimorphemic word) could hinder the lexical identification. Precisely, the irregular syllable length can hinder the word identification, as revealed in French (Magne et al., 2007) and Mandarin (Wang et al., 2005), researchers manipulated the syllable duration of multimorphemic spoken words to investigate the influence of rhythmic properties on semantic processing. French has been described as a language with the fixed accents at the end of the rhythmic units, which are characterized by lengthening the last syllables (Delattre, 1966). Magne et al. (2007) manipulated the duration of the lengthening last syllables to the penultimate syllables (non-accented and normal duration) of the trisyllabic words, as to create the ‘metrical violated condition’, in which the target words had an abnormal syllable duration and violated accents. For example, the word *canapé* (sofa) is normally accented on the last syllable *PE*, while the metrically violated target word *caNapé* may possess an incongruous

lengthening penultimate syllable **NA**, moreover, the semantic congruency of target words was also manipulated to create the semantic violated condition at the sentence level. The sentence contained the target words placed at the last position were aurally presented to native French speakers, results showed an N400-like effect elicited by target words with misplaced stress pattern and incongruous syllable duration. Further, the semantically incongruous words elicited the classical larger N400 in relation to the semantically congruous words in the metrical judgment task, the authors interpreted this enhanced N400-like effect as a reflection of disrupted accessing procedure in response to the semantic integration caused by the abnormal metrical pattern.

In addition, as illustrated in Chapter 4.3.2, a set of studies have investigated the different contribution made respectively by the neutral tone and the full tone in the lexical access. In an eye-tracking study, Yan et al. (2014) found that readers processed an upcoming parafoveal word less effectively when a neutral tone compound was fixated, compared with the fixation on a full tone compound. In an ERP study, Luo et al. (2016) observed a more-negative amplitude of N400 during the readings of the neutral tone compound in relation to the full tone compound. However, more evidence is required to examine whether additional cognitive resources would be consumed in the processing of neutral tones than in the processing of full tones.

7.2 Goal and hypothesis

The difference of syllabic duration contained in the two tones of a two-character neutral tone compound provide us a good test case to examine the function of prosodic information in lexical processing. However, in previous studies, fewer studies investigated the question of whether the violation of syllable duration could modulate the neutral tone compound's recognition in the auditory modality. In the current study, by manipulating the durations of both two constituents within neutral tone compounds, we aimed to examine the neural correlates of native listeners in response to the modulations of syllabic duration in the spoken neutral tone compound processing. We aimed to investigate to what extent the prosody pattern of neutral tone and full tone constrain lexical access. The goal of this research is twofold. Firstly, we wanted to examine if the rhythmic violation of the neutral tone compounds could be detected by native listeners. Secondly, because the trochaic prosody pattern is formed by the two constituents within the neutral tone compounds, another question raised is if the deviations of two constituents (1st full tone & 2nd neutral tone) would elicit different neural correlates in the recognition procedure, to what extent the prosody of full tone and neutral tone could respectively affect the processing. As mentioned earlier (Luo et al., 2016), the irregular syllable duration of neutral tones seems to be more problematic than the

incongruous full tones in the lexical access, we therefore expected to observe the stronger brain electrical responses in the EEG results to the rhythmic violations of neutral tones than full tones.

Two variables were manipulated in this research, including **the constituents of neutral tone compound** (2 levels: full tone C1 and neutral tone C2) and **syllable durations** (2 levels: standard and deviant). The deviant rhythmic pattern was created by lengthening the duration of the constituents within compounds as to construct the prosodic incongruent conditions. Thus, a 2x2 experimental design resulting into four conditions were constructed, including Condition 1: two-character neutral tone compounds whose two syllables are in standard phonological form; Condition 2: deviant C1 & standard C2; Condition 3: standard C1 & deviant C2; Condition 4: deviant C1 & deviant C2 (Fig.7.1). Because the ERP technique was considered to provide specific information about the brain responses in word recognition, the neural correlates of listeners will be recorded during the online processing of the violated rhythmic incongruency. In each trial of experiment, participants listened to the sentences which contain the target stimulus, and they must make the decision, if the rhythmic pattern of the sentence was correct or not.

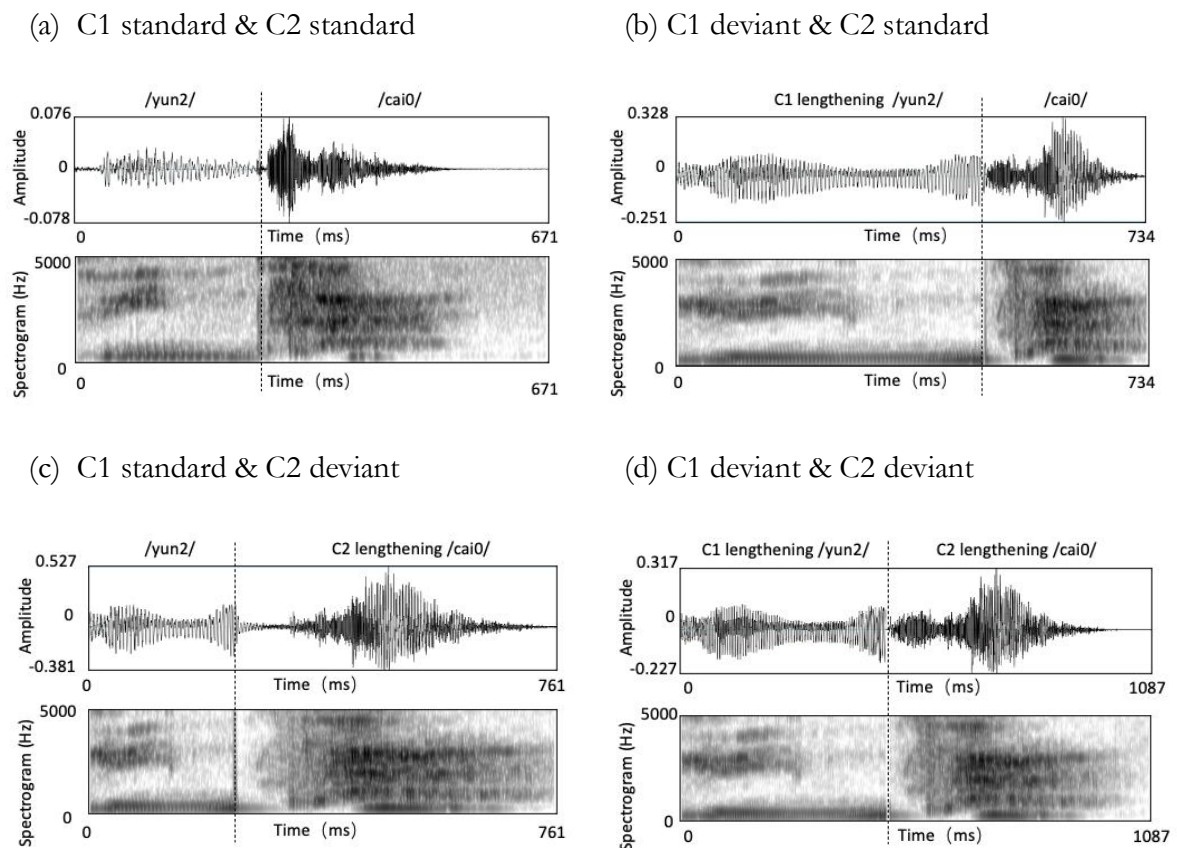


Fig.7.1. (a). The standard phonological form of neutral tone compound 云彩/yun2cai0/ ‘cloud’ was pronounced in a citation form. (b) The first constituent (C1) 云 yun (T2) is lengthening. (c) The second constituent (C2) 彩 cai (T0) is lengthening. (d) Two constituents 云 yun (T2) and 彩 cai (T0) are lengthening.

Accordingly, our goal was to investigate the contribution of syllable duration in the neutral tone compound processing. In our assumption, three ERP components would be involved in the processing of metrical violations. Firstly, as a typical ERP component that is related to the acoustic changes contained in auditory stimulus, the increased negativity of N200 should be elicited by the unexpected prosodic incongruity (Friedrich et al., 2001; Friedrich, Kotz, Friederici, & Alter, 2004; Näätänen, 1991; Näätänen & Alho, 1997; Näätänen et al., 2004; Näätänen & Winkler, 1999). Secondly, because MMN has been found in previous study as an ERP component corresponding to the mismatch between the standard phonological form of speech sounds and irregular auditory inputs, we assumed that the MMN should be yielded during the processing of incongruous metric structure of target stimuli (Näätänen, 1991; Näätänen & Alho, 1997; Näätänen et al., 2004; Winkler et al., 1999). Finally, based on the study of Magne et al. (2007), the N400 effect should be induced by the difficulties in the semantic processing due to the acoustic changes in the prosodic pattern, we assumed that prosodic violation could hinder the processing of neutral tone compounds, which should elicit the MMN, N200, and N400 effects due to the changes in the syllable duration. This could be viewed as evidence of the additional resource cost in the semantic processing when the listeners encountered the incongruous prosody in the speech input. In addition, considering the two prosodically different constituents within neutral tone compound (1st full tone & 2nd neutral tone), Wang (2004) revealed that duration is one of the most important acoustic cues to recognize a neutral tone syllable, we assumed that metrically incongruous neutral tone (C2) would induce the larger amplitude of MMN and N400 effects than the metrically incongruous full tone (C1).

7.3 Methodology

Participants Twenty-two native speakers (16 females and 6 males; age: 18 to 35, mean: 28.7 ± 4.37) of Chinese Mandarin who were college students in the universities in Paris were paid 10 euros per hour for their participation in the experimentation. None of them had lived in France for more than five years (1 to 5 years, mean 3.09 ± 1.26), and they spoke Mandarin frequently in their daily life (1 to 4 hours per day, mean 2.36 ± 1.04). Participants were right-handed with normal or corrected-to-normal vision and without a history of neurological, psychiatric, or any auditory impairment. Participants were clearly instructed on the content of tasks. They gave their written informed consent before the experiments. The research protocol was approved by the local ethical committee of Paris Nanterre University.

Materials 60 two-character Mandarin neutral tone compounds [Strong-weak] were firstly selected in Condition 1, e.g., compound ‘巴掌 ba (T1) zhang (T0)’ (hand palm). Regarding the previous acoustic research, the first full tone constituent (C1) ‘巴 ba (T1)’ within the neutral tone

compound is shorter than when it is pronounced as single character. More specifically, the duration of the first full tone constituent (C1) of a neutral tone compound was revealed to be about 50% of the length of the same constituent when pronounced in the monosyllabic form, based on our acoustic research in Chapter 5. This latter result clearly indicates that the duration of C1 is a reliable cue to distinguish a character produced in isolation from a character produced in a two-character neutral tone compound. Moreover, according to Lin and Yan (1990), the duration of the second neutral tone constituent (C2) ‘掌 zhang (T0)’ is shorter than its full tone counterpart ‘zhang (T3) 掌’ in a third tone (T3), i.e., the duration is almost 50% shorter than the phonological form of its full tone counterpart. In the present study, rhythmic violation of selected compounds was created by lengthening the syllable duration of the first full tone constituent (C1) and the second neutral tone constituent (C2). For this purpose, Praat (Boersma & Weenink, 2006) was used to prolong the target syllable within each stimulus, and the lengthening ratio that we applied was equal to 1.98, the syllable duration of both two constituents of the neutral tone compound was approximately doubled in the lengthening manipulation:

‘Deviant constituent 1 (C1^{deviant})’: duration of the first full tone constituent (C1) of neutral tone compound is 1.98 times as long as in its monosyllabic form.

‘Deviant constituent 2 (C2^{deviant})’: duration of the second neutral tone constituent (C2) of neutral tone compound is 1.98 times as long as in its full tone form.

Based on the durational differences, we manipulated the syllabic duration respectively for both the two constituents C1 and C2 within neutral tone compounds to create the metrical violated condition. Phonetically violated stimuli in the Condition 2/3/4 were constructed based on the 60 neutral tone compounds in Condition 1, including the 60 partially violated stimuli in Condition 2 (**deviant C1** + standard C2), 60 partially violated stimuli in Condition 3 (standard C1 + **deviant C2**), 60 fully violated stimuli in Condition 4 (**deviant C1** + **deviant C2**) (Table 7.1).

Table.7.1. Examples in the four experimental conditions.

Condition 1	Condition 2	Condition 3	Condition 4
Standard [C1 + C2]	C1 deviant [C1 ^{deviant} + C2(T0)]	C2 deviant [C1(full tone) + C2 ^{deviant}]	C1 deviant C2 deviant [C1 ^{deviant} + C2 ^{deviant}]

C1 + C2 (T0)	deviant C1 + C2 (T0)	C1 + deviant C2	deviant C1 + deviant C2
巴 掌 (T0)	巴 ⁺ 掌 (T0)	巴 掌 ⁺	巴 ⁺ 掌 ⁺
ba (T1) zhang (T0)	ba (T1) zhang (T0)	ba (T1) zhang (T3)	ba (T1) zhang (T3)
<i>hand palm</i>	<i>hand palm</i>	<i>hand palm</i>	<i>hand palm</i>

Sixty neutral tone compounds in Condition 1 were selected from ‘Contemporary Chinese Dictionary’ (Lu & Ding, 2008) the lexical frequency was matched for word frequency (mean 289.7 occurrences per 10 million, ± 22.89) in the Corpus of contemporary Mandarin by the China State language commission (<http://corpus.zhonghuayuwen.org/Resources.aspx>)³ (Table in Appendix 4). The time-stretching function in Praat was applied (Boersma & Weenink, 2006), which allowed us to manipulate the duration of the acoustic signals without changing the fundamental frequency of target syllables. The metrical deviant stimuli in Condition 2/3/4 were created by lengthening respectively the two constituents of the neutral tone compounds in Condition 1 (standard: C1^{standard} + C2^{standard}), the duration of C1 was lengthened (partially violated: C1^{deviant} + C2^{standard}) in Condition 2, the duration of C2 was lengthened (partially violated: C1^{standard} + C2^{deviant}) in Condition 3, and the duration of both the two syllables was lengthened in Condition 4 (fully violated: C1^{deviant} + C2^{deviant}). Thus, each of the four experimental conditions resulted in 60 target stimuli compounds that were embedded in the semantically congruent utterances. The 240 sentences (60*4=240) were selected from the corpus: BLCU Corpus center (<http://bcc.bcu.edu.cn/>), and the number of characters contained in the selected sentences were matched (mean 22.8 \pm 4.35). All the 240 sentences were recorded in the declarative mode by a native female speaker who spoke standard Mandarin in an anechoic chamber using a digital audiotape.

7.4 ERP recording and analyses

Experiment procedure The participants were explicitly instructed to pay attention to the prosody pattern of words in the sentences, and they were instructed to decide whether the words were rhythmically incongruous or not. Each trial contained one sentence, one critical stimulus was embedded in it, the duration of stimuli was either lengthened (the first or the second constituent, or both) or not (a standard neutral tone compound). Each trial in the experiment started with a fixation cross at the center of the screen for 1000ms and followed by a blank screen for 1000ms. One sentence was played after the blank screen, and after the offset of each sentence, a “?” was displayed on the screen to cue the participant to answer the question, if there was a rhythmically

³ Corpus of contemporary Mandarin is an on-line word-text corpus, which comprises commonly used written and spoken Chinese from a wide range of sources, including approximately 10 million words (20 million in characters).

incongruous word contained in the sentence or not, and the listeners were asked to make the ‘yes’ or ‘no’ decision by pressing the buttons on the keyboard, their response delay was 3000ms, which would be interrupted as soon as the participants gave their responses. The intertrial interval (ITI) was 1500ms. We collected behavioral results and continuous EEG data in the task, the reaction times of each participant were recorded as the behavioral results in metrical pattern discrimination tasks, the EEG signals were recorded in the experiment, and ERP results were extracted from the onset of target stimuli to the offset of the stimuli in the preprocessing. And the order of sentences was randomized across the four conditions and twenty-two participants.

EEG recordings and preprocessing EEG signals were recorded from the scalp by using the 64-channel Biosemi ActiveTwo AD box with a sampling rate of 512 Hz. An online bandpass filter was applied with a lowpass cutoff was 100Hz and a highpass cutoff was 0.05Hz. To measure the eye movements from the two eyes, four electrodes were applied for electrooculogram (EOG), the measurements were obtained from below versus above the left eye (vertical EOG) and the left versus right lateral orbital rim (horizontal EOG), moreover, two additional electrodes were placed over the right and left mastoids. After the recordings, we used Python-MNE for the preprocessing of EEG data. Firstly, through a visual inspection, the bad channels which contained EMG, or the other artifacts (not eye-movements or blinks) were manually removed. Then, we used the algebraic average of the left and right mastoids (M1 & M2) to perform the interpolation of the bad channels, and each electrode was re-referenced offline to the algebraic average of the left and right mastoids (M1 & M2). Secondly, the remaining data (except M1 & M2) were offline bandpass-filtered from 0.05 to 30 Hz, and the data were corrected for the eye-blink artifacts by using the ICA to identify and subtract the oculomotor artifacts. Finally, cleaned data were segmented into epochs spanning -200ms to 800ms corresponding to the onset and offset of target stimuli, another visual inspection was realized to reject manually the epochs containing the excessive artifacts. As for the remaining epochs, a baseline correction was run to the pre-stimulus time window from -200ms to 0ms. To analyze the ERP results, the collected epochs were averaged within and across the participants to obtain the grand averages in the four experimental conditions.

EEG results analyses Preprocessed ERP data were analyzed with the Repeated measures ANOVAs in RStudio. We extracted the mean amplitudes within the time window of 120-200ms for the MMN component, in the time window of 220-300ms for the N200 component, and in the time windows of 350-550ms for the N400 component. The selection of time windows was based on the ERP waves which were triggered by the rhythmically congruous or incongruous stimuli in the four conditions. The mean amplitudes were extracted from 9 fronto-central and centro-parietal electrodes (F3, Fz, F4/C3, Cz, C4/P3, Pz, P4). Firstly, the amplitude at the three midline electrodes

(Fz/Cz/Pz) was entered in the ANOVAs. Then, in the lateral electrodes, we selected these three regions of interest (anterior, central, and posterior), and six lateral electrodes were selected for the three regions (anterior: F3&F4; central: C3&C4; posterior: P3&P4), the mean amplitudes of lateral electrodes were entered the ANOVAs by measuring the mean amplitudes of the three regions of interests, including the average of F3 & F4 for the anterior region, the average of C3 & C4 for the central region and the average of P3 & P4 for the posterior region. To adjust for the violation of sphericity (Greenhouse & Geisser, 1959), the Greenhouse-Geisser correction was applied when there was more than one degree of freedom in the numerator. To investigate the influences of the incongruent prosody patterns in the processing of neutral tone compounds, Repeated two-way ANOVAs were conducted for the mean amplitudes at both the midline and lateral electrodes. In the results of ANOVAs, when one significant main effect or interaction was observed, Post-hoc Tukey tests were performed to interpret the result.

7.5 Results

7.5.1 Behavioral results

We examined the reaction times (RT) of the participants for the four conditions as well as accuracy, and the reaction times which were greater than two SDs beyond the global mean were discarded (10.6%). The accuracy rate was 82.62 % for all the responses. As for the error rates of the four conditions, participants made significantly fewer errors to the stimuli in Condition 1 that had no phonological incongruency (17.48%), more errors were made when the first constituent (C1) was phonologically incongruent in Condition 2 (33.25 %), and when the second constituent (C2) was incongruent in Condition 3 (32.07%), finally, when both the two constituents were incongruent in Condition 4, the error rate was the highest (39.38%). Moreover, the Repeated one-way ANOVAs were performed, and no significant difference was revealed between the reaction times in the four conditions [$F(1, 21) = 1.521$, $SD = 0.122$, $p = 0.218$, $p > .05$] (Fig.7.2).

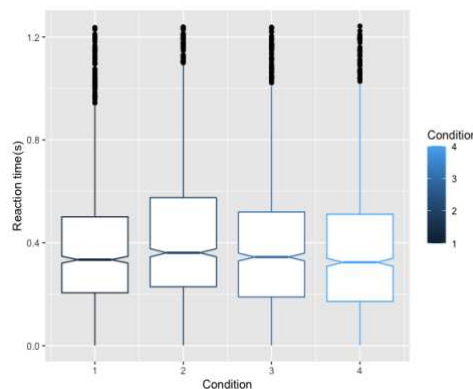


Fig.7.2. Boxplot for the reaction times in the four conditions ($p > .05$).

7.5.2 ERP results

The mean amplitudes of the three time-windows at both midline and lateral electrodes are presented in Table.7.2, results of Repeated two-way ANOVAs [Condition (1/2/3/4) * Electrodes (midline: Fz/Cz/Pz; lateral: anterior, central, posterior)] are presented in the Table.7.3. In Fig.7.3 and Fig.7.4, the ERP waveforms and global topographical figures of four conditions are presented, in Fig.7.5, Fig.7.6, Fig.7.7, biphasic waveforms and biphasic topographical figures are presented for the comparisons respectively between the Condition1 (standard) and Condition 2 (C1^{deviant}) / Condition 3 (C2^{deviant}) / Condition 4 (C1^{deviant} C2^{deviant}).

Table 7.2. Mean amplitudes of three time windows at both midline and lateral electrodes (μV , mean \pm *SD*).

Condition	Electrodes	Time window 120-200ms	Time window 220-300ms	Time window 350-550ms
Site midline (Fz/Cz/Pz)				
Condition 1	Anterior	0.69 \pm 0.036	1.41 \pm 0.297	1.38 \pm 0.253
Standard	Central	1.14 \pm 0.112	1.52 \pm 0.243	2.24 \pm 0.807
	Posterior	1.77 \pm 0.606	2.25 \pm 0.739	3.77 \pm 0.526
Condition 2	Anterior	1.18 \pm 0.062	1.32 \pm 0.002	2.11 \pm 0.485
C1 deviant	Central	1.63 \pm 0.193	1.88 \pm 0.398	2.98 \pm 0.888
	Posterior	2.99 \pm 0.368	3.99 \pm 0.561	5.56 \pm 0.626
Condition 3	Anterior	-1.25 \pm 0.005	-1.94 \pm 0.084	-0.12 \pm 0.323
C2 deviant	Central	-0.91 \pm 0.353	-1.04 \pm 0.589	1.39 \pm 0.828
	Posterior	-0.15 \pm 0.766	0.18 \pm 0.983	3.09 \pm 1.165
Condition 4	Anterior	0.53 \pm 0.293	0.89 \pm 0.064	0.98 \pm 0.029
C1C2 deviant	Central	0.12 \pm 0.270	0.92 \pm 0.112	0.99 \pm 0.029
	Posterior	0.21 \pm 0.322	1.86 \pm 0.398	3.09 \pm 0.604
Site lateral (anterior F3; F4/ central C3; C4/ posterior P3; P4)				
Condition 1	Anterior	0.35 \pm 0.109	1.06 \pm 0.802	1.25 \pm 0.762
Standard	Central	0.92 \pm 0.031	1.28 \pm 0.007	2.08 \pm 0.261
	Posterior	2.05 \pm 0.036	2.28 \pm 0.267	3.71 \pm 0.343
Condition 2	Anterior	0.74 \pm 0.237	0.58 \pm 0.001	0.69 \pm 0.364
C1 deviant	Central	1.53 \pm 0.276	2.01 \pm 0.206	3.22 \pm 0.039

	Posterior	2.80 ±0.679	4.01 ±0.787	5.13 ±0.593
Condition 3	Anterior	-1.92 ±0.907	-2.37 ±0.401	-1.07 ±0.096
C2 deviant	Central	-1.13 ±0.390	-1.47 ±0.322	0.79 ±0.071
	Posterior	0.24 ±0.189	0.01 ±0.007	3.33 ±0.394
Condition 4	Anterior	0.36 ±0.303	0.45 ±0.214	0.68 ±0.498
C1C2 deviant	Central	-0.76 ±0.627	-0.13 ±0.624	0.59 ±0.461
	Posterior	0.34 ±0.247	0.14 ±0.148	3.35 ±0.115

Table.7.3. Summary of ANOVAs analysis computed at the midline and lateral electrodes in the three time windows. *F* is for F value, *p* is for p-value, *d* is for difference in mean amplitude (μV). Significance coding: *p*<.001, *p*<.01, *p*<.05.

Factors	df	Time window			Time window			Time window		
		120-200ms			220-300ms			350-550ms		
		<i>F</i>	<i>p</i>	<i>d</i> (μV)	<i>F</i>	<i>p</i>	<i>d</i> (μV)	<i>F</i>	<i>p</i>	<i>d</i> (μV)
Site midline (Fz, Cz, Pz)										
Condition	3	63.07	<.001	-	56.51	<.001	-	11.40	<.001	-
C1deviant-Standard		-	.058	0.73	-	.116	0.67	-	.054	1.12
C2deviant-Standard		-	<.001	-1.94	-	<.001	-2.66	-	.091	-0.98
C1C2deviant-Standard		-	.003	-0.92	-	.300	-0.51	-	.594	-0.47
C2deviant-C1deviant		-	<.001	-2.67	-	<.001	-1.47	-	<.001	-2.1
C1C2deviant-C1deviant		-	<.001	-1.65	-	.004	-1.18	-	.005	-1.59
C2deviant-C1C2deviant		-	.001	1.02	-	<.001	-2.15	-	.556	-0.51
Electrodes	2	15.57	<.001	-	26.89	<.001	-	41.55	<.001	-
Condition*Electrodes	6	3.16	.042	-	2.01	.142	-	0.539	.768	-
Residuals: 12										
Site lateral (anterior F3; F4/ central C3; C4/ posterior P3; P4)										
Condition	3	45.69	<.001	-	69.41	<.001	-	29.52	<.001	-
C1deviant-Standard		-	.132	0.59	-	.074	0.66	-	.050	0.66
C2deviant-Standard		-	<.001	-2.03	-	<.001	-2.62	-	<.001	-1.34
C1C2deviant-Standard		-	.003	-1.11	-	.015	-0.88	-	.019	-0.81
C2deviant-C1deviant		-	<.001	-2.62	-	<.001	-3.28	-	<.001	-2.00

C1C2deviant-C1deviant	-	<.001	-1.7	-	<.001	-1.54	-	<.001	-1.47	
C2deviant-C1C2deviant	-	.012	-0.92	-	<.001	-1.74	-	.156	-0.53	
Region	2	27.59	<.001	-	61.65	<.001	-	57.82	<.001	-
Condition*Region	6	3.53	.029	-	4.47	.013	-	5.86	.004	-
Residuals: 12										

120-200 ms time window

Midline electrodes:

A main effect of **Condition** was revealed ($F(3,12) = 63.07, p < .001, MSE = 0.331, \eta^2 p = 0.75$). Further TukeyHSD was created for the detailed effects of the four conditions. This test indicated that the comparisons of the mean amplitudes between the C1^{deviant} condition and standard condition didn't reach the significance ($d = 0.73\mu V, p > .05$). The mean amplitudes of the C2^{deviant} condition and C1^{deviant}C2^{deviant} condition were significantly more negative than standard condition ($d = -1.94\mu V, p < .001; d = -0.92\mu V, p < .01$, respectively), and the mean amplitudes elicited by the C2^{deviant} condition were more negative than C1^{deviant}C2^{deviant} condition ($d = 1.02\mu V, p < .01$).

Moreover, results of ANOVAs showed a significant interaction between **Condition** and **Electrodes** ($F(6,12) = 3.16, p < .05, MSE = 0.063, \eta^2 p = 0.08$), indicating that the C2^{deviant} condition elicited a significantly larger MMN effect than the standard condition / C1^{deviant} condition / C1^{deviant}C2^{deviant} condition at the anterior electrode (Fz) in relation to the other midline electrodes. And the C2^{deviant} condition elicited the significantly larger MMN amplitude than the standard condition / C1^{deviant} condition in the central electrode (the mean amplitude of Cz of the C2^{deviant} condition was more negative than the mean amplitude of Cz of the standard condition / C1^{deviant} condition, $p = 0.012, p = 0.001$, respectively).

Lateral electrodes:

A significant main effect of **Condition** was revealed ($F(3,12) = 45.69, p < .001, MSE = 0.665, \eta^2 p = 0.61$). Tukey HSD tests revealed that the comparison of mean amplitude between the C1^{deviant} condition and standard condition didn't reach the significance ($d = 0.59\mu V, p > .05$), the mean amplitudes of the C2^{deviant} condition and C1^{deviant}C2^{deviant} condition were significantly larger than standard condition ($d = -2.03\mu V, p < .001; d = -1.11\mu V, p < .01$, respectively), and the C2^{deviant} condition elicited the larger mean amplitudes than C1^{deviant}C2^{deviant} condition ($d = -0.92\mu V, p < .05$). The results indicate that within the 120 - 200ms time window, participant's neutral correlates were modulated by the incongruent syllable duration of the stimuli in the C2^{deviant} condition and C1^{deviant}

C2^{deviant} condition, whereas listener's brain response wasn't significantly modulated by the stimuli in the C1^{deviant} condition.

Additionally, a significant interaction between the **Condition** and **Electrodes** was reported ($F(6,12) = 45.69, p < .05, MSE = 0.091, \eta^2 p = 0.09$), indicating that C2^{deviant} condition elicited the larger MMN amplitudes than standard condition / C1^{deviant} condition / C1^{deviant} C2^{deviant} condition in the anterior region (the mean amplitudes of C2^{deviant} condition in the anterior region were more negative than the mean amplitudes of the standard condition / C1^{deviant} condition / C1^{deviant} C2^{deviant} condition, $p = 0.005; p = 0.001; p = 0.005$, respectively), and the C2^{deviant} condition elicited the larger MMN amplitudes than the standard condition / C1^{deviant} condition in the central region (the mean amplitudes of C2^{deviant} condition in the central region were more negative than the standard condition / C1^{deviant} condition, $p = 0.01; p = 0.001$, respectively).

220-300 ms time window

Midline electrodes:

A main effect of **Condition** was revealed ($F(3,12) = 56.51, p < .001, MSE = 0.718, \eta^2 p = 0.69$). Tukey HSD test indicated that the mean amplitudes elicited by the C2^{deviant} condition were significantly larger than the standard condition ($d = -2.66 \mu V, p < .001$), the comparisons of mean amplitudes between the C1^{deviant} and C1^{deviant} C2^{deviant} conditions vs. standard condition ($d = 0.67 \mu V, p > .05; d = -0.51 \mu V, p > .05$, respectively) didn't reach the significance, the stimuli in the C2^{deviant} condition elicited larger mean amplitudes than C1^{deviant} C2^{deviant} condition ($d = -2.15 \mu V, p < .001$).

Lateral electrodes:

A main effect of **Condition** was reported ($F(3,12) = 69.41, p < .001, MSE = 0.665, \eta^2 p = 0.56$). Tukey HSD tests revealed that the comparisons between the mean amplitude elicited by the C1^{deviant} condition and standard condition didn't reach the significance ($d = 0.66 \mu V, p > .05$), the mean amplitudes elicited by the C2^{deviant} condition and C1^{deviant} C2^{deviant} condition were significantly larger than the standard condition ($d = -2.62 \mu V, p < .001; d = -0.88 \mu V, p < .05$, respectively), the mean amplitudes elicited by the C2^{deviant} condition were significantly larger than the C1^{deviant} C2^{deviant} condition ($d = -1.74 \mu V, p < .001$). The results indicate that within the 220-300ms time window, participant's neural correlates were significantly modulated by the incongruent syllabic duration of the stimuli in the C2^{deviant} condition and C1^{deviant} C2^{deviant} condition, listener's brain response was not significantly modulated by the stimuli in the C1^{deviant} condition.

Moreover, an interaction between **Condition** and **Electrodes** was revealed ($F(6,12) = 4.47, p < .05, MSE = 0.087, \eta^2 p = 0.07$), the stimuli in the C2^{deviant} condition elicited the larger amplitudes of N2b than the standard condition / C1^{deviant} condition / C1^{deviant} C2^{deviant} condition in the anterior region (the mean amplitudes of the C2^{deviant} condition in the anterior region were larger than the standard condition / C1^{deviant} condition / C1^{deviant} C2^{deviant} condition, $p < .001; p < .001; p < .001$, respectively), and the stimuli in the C2^{deviant} condition elicited the larger amplitudes of N2b than the standard condition / C1^{deviant} condition in the central region (the mean amplitude of the C2^{deviant} condition in the central region was larger than standard condition / C1^{deviant} condition, $p < .001; p < .001$, respectively), and this interaction at the midline electrodes didn't reach the significance ($F(6,12) = 2.016, p > .05$).

350-550 ms time window

Midline electrodes:

The ANOVAs showed a significant main effect of **Condition** ($F(3,12) = 11.40, p < .001, MSE = 0.211, \eta^2 p = 0.26$). Tukey HSD tests revealed that the comparisons of the mean amplitudes between the C1^{deviant} / C2^{deviant} / C1^{deviant} C2^{deviant} conditions vs. standard condition did not reach the significance ($d = 1.12\mu V, p > .05; d = -0.98\mu V, p > .05; d = -0.47\mu V, p > .05$, respectively).

Lateral electrodes:

A significant main effect of **Condition** was revealed ($F(3,12) = 29.52, p < .001, MSE = 2.387, \eta^2 p = 0.2$). Tukey HSD tests indicated that the mean amplitudes of the C1^{deviant}, C2^{deviant} and C1^{deviant} C2^{deviant} conditions were significantly larger than the standard condition ($d = 0.66\mu V, p < .05; d = -1.34\mu V, p < .001; d = -0.81\mu V, p < .05$, respectively), and the C1^{deviant} C2^{deviant} condition elicited the larger mean amplitudes than the C1^{deviant} condition ($d = -1.47\mu V, p > .001$). The results revealed that within the 350–550ms time window, participants' neutral correlates were significantly modulated by the incongruent syllable duration of the stimuli in the C1^{deviant}, C2^{deviant} and C1^{deviant} C2^{deviant} conditions.

Furthermore, an interaction between **Condition** and **Electrodes** was revealed ($F(6,12) = 5.86, p < .01, MSE = 0.078, \eta^2 p = 0.08$), the largest **Condition** effect of the N400 was revealed in the anterior region (the mean amplitudes of the C2^{deviant} condition were more negative than the standard condition / C1^{deviant} condition / C1^{deviant} C2^{deviant} condition, $p < .001; p < .05; p < .05$, respectively), and this interaction at the midline electrodes didn't reach the significance ($F(6,12) = 0.539, p > .05$).

ERP waveform amplitude analyses

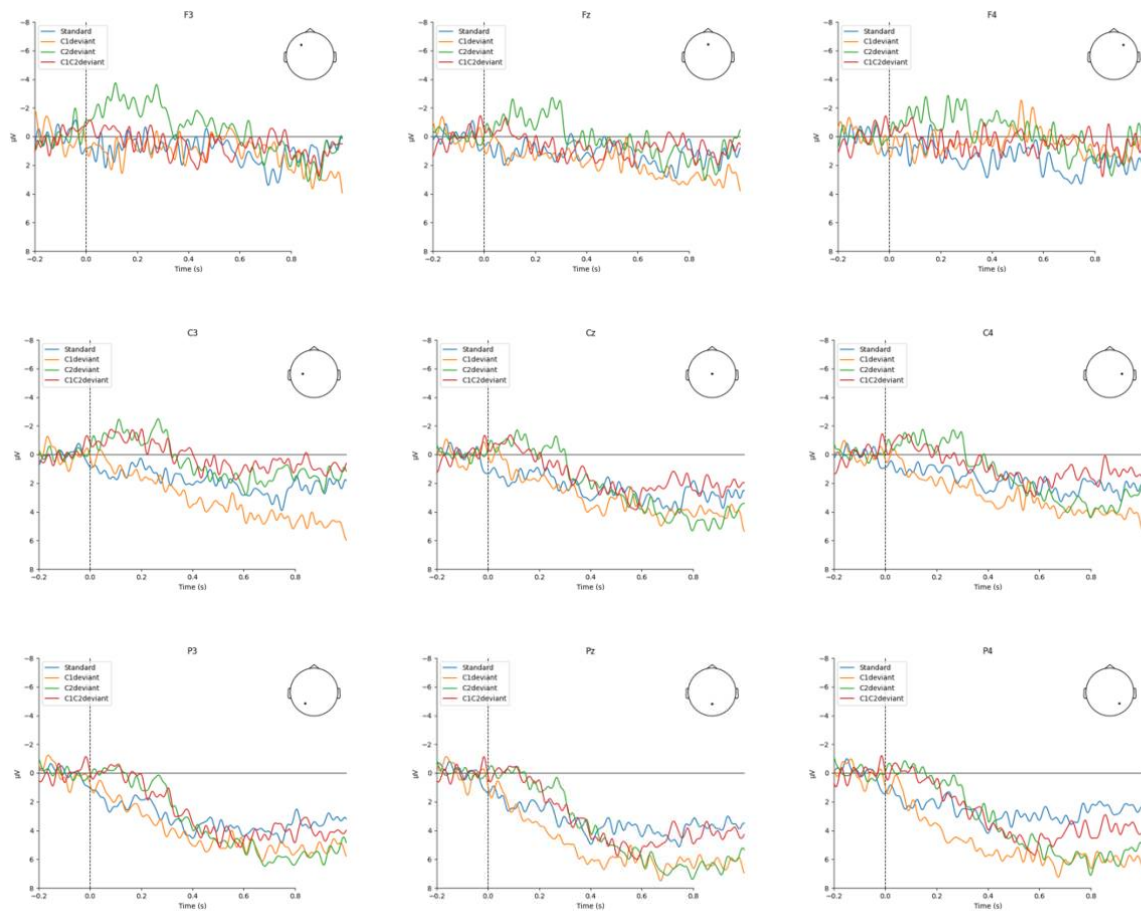
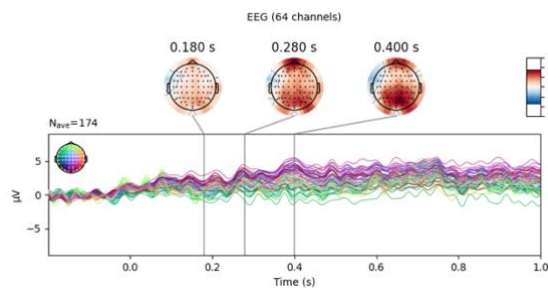


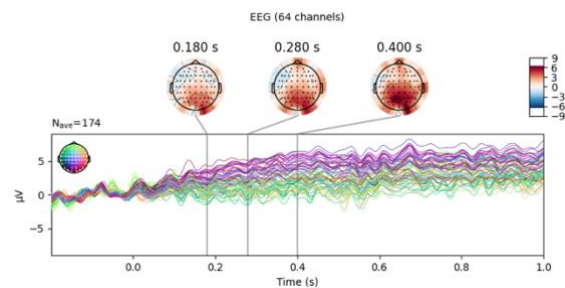
Fig.7.3. Grand-averaged ERPs time-locked to the onset of the target stimuli in the four conditions (blue: standard; orange: C1 deviant; green: C2 deviant; red: C1deviant C2deviant), which were recorded at the midline electrodes (Fz, Cz, Pz) and at the lateral electrodes (anterior: F3, F4; central: C3, C4; posterior: P3, P4). In each ERP waveform figure, negativity is plotted upwards, and positivity is plotted downwards.

Spatial-central segmentation analysis

(a) Standard



(b) C1 deviant



(c) C2 deviant

(d) C1 deviant C2 deviant

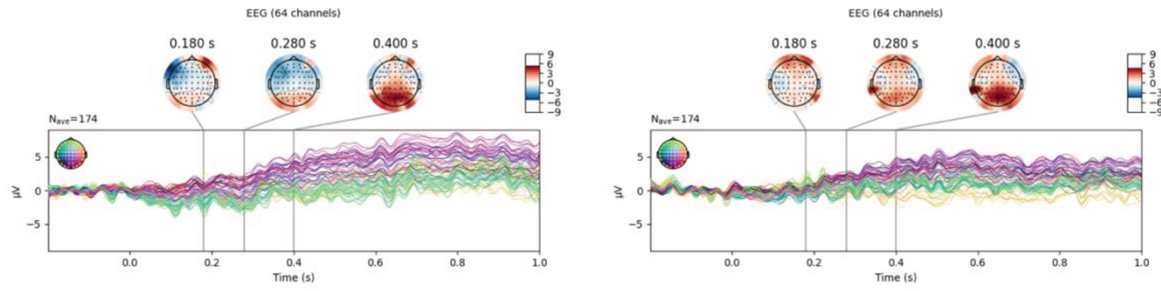


Fig.7.4. Topography of mean amplitude for three timings of interests in the four conditions (a: standard; b: C1 deviant; c: C2 deviant; d: C1 deviant C2deviant), the central distribution is revealed in each topographic map by the spatial-central segmentation analysis, negativity is plotted downwards, and positivity is plotted upwards.

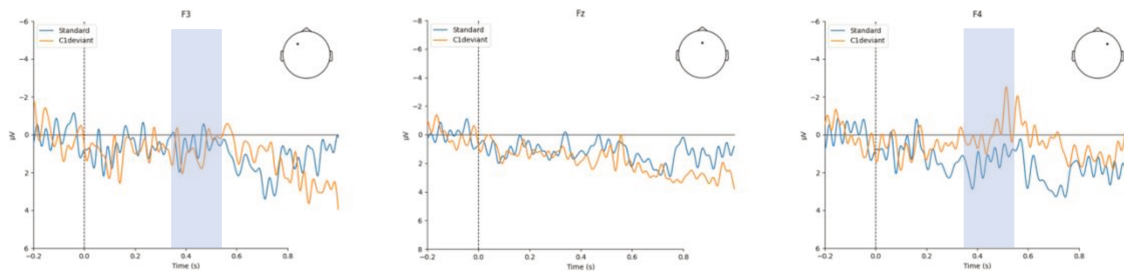
7.5.3 Comparison of Standard (Condition 1) vs. C1 deviant (Condition 2)

MMN: In the 120–200ms time window, Tukey HSD revealed no significant MMN effect between the standard condition and the C1^{deviant} Condition in the midline and lateral electrodes.

N2b: In the 220–300ms time window, Tukey HSD revealed no significant N2b effect between the standard condition and C1^{deviant} Condition in the midline and lateral electrodes.

N400: In the 350–550ms time window, Tukey HSD revealed the significantly larger N400 effect elicited by the C1^{deviant} Condition than the standard condition in the anterior region of the lateral electrodes ($p = 0.041$). A late N400 was observed at the lateral electrodes (F3, F4) with the peak amplitude at around 550ms after the stimuli onset, which is much later than the typical N400 component that peaks at 400ms post-stimuli (Fig.7.5).

(a)



(b)

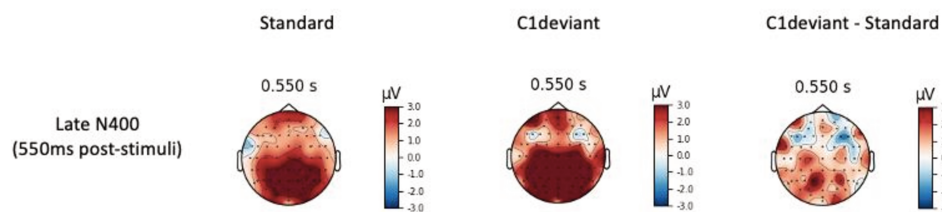


Fig.7.5. (a) Grand-averaged ERPs time-locked to the onset of stimuli in the two conditions (blue: standard; orange: C1deviant). The Time window of the N400 component is indicated in blue. (b) Scalp maps show the topographies of the standard condition and C1 deviant condition for the late N400 component in the 550ms post-stimuli.

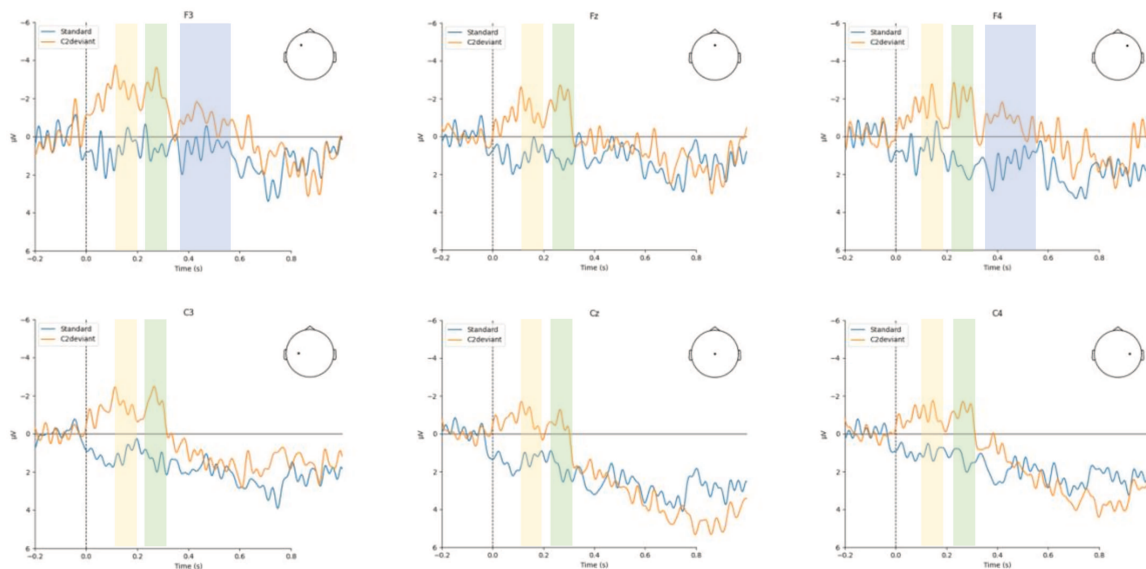
7.5.4 Comparison of Standard (Condition 1) vs. C2 deviant (Condition 3)

MMN: In the 120–200ms time window, Tukey HSD revealed a significant MMN effect between the standard condition and the C2^{deviant} Condition at the midline electrodes and the lateral electrodes, significant MMN effects were distributed in the anterior region (p [midline] = 0.004; p [lateral] = 0.005) and the central region (p [midline] = 0.003, p [lateral] = 0.012).

N2b: In the 220–300ms time window, Tukey HSD test revealed a significant N2b effect between the standard condition and the C2^{deviant} Condition at the midline and the lateral electrodes, significant N2b effects were distributed in the anterior region (p [midline] < .001; p [lateral] < .001) and the central region (p [midline] = 0.004; p [lateral] < .001).

N400: In the 350–550ms time window, Tukey HSD test revealed a significant N400 effect between the standard condition and the C2^{deviant} Condition at the lateral electrodes, the significant N400 effects were distributed in the anterior region (p = 0.002) (Fig. 7.6).

(a)



(b)

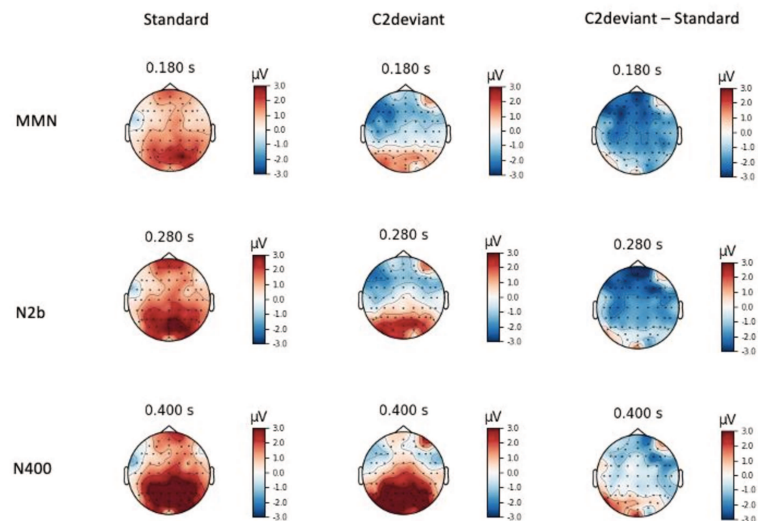


Fig.7.6. (a) Grand-averaged ERPs time-locked to the onset of stimuli in the two conditions (blue: standard; orange: C2 deviant). Time windows of MMN, N2b, and N400 are indicated by yellow, green, and blue. (b) Scalp maps show the topographies of standard condition and C2 deviant condition for MMN, N2b, and N400 components.

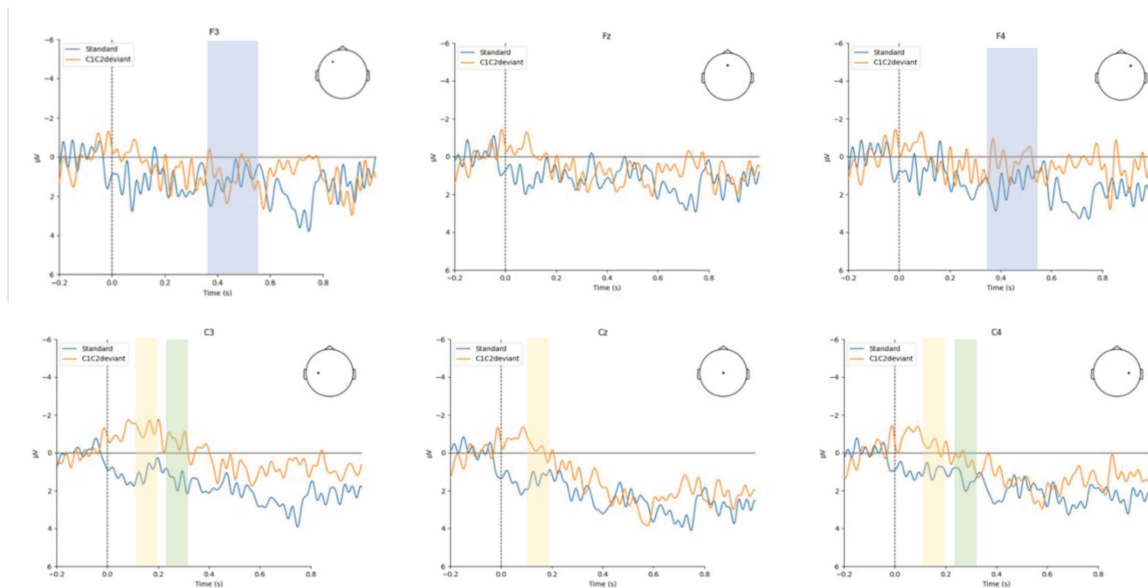
7.5.5 Comparison of Standard (Condition 1) vs.C1deviant C2 deviant (Condition 4)

MMN: In the 120–200ms time window, Tukey HSD revealed a significant MMN effect between the standard condition and C1^{deviant} C2^{deviant} Condition at the midline electrodes and in the lateral electrodes, the significant MMN effects were distributed in the central region (p [midline] = 0.03; p [lateral] = 0.046).

N2b: In the 220–300ms time window, Tukey HSD showed significant N2b effect between the standard condition and the C1^{deviant} C2^{deviant} condition at the lateral electrodes, significant N2b effects were distributed in the central region (p = 0.048).

N400: In the 350–550ms time window, Tukey HSD revealed the significant N400 effect between the standard condition and C1^{deviant} C2^{deviant} condition at the lateral electrodes, significant N400 effects were distributed in the anterior region (p = 0.05) (Fig. 7.7).

(a)



(b)

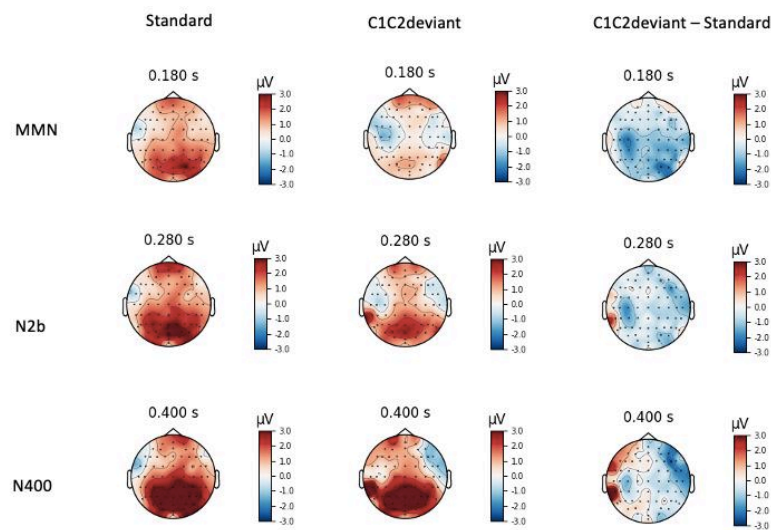


Fig.7.7. (a) Grand-averaged ERPs time-locked to the onset of stimuli in the two conditions (blue: standard; orange: C1 deviant C2 deviant). The time windows of MMN, N2b, and N400 are indicated by yellow, green, and blue. (b) Scalp maps show the topographies of standard condition and C1 deviant C2 deviant condition for the MMN, N2b, and N400 components.

7.6 Discussion

In this study, we addressed the question about the effect of incongruent metrical patterns on the auditory processing of two-character neutral tone compounds. A deviant metrical structure was created by lengthening the syllable duration of both two constituents within compound words. Participants listened to the sentences in which the standard or prosodically deviant stimuli were embedded. The behavioral results revealed no significant effect of incongruent prosody patterns on the reaction times in the task. The ERP data revealed that metrical violations, like the irregular syllable duration of constituents can hinder lexical processing. Our results show that the three ERP components of interest, the MMN, N2b, and N400, which correspond respectively with the time windows of 120-200ms, 220-300ms, and 350-550ms after the stimuli onset, were modulated by the metrically incongruous stimuli. Among the four categories of stimuli, the incongruous metrical pattern of the second neutral tone constituent elicited the most pronounced ERP effects compared with the other categories. Taken together, our results are consistent with previous studies, which indicate that syllable duration plays a role in lexical processing. Whereas, beyond the classical result that incongruent metric pattern hinders lexical access, here, we provide a new insight into the role of prosody in word recognition. Manipulation of the duration of neutral tone within the Mandarin compounds seems to be more disturbing than similar manipulation of full tone in the recognition task. The observation of the largest N400 response was in the stimuli which contained the irregular

second neutral tones, this result lends support to the conclusion that the prosodic pattern of the second neutral tone might be a determining prosodic cue in the Mandarin compound processing.

In the current research, the first major finding was to reveal the role of metric structure in the auditory processing of two-character neutral tone compounds. In our EEG results, concerning the two lengthened constituents of neutral tone compounds, we observe that lengthened second neutral tone elicited the more negative amplitudes of MMN and N2b components, with an anterior and central distribution, in relation to the stimuli with a lengthened first full tone (the effects were absent). Interestingly, when the syllabic lengthening concerned only the second neutral tones, the larger MMN and N2b effect were revealed in relation to the stimuli in which both two constituents within the compounds were lengthened. To our best knowledge, few studies have reported this observation regarding the metrical deviation of the second neutral tone in the auditory processing. The increased negative amplitude of MMN and N2b effects are taken as evidence for the automatic phonological processing of irregular metrical patterns, which could also be viewed as a reflection of the listener's sensitivity to the unusual, lengthened stimulus.

Previous studies called the early negativities in the time window of 120ms-200ms as MMN (mismatch negativity), it is a change-specific component and considered to represent phonological discrimination in the brain. The enhanced MMN effect was revealed to reflect the detection of the metrical errors, or rhythmic irregularities contained in the auditory inputs. The mismatch between the correct metrical pattern and prosodically deviant stimulus can induce the mismatch negativity (MMN). Moreover, in the N200 family, the N2b is another important component to examine the phonological activation and representation. Listeners may analyze the phonological forms of the incoming stimuli and create the memory traces as to recognize them. In our study, the observation of the N2b effect reflects the detection of the deviated syllable duration contained in the stimulus, which confirms the cognitive sensitivity of participants to the subtle metrical violations. When the prosodically incongruous stimuli were processed, additional cognitive resources would be required to access the prosodically deviant input. Our results are consistent with the previous works, which revealed that the small deviations contained in the rhythmic pattern can be detected by listener's brains (Rothermich et al., 2012; Rothermich et al., 2010; Schmidt-Kassow & Kotz, 2009).

The second major finding in this EEG experiment is that the incongruous metrical pattern of the second neutral tone can hinder the lexical-semantic processing of neutral tone compounds. As we know that N400 is a centroparietal negative-going component, its peak amplitude is at about 400ms after the deviant stimulus. According to previous studies, the N400 effect can be induced by a stimulus which is semantically inappropriate to the expected context (Hillyard & Kutas, 1983; Kutas & Federmeier, 2011). In this study, a typical N400 effect was induced by the incongruous

metrical structure of the second neutral tone within compound words, which showed a lateralized anterior topography, and the irregular longer duration of the second neutral tones elicited the larger N400 effect than the compounds containing two irregular longer syllables. Our results reveal that incongruous metrical pattern (lengthened syllable duration) of the neutral tone compounds could hinder the processing of the lexical-semantic properties of words. These findings are analogous to the previous studies in French (Magne et al., 2007; Marie et al., 2011), German (Friedrich, Kotz, Friederici, & Gunter, 2004; Knaus et al., 2007), English (Cooper et al., 2002) and Spanish (Soto-Faraco et al., 2001). The unnatural prosody pattern induced the increased N400 effect, which can be interpreted as a cost of more cognitive resources in lexical access. Two propositions may shed light on the difficulties induced by the metrically incongruous stimulus. Firstly, when the listeners encounter a metrically deviant stimulus in the online input, a recognition problem will be induced by the mismatches between the expectancy of metrically congruent neutral tone compounds and the deviated acoustic changes contained in the metrically stretched stimuli, the unexpected metrical structure thus can hinder the phonological representation in the mental lexicon. Previous studies revealed that the mismatch between the expectancy and the incongruous prosodic information in the input signal would cost additional cognitive resources due to rhythmic deviation (Bohn et al., 2013; Rothermich et al., 2012; Rothermich et al., 2010). Secondly, the deviated syllable duration of stimuli prevented the lexical retrieval of phonological representation, and more resources would be required in order to realize the semantic integration. Our finding is in line with previous studies and confirms the importance of syllable duration in the processing of spoken words (Magne et al., 2007; Wang et al., 2005). In addition, in our results, a late N400 effect was elicited by the stimulus with a deviant first full tone consistent, which peaks at around 550ms with an anterior distribution. This observation is consistent with previous studies such as Desroches et al. (2009), Dufour et al. (2013), and Brunellière et al. (2013), which have highlighted the sensitivity of a late N400 to the metrical incongruity in the auditory stimulus within the late time window of 520 - 600ms after the stimulus onset.

Moreover, in our result, a larger N400 effect was elicited when participants processed the stimuli with lengthened neutral tones within neutral tone compounds, in relation to stimuli with lengthened first full tones. It is noteworthy that the different mean amplitudes of the N400 effects indicated the degrees of lexical representation difficulties induced by the two constituents within neutral tone compounds. In relation to the stimuli containing an incongruous first full tone, the deviant rhythmic pattern of the second neutral tone elicited more negative amplitudes of the MMN and N2b effects in the fronto-central region, and more negative amplitudes of the N400 effect in the anterior region. This finding shed new light on the role of neutral tone and full tone syllables

in the compound processing. During the recognition of two-character neutral tone compounds, the deviant duration of the second neutral tone evoked more negativities in relation to the first full tone, both the early auditory negativities MMN and N2b, and the semantic N400 effect are larger, this suggests that metrical abnormality contained in the neutral tones could induce more difficulties in the semantic integration. Depending on the previous study, this result had never been observed, which could help us to account for the important role of neutral tone in auditory processing. Our study extends the findings that the larger amplitudes evoked by the lengthened neutral tone reflects the larger difficulty in the lexical access. Considering the main acoustic cues of neutral tone in the perception, Wang (2004) have claimed the crucial role of duration and pitch in the perception of Mandarin lexical tones. In our study, we considered that the different negative amplitudes of the MMN, N2b and N400, which were elicited by both the deviant neutral tone and full tone within the compounds, can shed light on this topic. Listeners are found to be more sensitive to detect the acoustic incongruities contained in the second neutral tones. In relation to the first full tones, the lengthened duration of neutral tones created a cognitive bias, that made the listeners deviate from the correct lexical representation and hinder the word recognition. Our result can complement the research about the functions of suprasegmental cues in lexical access, particularly for the function of syllable duration. The acoustic changes in the phonemes could modulate the neural correlates and elicited the N400 effect, this suggests that the second neutral tone constituent is activated in the representation procedure, and the violations in the sub-constituent's duration of neutral tones disrupt the semantic access. Overall, it suggests that neutral tone compound can be decomposed into their individual constituents in the auditory lexical access.

7.7 Conclusion

In summary, we observed that the incongruous metrical pattern of two-character neutral tone compounds has evoked the negative ERP effects of MMN, N2b, and N400, and the deviant duration of the second neutral tone elicited larger MMN, N2b and N400 effects, and the deviant duration of the first full tone elicited a late N400 effect. The observation of MMN and N2b effects revealed that deviant syllable duration, which is not congruent with the correct Mandarin prosody pattern, can be detected by native listeners, which can interrupt lexical access in the phonological representation. The dysfunction in the early stage of phonological processing, particularly when it concerns the lengthened neutral tone carried by the second constituent within compounds, might be reflected in the processing of the semantic properties of compounds, as revealed by the largest N400 effect induced by the stimuli containing the incongruous second neutral tone. These results suggest that native listeners are more sensitive to the acoustic changes in the second neutral tone

in relation to the first full tone in auditory processing of two-character neutral tone compounds, the suprasegmental information, particularly the syllable duration of neutral tones may contribute more to the lexical access.

8 General discussion and conclusion

The present dissertation analyzed the acoustic correlates of two-character compounds with a neutral tone in Mandarin, and electroencephalography (EEG) with event-related potentials was used to examine the native listeners' lexical processing of neutral tone compounds in the auditory modality. The main contribution of this research is three-fold. Firstly, we explored the acoustic characteristics of neutral tone compounds. Secondly, we examined whether two-character neutral tone compounds would be morphologically decomposed in the auditory processing, and how the lexical frequency and prosody pattern affect the auditory lexical access. Finally, we investigated to what extent the suprasegmental information (syllable duration) can modulate word recognition. Two categories of two-character compound words are concerned: (1) neutral tone compounds constituted the critical linguistic materials, a full tone syllable carried by its first constituent and a neutral tone syllable carried by its second constituent and form a trochaic prosody pattern [Strong-weak]; (2) full tone compounds are also involved, which have two full tones carried respectively by the first and the second constituent and form a spondee prosody pattern [Strong-strong].

In the current research, two types of methodologies were used, including (1) analysis of acoustic correlates, (2) Encephalography (EEG) technique, and in total four experiments were constructed. The two acoustic analyses concerned the phonetic feature of two-character neutral tone compound, in which we applied a multi-speaker production method and examined the three principal acoustic cues (duration, pitch, and intensity), concerning the second neutral tone syllable within compounds. Based on the recordings produced by the participants, we used the Praat and RStudio and analyzed the acoustic correlates of target syllables. In the first acoustic analysis, 160 (80 pairs) two-character neutral tone compounds [Strong-weak] and full tone compounds [Strong-strong] were chosen, for each minimal pair, their second syllables have the same syllable structure carrying either a neutral tone or a full tone, i.e., 巴掌 /ba1zhang0/ (clap) and 手掌 /shou3zhang3/ (hands), their second constituents have the identical syllabic structure /zhang/ and in either neutral tone (Tone 0) or full tone (Tone 3). We focused on the acoustic characteristics of the second neutral tone in contrast with the second full tone. Our results indicate two major findings. Firstly, the second neutral tone is revealed to be significantly reduced compared with the second full tone, and the intensity of the neutral tone is lower than full tone, except in the intrinsic tonal context of Tone 3. These results are consistent with the findings reported in the previous studies (Lin & Yan, 1990; Shen, 1992). In comparison with a full tone, the second neutral tone has a significantly lower pitch level in Tone 4 and a narrower pitch range in Tone 3 and Tone 4. Moreover, the pitch (F0) of the second neutral tone was revealed to co-vary with the preceding full tone, we observed that the pitch onset of the neutral tone is dependent on the pitch offset value of the preceding full

tone, except in the intrinsic tonal context of Tone 4. Consistent with the findings in the previous studies (Duanmu, 2007; Lee & Zee, 2008), this suggests the presence of a phonetic interaction between the first and the second syllables. Secondly, when the constituent is in a neutral tone (Tone 0), its full tone counterpart constituent may be in one of the four full tones (Tone 1/2/3/4), e.g., 彩/cai0/ within neutral tone compound 云彩/yun2cai0/(cloud) carries a Tone 0, and 彩/cai3/ carries a Tone 3 within full tone compound 色彩/se4cai3/(color); 家/jia0/ carries a Tone 0 within neutral tone compound 亲家/qin1jia0/(relatives by marriage), and 家/jia1/ carries a Tone 1 within full tone compound 回家/hui2jia1/ (go home). Despite the different intrinsic tone patterns of neutral tones, the duration and pitch of the second neutral tones within compounds reveal barely any significant differences across the four tones, suggesting that neutral tones have a shorter duration, lower pitch target, and narrower pitch range. In contrast, the intensity of neutral tone demonstrated significant differences across the four intrinsic full tones, the intensity of the second neutral tone with an intrinsic Tone 3 is significantly lower than the neutral tone in the other intrinsic tone patterns.

In the second acoustic analysis, we examined the coarticulation effect with respect to the second neutral tone within a two-character compound in order to investigate the extent to which the acoustic identity of neutral tones varies with the adjacent tonal contexts. Firstly, 80 neutral tone compounds in the four categories were selected. The second neutral tones were combined respectively with the four full tones based on the preceding tonal contexts. Moreover, 20 full tone compounds were chosen as the control group, and the carry-over effects from the first full tone to the phonological form of the second neutral tone were explored. In our results, significant carry-over effects from the preceding tonal contexts to the second neutral tones were observed, and the acoustic features of the second neutral tones were coextensive with the first full tones. Results revealed the main effects of the preceding tonal contexts on the duration and intensity of the second neutral tones, the dependent contextual characteristics were observed in the pitch range and pitch level of the second neutral tones across four full tones patterns, which confirms that the second neutral tones co-vary with the preceding tonal contexts (Tang, 2014). Compared with the full tone compounds in the control group, the second neutral tones were revealed to have a shorter duration with a lower pitch level and a narrower pitch range. But the pitch transition between the offset of the first full tone and the onset of the second neutral tone revealed barely any significant variation, suggesting that the pitch target of neutral tone was dependent on the preceding full tone. Secondly, we provided the detailed tonal variations of the first full tones under the influences of the following tonal contexts of the second constituents, and the same first full tone constituent was paired respectively with the second constituents in Tone 1/2/3/4 and in Tone 0, accordingly.

100 two-character compound words were selected, and the 20 first constituents were produced in isolation as a control group. In our results, barely any significant spreading influence is found from the rightmost weak neutral tone to its preceding strong full tone. As for the acoustic properties of the first full tones, no significant difference was found between the first constituent followed by a neutral tone or by any of the four full tones, no significant spreading influence was observed from the second syllable to the first full tone within two-character compounds. The only evidence of an anticipatory effect was observed in the pitch range of the first constituents, which was coextensive with the tonal target of the second constituents.

Two ERP experimentations were carried out to investigate to what extent the frequency effect, prosody pattern, and syllable length can modulate the auditory processing of two-character neutral tone compounds. The first EEG experiment found that the frequency of whole-compound and first constituent within two-character compounds modulated the priming effect of the first constituent. Moreover, we also found that the neutral tone works as a determining cue in auditory lexical processing. Concerning the role of syllabic duration in the lexical representation, the second EEG experiment revealed that the MMN, N2b, and N400 effects were elicited in response to the deviated syllabic durations of the two constituents within neutral tone compounds.

Semantic congruity effect: N400

In the first EEG study, by presenting the listeners with two-character spoken neutral tone compounds [Strong-weak] and full tone compounds [Strong-strong], which were manipulated in their whole-compound frequency and first-constituent frequency, participants' brain responses during the recognition of visual target were recorded and examined. Through the analysis of the N400 component, we found that the amplitude of the N400 effect was modulated by the frequency effects of the whole compound and first constituent, a relatively smaller priming effect was found in the rare compounds and the compounds with low-frequency first constituents. Moreover, the second neutral tone was revealed to be involved in the lexical access of compounds, and a smaller priming effect was observed in the neutral tone compounds with low-frequency first constituents, as reflected by increased N400 effect with a central distribution at the surface of the scalp. Because N400 component is a centro-parietally located negativity, which typically peaks at around 400 ms post-stimuli after the detection of a semantic incongruity in a sentence, the enhanced N400 effect could reflect the difficulty that occurred in the lexical-semantic integration procedure (Hagoort et al., 1993; Kutas & Hillyard, 1980). Our result is consistent with the previous study conducted by Wang et al. (2017), which claimed that due to a lower morpheme frequency, the representation of a single constituent within the multi-morphemic compounds consumes more cognitive resources to realize the lexical retrieval in the competition, which could be reflected by enhanced N400 effect

in the processing of compounds with low-frequency first constituents. It is noteworthy that a larger N400 effect was elicited in the lexical access of neutral tone compounds with the low-frequency first constituents. Due to the semantic integration dependence from the second neutral tone to the first full tone, lexical representation of the second neutral tone must be realized through the combination with the first full tone for the semantic integration in the auditory modality, when the first constituent is decomposed and activated, a global processing would be activated for the whole word form of neutral tone compound, resulting into an enhanced N400. Our finding is consistent with Luo et al. (2016), in which the more negative amplitudes of the N400 effect were revealed in the readings of neutral tone compounds, in relation to the readings of full tone compounds.

In the second EEG study, the metrical incongruity was created by lengthening the syllable duration for both two constituents within two-character neutral tone compounds, deviant metric structure of the second neutral tone elicited more negative N400 amplitudes with a fronto-centrally located negative deflection, the deviant metric structure of first full tone elicited a delayed N400 effect with the peak latency at around 550 ms after the stimuli onset and with a frontal distribution. Our data are consistent with previous studies suggesting that metrical information interplays with the lexico-semantic processing, and the metrical violation can disrupt the access to the semantic representation in the long-term memory (Bohn et al., 2013; Magne et al., 2007; Marie et al., 2011; Rothermich et al., 2012; Rothermich et al., 2010). In our study, the incongruous metric structure hinders the semantic representation in the online processing, additional cognitive resources were consumed due to the deviant metrical pattern in the lexical retrieval procedure, as indexed by the enhanced N400 effect. Accordingly, our results indicate that suprasegmental information has been encoded at the lexical level, which takes part in the access. Listener's speech comprehension will be impaired, when an analysis of an unexpected metrical template impacts the semantic integration procedure. This finding is analogous to the studies of Soto-Faraco et al. (2001) and Cutler (2005), in which the important role of suprasegmental information in the lexical access is indicated in the word recognition. When the inputs are mapped onto a sequence of phonemic items, the duration of phonemes could facilitate the phonological activation in the competition procedure. Another important finding in this research is the late latency of the N400 effect which was elicited by the metrical deviations of the first full tones. As reported by Brunellière et al. (2013) and Dufour et al. (2013), the metrically deviant stimulus could disrupt the semantic integration procedure and elicit an increased late N400 effect.

Exogenous metrical effect: N2b

In the second EEG experiment, we created the incongruous metrical structure for the two-character neutral tone compounds by lengthening the syllable duration of both two constituents

(the first full tone and the second neutral tone). The N2b effect was found to be larger in response to the acoustic changes in the second neutral tones across the fronto-central sites (Fz - Cz). Within the neutral tone compounds, the deviant duration of the second neutral tone has elicited the largest negative amplitudes of N2b in relation to the stimuli with two deviant constituents. Interestingly, the deviations of the first full tone did not elicit a significant modulation of the N2b amplitudes. Because N2b is an ERP component that is thought to mark the acoustic changes and the access to the phonological representation in word recognition, the observation of N2b effect reflects the detection of the irregular phonetic changes contained in the online inputs, the larger N2b effects elicited by the second neutral tones are interpreted as reflecting the processing difficulty to access/activate the appropriate phonological form of compounds. When listeners encountered the deviant stimuli with a lengthened neutral tone, some difficulties were cropping up during the creation of a correct echoic memory trace as to realize the matched phonological representation in the mental lexicon (Pulvermüller, 1996; Pulvermüller & Schumann, 1994). Our results are consistent with the previous works of Näätänen (1982) and Näätänen and Gaillard (1983), in which the N2b effect was indicated to be associated with the detection of pre-perceptual acoustic changes, which can be related with involuntary attention switch caused by the unexpected changes. In our study, the N2b effect was observed on the stimuli with the deviant second neutral tones or with two deviant tones, and not on the stimuli possessing the deviant first full tones. This difference was considered to reflect the listener's sensitivity to the sound changes contained in the second neutral tone, while the metrical violation contained in the first full tone does not seem to have as much of an impact on the access to the phonological representation of compounds. Critically, our results showed that native Mandarin listeners were more sensitive to the phonetic deviations in the neutral tones than in the full tones in auditory lexical access.

Mismatch negativity effect: MMN

In the second EEG experiment, the Mismatch negativity effect (MMN) was elicited by the deviations in the second neutral tones and the deviations in both two tones within two-character neutral tone compounds. It is noteworthy that different negative amplitudes of MMN effects were observed, i.e., the deviant second neutral tones elicited larger MMN effects than the stimuli with two deviant tones. Moreover, no MMN effect was observed in response to the processing of the compounds with deviant first full tones. As we know that MMN component is a change-specific ERP component, which is considered to represent the neutral processes of the automatic auditory perception, its peak latency is usually within the time window of 100 to 250ms from the stimulus onset with a frontocentral topography, the MMN component has been identified as a detection of the metrical error or the rhythmic irregularity contained in the auditory inputs (Escera et al., 1998;

Giard et al., 1990; Näätänen et al., 1978; Näätänen & Michie, 1979; Rinne et al., 2000). After the inputs of incoming stimulus, listeners would detect the acoustic changes contained in the signals, some difficulties were triggered by the unmatched deviant metric structure of stimulus in relation to the correct prosody patterns in the mental storage. Consequently, to recognize and categorize the deviant metrical patterns, additional costs in the phonological representation would evoke an increased MMN effect. MMN effect found in this study are in line with the previous studies, the changes in the syllable duration could evoke the MMN effects in the recognition of the disyllabic compounds (Tervaniemi et al., 2006). Furthermore, in our result, within two-character neutral tone compounds, the two constituents had elicited different amplitudes of mismatch negativity effects, as indexed by the amplitudes of the MMN component, the deviant metrical pattern of the second neutral tones elicited larger MMN effects than the deviant metric pattern of the first full tones. In spoken word recognition, native listeners don't discriminate the unmatched deviant metric pattern of the first full tones, their brains didn't display the significant MMN responses to the durational changes of the first full tone.

In general, our studies disentangle the role of lexical frequency effect and prosody pattern in Mandarin neutral tone compound's processing. In the first EEG study, a significant repetition priming effect of the first constituent was observed in the cross-modal priming experiment. The reduced priming effect was observed in rare compounds and compounds with low-frequency first constituents, and no similar effects were observed in the frequent compounds and the compounds with high-frequency first constituents. These findings suggest that the frequent compounds and the compounds with high-frequency first constituents are turned to be represented in their whole-form, whereas the rare compounds and the compounds with low-frequency first constituents might be represented and processed via their individual constituent. In addition, we also provide evidence of prosodic processing related to the second neutral tones, a reduced priming effect of the first constituent was revealed in neutral tone compounds with low-frequency first constituents, and no such effect was revealed in full tone compounds. When the low-frequency first constituent is decomposed and activated, because the representation of the second neutral tone constituents must be realized in combination with first full tone constituents in the auditory modality, additional cost in the cognitive system would be consumed to recognize the whole word form of neutral tone compound after the decomposition. Thus, we consider that suprasegmental information contained in the individual constituent is stored in the mental lexicon of the native listeners, which could be activated in the representation. Overall, considering the trochaic prosody structure [Strong-weak] formed by the first full tone and the second neutral tone, our results support the prediction that the second neutral tone plays a role in the access. In auditory modality, when the first constituent

is activated, and the neutral tone compound would be accessed via their individual constituents, there will be a secondary holistic processing for the whole compound's form to realize the semantic representation, and a global processing must be boosted by combining both two constituents, this additional procedure in the neutral tone compound's processing will consume more as reflected by a larger N400 effect.

In the second EEG experiment, when the stimuli were presented without their expected metrical templates, participants encountered the semantic integration difficulty due to the metrical deviation. Considering the two tones within neutral tone compound, native listeners demonstrated a decreased sensitivity to the incongruous metrical structure of the first full tone, and an increased sensitivity to the second neutral tone. Lexical access was significantly inhibited by the lengthened duration of the second neutral tones, in relation to first full tones. This suggests that listeners relied more on the suprasegmental information of neutral tones in the spoken word recognition, within neutral tone compounds, the syllable duration of the second neutral tone contributes more to the lexical representation. Depending on previous studies about the perception cues of the Mandarin lexical tones, the pitch (F0) was claimed as the primary suprasegmental cue, and syllable duration was considered as the secondary cue (Howie, 1976; Lai et al., 2010; Liu & Samuel, 2004; Shen, 1993). However, concerning the acoustic characteristics of the Mandarin stressed and unstressed syllables, the syllable duration was claimed as the reliable and primary cue to the contrastive stress patterns (Duanmu, 2007). In the current thesis, we found that within two-character compounds, the syllable duration of neutral tone could hinder the lexical representation, however, the full tone can still be recognized despite the incongruent metrical templates. When the duration of individual constituent was changed, the lexical activation will be disrupted, which has hindered the retrieval of the potential candidates in the lexical storage. Our findings showed that the second neutral tone was decomposed and activated in the representation stage, and the acoustic changes contained in the phonemic units will interrupt the semantic accessing of the whole compounds, which can be evidence of a decomposition processing of neutral tone compounds.

Taken together, this doctoral dissertation investigated the acoustic correlates of the two-character neutral tone compounds. Based on the neurophysiological results, our study shed new light on the lexical representation of the Mandarin compounds which consist of two characters. In future studies, some more research will be required to better understand the neural mechanisms underlying the processing of neutral tone compounds. Concerning the lexical frequency effects, a study concerning the interactions between whole-compound frequency and individual constituent frequency of neutral tone compound will be required, due to the lack of evidence of the interaction effect in the current study. Moreover, among the main suprasegmental cues, in this dissertation,

we only explored the contribution of syllable duration in lexical access of neutral tone compounds, a supplementary investigation could be constructed to explore the other suprasegmental cues, such as the deviant pitch and intensity in the compound word processing.

Appendices

Appendices 1⁴

No.	Neutral tone compound	Frequency	Occurrence Times	Stroke number	No.	Full tone compound	Frequency	Occurrence Times	Stroke number
1	畜生 /chu 4 sheng 0/	0.0005	50	15	1	诞生 /dan 4 sheng 1/	0.0043	407	13
2	下巴 /xia 4 ba 0/	0.0008	80	7	2	嘴巴 /zui3 ba 1/	0.0033	314	20
3	功夫 /gong 1 fu 0/	0.0029	273	9	3	丈夫 /zhang4 fu 1/	0.0101	961	7
4	亲家 /qin 4 jia 0/	0.0006	60	19	4	作家 /zuo 4 jia 1/	0.0138	1323	17
5	秧歌 /yang 1 ge 0/	0.0005	51	24	5	唱歌 /chang 4 ge 1/	0.0021	205	25
6	吆喝 /yao 1 he 0/	0.0008	79	18	6	吆喝 /chi 1 he 1/	0.0006	58	18
7	意思 /yi 4 si0/	0.0142	1358	22	7	构思 /gou4 si 1/	0.002	193	17
8	委屈 /wei 3 qu 0/	0.0014	133	16	8	冤屈 /yuan 1 qu 1/	0.0005	50	18
9	冒失 /mao 1 shi 0/	0.0005	50	15	9	过失 /guo 4 shi 1/	0.0009	83	11
10	比方 /bi 3 fang 0/	0.0009	85	8	10	立方 /li 4 fang 1/	0.0006	60	9
11	晃悠 /huang 4 you 0/	0.0005	50	19	11	忽悠 /hu 1 you 1/	0.0005	50	19
12	官司 /guan 1 si0/	0.0005	50	13	12	上司 /shang 4 si 1/	0.0005	52	8
13	打发 /da 3 fa 0/	0.001	98	10	13	批发 /pi 1 fa 1/	0.0037	353	12
14	耷拉 /da 1 la 0/	0.0006	53	17	14	拖拉 /tuo 1 la 1/	0.0005	50	16
15	嫁妆 /jia 4 zhuang 0/	0.0005	50	19	15	化妆 /hua 4 zhuang 1/	0.0005	50	10
16	消息 /xiao 1 xi 0/	0.0098	936	20	16	栖息 /qi 1 xi 1/	0.0009	89	20
17	招呼 /zhao 1 hu 0/	0.0037	355	16	17	欢呼 /huan 1 hu 1/	0.0012	111	14
18	答应 /da 1 ying 0/	0.0067	641	19	18	顺应 /shun4 ying 1/	0.0008	80	16
19	称呼 /cheng 1 hu 0/	0.0012	112	18	19	高呼 /gao1 hu 1/	0.0006	57	18
20	弟兄 /di4 xiong0/	0.0024	230	12	20	老兄 /lao3 xiong1/	0.0009	89	11
21	凑合 /cou1 he0/	0.0005	50	17	21	集合 /ji2 he2/	0.0019	183	18
22	眉毛 /mei2 mao0/	0.0017	166	13	22	羽毛 /yu3 mao2/	0.0021	197	10
23	苍蝇 /cang1 ying0/	0.0013	121	21	23	蚊蝇 /wen2 ying2/	0.0005	50	24
24	锄头 /chu2 tou0/	0.0008	80	17	24	矛头 /mao2 tou2/	0.0005	52	10
25	棒槌 /bang4 chui0/	0.0006	62	25	25	鼓槌 /gu3 chui2/	0.0005	50	26
26	灯笼 /deng1 long0/	0.0009	85	17	26	牢笼 /lao2 long2/	0.0005	50	18
27	胳膊 /ge1 bo0/	0.0026	249	24	27	赤膊 /chi4 bo2/	0.0005	50	21
28	姑娘 /gu1 niang0/	0.0192	1834	18	28	新娘 /xin1 niang2/	0.0016	156	23
29	便宜 /pian2 yi0/	0.0022	211	17	29	适宜 /shi4 yi2/	0.0044	424	17
30	合同 /he2 tong0/	0.0085	810	12	30	陪同 /pei2 tong2/	0.0011	107	16
31	快活 /kuai4 huo0/	0.0015	144	16	31	生活 /sheng1 huo2/	0.091	8694	14
32	糊涂 /hu2 tu0/	0.0023	218	25	32	滩涂 /tan1 tu2/	0.0005	50	23
33	机灵 /ji1 ling0/	0.001	92	13	33	心灵 /xin1 ling2/	0.005	479	11
34	见识 /jian4 shi0/	0.0009	84	11	34	常识 /chang2 shi2/	0.0015	145	18
35	交情 /jiao1 qing0/	0.0005	50	17	35	表情 /biao1 qing2/	0.0049	468	19
36	豁达 /liu1 da0/	0.0005	50	19	36	发达 /fa1 da2/	0.0116	1113	11
37	麻烦 /ma2 fan0/	0.0022	210	21	37	厌烦 /yan4 fan2/	0.0005	50	16
38	奴才 /nu2 cai0/	0.0006	56	8	38	人才 /ren2 cai2/	0.0115	1096	5
39	暖和 /nuan3 he0/	0.0008	72	21	39	温和 /wen1 he2/	0.0017	160	20
40	招牌 /zhaol pai0/	0.0009	89	20	40	名牌 /ming2 pai2/	0.0007	71	18

⁴ **Frequency:** the word's frequency is calculated by counting the occurrence times per 10 million in the Corpus of Contemporary Mandarin; **Stroke number:** the number of strokes contained in the two characters of a compound.

No.	Neutral tone compound	Frequency	Occurrence Times	Stroke number	No.	Full tone compound	Frequency	Occurrence Times	Stroke number
41	下场/xia4 chang0/	0.0009	82	9	41	市场/shi4 chang3/	0.0427	4082	11
42	财主/cai2 zhu0/	0.0007	65	12	42	民主/min2 zhu3/	0.0131	1249	10
43	耳朵/er3 duo0/	0.0055	524	12	43	花朵/hua1 duo3/	0.0018	175	13
44	闺女/gui1 nv0/	0.0013	129	12	44	子女/zi3 nv3/	0.006	576	6
45	家伙/jia1 huo0/	0.0043	413	16	45	合伙/he2 huo3/	0.0005	51	12
46	戒指/jie4 zhi0/	0.0006	60	16	46	手指/shou3 zhi3/	0.0036	348	13
47	利索/li4 suo0/	0.0005	50	17	47	探索/tan4 suo3/	0.0076	724	21
48	马虎/ma3 hu0/	0.0006	58	11	48	老虎/lao3 hu3/	0.0025	237	14
49	朋友/peng2 you0/	0.0143	1365	12	49	战友/zhan4 you3/	0.0022	215	13
50	屁股/pi4 gu0/	0.0023	223	15	50	持股/chi2 gu3/	0.0005	50	15
51	清楚/qing1 chu0/	0.0164	1571	24	51	酸楚/suan1 chu3/	0.0005	50	27
52	牲口/sheng1 kou0/	0.0019	184	12	52	出口/chu1 kou3/	0.0066	634	8
53	舒坦/shu1 tan0/	0.0005	50	20	53	平坦/ping2 tan3/	0.0013	123	13
54	挖苦/wa1 ku0/	0.0005	50	17	54	吃苦/chi1 ku3/	0.0009	86	14
55	胭脂/yan1 zhi0/	0.0005	50	20	55	油脂/you2 zhi3/	0.0008	77	18
56	云彩/yun2 cai0/	0.0006	57	15	56	色彩/se4 cai3/	0.009	864	17
57	指甲/zhi3 jia0/	0.0007	71	14	57	盔甲/kui1 jia3/	0.0005	50	12
58	豆腐/dou4 fu0/	0.0023	218	21	58	贪腐/tan1 fu3/	0.0005	50	21
59	烧饼/shao1 bing0/	0.0005	50	19	59	糕饼/gao1 bing3/	0.0005	50	19
60	规矩/gui1 ju0/	0.0017	161	17	60	逾矩/yu4 ju3/	0.0005	50	21
61	对付/dui4 fu0/	0.0032	308	10	61	支付/zhi1 fu4/	0.003	282	9
62	部分/bu4 fen0/	0.0348	3328	14	62	成分/cheng2 fen4/	0.0128	1222	8
63	动静/dong4 jing0/	0.0012	117	20	63	平静/ping2 jing4/	0.0046	437	19
64	记性/ji4 xing0/	0.0005	50	13	64	个性/ge4 xing4/	0.0086	820	11
65	聚餐/lei2 zhui0/	0.0005	50	25	65	招赘/zhaol zhui4/	0.0005	50	22
66	叫唤/jiao4 huan0/	0.0006	57	15	66	召唤/zhaol huan4/	0.0007	66	15
67	和尚/he2 shang0/	0.0028	266	18	67	高尚/gaol shang4/	0.0027	261	18
68	客气/ke4 qi0/	0.0015	146	13	68	生气/sheng1 qi4/	0.0045	426	9
69	漂亮/piao4 liang0/	0.0056	536	23	69	闪亮/shan3 liang4/	0.0008	75	14
70	商量/shang1 liang0/	0.0049	467	23	70	质量/zhi4 liang4/	0.0292	2794	20
71	唾沫/tuo4 mo0/	0.0007	71	19	71	泡沫/pao4 mo4/	0.0007	71	16
72	主意/zhu3 yi0/	0.0039	373	18	72	乐意/le4 yi4/	0.001	91	18
73	头发/tou2 fa0/	0.0069	660	10	73	毛发/mao2 fa4/	0.0007	63	9
74	特务/te4 wu0/	0.0034	326	15	74	任务/ren4 wu4/	0.0335	3204	11
75	算计/suan4 ji0/	0.0005	50	18	75	设计/she4 ji4/	0.0127	1211	10
76	欺负/qi1 fu0/	0.001	92	18	76	辜负/gu1 fu4/	0.0006	62	15
77	厉害/li4 hai0/	0.0043	414	15	77	迫害/po4 hai4/	0.0019	182	18
78	凉快/liang2 kuai0/	0.0005	50	17	78	飞快/fei1 kuai4/	0.0011	109	10
79	差事/chai1 shi0/	0.0005	50	17	79	本事/ben3 shi4/	0.0017	166	17
80	买卖/mai3 mai0/	0.0037	353	14	80	出卖/chu1 mai4/	0.0023	216	13

Appendices 2-1

No.	Neutral tone compound	Frequency	Occurrence Times	Stroke Number	No.	Neutral tone compound	Frequency	Occurrence Times	Stroke Number
1	稀罕/xi1 han0/	0.0005	50	19	41	里头/li3 tou0/	0.0013	125	12
2	疙瘩/ge1 da0/	0.0013	126	22	42	喇叭/la3 ba0/	0.0017	160	17
3	窟窿/ku1 long0/	0.0006	62	29	43	使唤/shi3 huan0/	0.0005	50	18
4	甘蔗/gan1 zhe0/	0.0009	86	19	44	爽快/shuang3 kuai0/	0.0006	57	18
5	嘟囔/du1 nang0/	0.0005	50	38	45	尾巴/wei3 ba0/	0.0053	505	11
6	舒服/shu1 fu0/	0.0032	306	20	46	稳当/wen3 dang0/	0.0005	50	20
7	巴结/ba1 jie0/	0.0005	50	15	47	哑巴/ya3 ba0/	0.0005	50	13
8	思量/si1 liang0/	0.0005	50	21	48	枕头/zhen3 tou0/	0.0014	130	13
9	丫头/ya1 tou0/	0.001	93	8	49	嘴巴/zui3 ba0/	0.0033	314	20
10	衣服/yi1 fu0/	0.0124	1188	14	50	眨巴/zha3 ba0/	0.0005	51	13
11	冤枉/yuan1 wang0/	0.0005	51	18	51	养活/yang3 huo0/	0.0011	103	18
12	扎实/zha1 shi0/	0.0007	64	12	52	晚上/wan3 shang0/	0.0135	1288	14
13	将就/jiang1 jiu0/	0.0005	50	15	53	女婿/nv3 xu0/	0.0013	125	15
14	折腾/zhe1 teng0/	0.001	91	20	54	老实/lao3 shi0/	0.002	188	14
15	芝麻/zhi1 ma0/	0.0005	51	17	55	喇嘛/la3 ma0/	0.0005	50	26
16	庄稼/zhua1 jia0/	0.0024	234	21	56	火候/huo3 hou0/	0.0005	50	14
17	作坊/zuo1 fang0/	0.0008	76	14	57	扁担/bian3 dan0/	0.0008	79	17
18	结巴/jie1 ba0/	0.0005	50	13	58	伙计/huo3 ji0/	0.0012	111	10
19	街坊/jie1 fang0/	0.0005	50	20	59	打量/da3 liang0/	0.0022	208	17
20	妖精/yao1 jing0/	0.0005	50	21	60	指头/zhi3 tou0/	0.0012	114	14
21	蛤蟆/ha2 ma0/	0.0006	56	28	61	骆驼/luo4 tuo0/	0.0017	164	17
22	萝卜/luo2 bo0/	0.0009	90	13	62	罐头/guan4 tou0/	0.0008	81	28
23	狐狸/hu2 li0/	0.001	98	18	63	应酬/ying4 chou0/	0.0005	50	20
24	咳嗽/ke2 sou0/	0.0026	252	23	64	扫帚/sao4 zhou0/	0.0005	50	14
25	模糊/mo2 hu0/	0.0039	369	29	65	事情/shi4 qing0/	0.0234	2235	19
26	忙活/mang2 huo0/	0.0005	50	13	66	厚道/hou4 dao0/	0.0005	50	21
27	脾气/pi2 qi0/	0.0032	307	16	67	吓唬/xia4 hu0/	0.0005	52	17
28	拳头/quan2 tou0/	0.0022	208	15	68	相声/xiang4 sheng0/	0.0006	61	16
29	粮食/liang2 shi0/	0.013	1242	19	69	别扭/bie4 niu0/	0.0006	62	14
30	媳妇/xi2 fu0/	0.0049	469	19	70	钥匙/yao4 shi0/	0.0017	167	20
31	门道/men2 dao0/	0.0005	50	15	71	痢疾/li4 ji0/	0.0005	50	22
32	衙门/ya2 men0/	0.0009	89	16	72	月饼/yue4 bing0/	0.0005	50	13
33	琢磨/zuo2 mo0/	0.0015	143	28	73	状元/zhua1 yuan0/	0.0006	59	11
34	时候/shi2 hou0/	0.068	6500	17	74	壮实/zhua1 shi0/	0.0005	50	14
35	能耐/neng2 nai0/	0.0007	65	17	75	转悠/zhuan4 you0/	0.0005	50	19
36	石头/shi2 tou0/	0.0072	687	10	76	栅栏/zha4 lan0/	0.0005	50	18
37	拾掇/shi2 duo0/	0.0005	50	20	77	帐篷/zhang4 peng0/	0.002	192	23
38	舌头/she2 tou0/	0.0024	228	11	78	木头/mu4 tou0/	0.0018	174	9
39	馒头/man2 tou0/	0.0014	132	19	79	对头/dui4 tou0/	0.0005	50	10
40	京快/liang2 kuai0/	0.0005	50	17	80	晃悠/huang4 you0/	0.0005	50	21

Full tone compound	Frequency	Occurrence Times	Stroke number
歌声/ge1 sheng1/	0.0026	245	21
剪刀/jian3 dao1/	0.0005	51	13
猖狂/chang1 kuang2/	0.0005	51	18
军事/jun1 shi4/	0.0159	1522	14
凑巧/cou4 qiao3/	0.0005	51	16
龙头/long2 tou2/	0.0008	74	16
香火/zhao2 huo3/	0.0008	72	15
时髦/shi2 mao2/	0.0013	122	21
酒精/jiu3 jing1/	0.0027	255	24
眼前/yan3 qian2/	0.0074	704	20
甜蜜/tian2 mi4/	0.001	94	25
安静/an1 jing4/	0.0028	272	20
气息/qi4 xi1/	0.0018	169	14
制服/zhi4 fu2/	0.0005	52	16
慈祥/ci2 xiang2/	0.0005	51	23
设置/she4 zhi4/	0.0038	365	19
吵闹/cha3 nao4/	0.0006	56	15
大娘/da4 niang2/	0.0039	374	13
冰箱/bing1 xiang1/	0.0005	51	21
废物/fei4 wu4/	0.0012	115	16

Appendices 2-2

No.	Compound	Frequency	Occurrence Times	Stroke Number	No.	Compound	Frequency	Occurrence Times	Stroke Number
1	窗花 /chuang1 hua1/	0.0005	50	19	26	跳出 /tiao4 chu1/	0.0012	113	18
2	窗前 /chuang1 qian2/	0.0009	89	20	27	跳槽 /tiao4 cao2/	0.0005	50	28
3	窗口 /chuang1 kou3/	0.0024	228	15	28	跳舞 /tiao4 wu3/	0.0023	224	27
4	窗外 /chuang1 wai4/	0.0032	303	17	29	跳跃 /tiao4 yue4/	0.002	190	24
5	窗户 /chuang1 hu0/	0.0024	233	15	30	跳蚤 /tiao4 zao0/	0.0005	50	22
6	笑声 /xiao4 sheng1/	0.0025	235	17	31	认真 /ren4 zhen1/	0.0143	1363	9
7	笑容 /xiao4 rong2/	0.0028	271	20	32	认为 /ren4 wei2/	0.0599	5720	8
8	笑脸 /xiao4 lian3/	0.0012	112	19	33	认可 /ren4 ke3/	0.0012	112	9
9	笑料 /xiao4 liao4/	0.0005	50	17	34	认定 /ren4 ding4/	0.0019	183	17
10	笑话 /xiao4 hua0/	0.0018	168	18	35	认识 /ren4 shi0/	0.0434	4151	11
11	本身 /ben3 shen1/	0.018	1720	12	36	骨灰 /gu3 hui1/	0.0005	50	15
12	本来 /ben3 lai2/	0.011	1055	12	37	骨骼 /gu3 ge2/	0.0014	131	24
13	本体 /ben3 ti3/	0.0012	116	12	38	骨血 /gu3 xue3/	0.0005	50	15
14	本位 /ben3 wei4/	0.0008	79	12	39	骨干 /gu3 gan4/	0.0029	280	12
15	本事 /ben3 shi0/	0.0017	166	13	40	骨头 /gu3 tou0/	0.0018	171	14
16	收割 /shou1 ge1/	0.0009	86	18	41	名声 /ming1 sheng1/	0.0007	63	13
17	收回 /shou1 hui2/	0.0024	229	16	42	名人 /ming2 ren2/	0.0011	107	8
18	收买 /shou1 mai3/	0.001	97	22	43	名帖 /ming2 tie3/	0.0005	50	14
19	收购 /shou1 gou4/	0.005	481	15	44	名胜 /ming2 sheng4/	0.0006	53	15
20	收拾 /shou1 shi0/	0.003	291	15	45	名堂 /ming2 tang0/	0.0006	58	17
21	地方 /di4 fang1/	0.0532	5081	10	46	大厅 /da4 ting1/	0.0017	160	7
22	地球 /di4 qiu2/	0.0216	2062	17	47	大型 /da4 xing2/	0.0048	459	12
23	地理 /di4 li3/	0.0052	493	17	48	大小 /da4 xiao3/	0.0152	1451	13
24	地位 /di4 wei4/	0.0283	2702	9	49	大众 /da4 zhong4/	0.002	188	9
25	地道 /di4 dao0/	0.0006	55	18	50	大意 /da4 yi0/	0.0005	50	16

No.	Compound	Frequency	Occurrence Times	Stroke Number	No.	Compound	Frequency	Occurrence Times	Stroke Number
51	东方/dong1 fang1/	0.0037	352	9	76	石灰/shi2 hui1/	0.0016	155	11
52	东洋/dong1 yang2/	0.0008	72	14	77	石油/shi2 you2/	0.0099	943	13
53	东北/dong1 bei3/	0.0052	497	10	78	石板/shi2 ban3/	0.0009	88	15
54	东面/dong1 mian4/	0.0006	59	15	79	石块/shi2 kuai4/	0.0016	152	12
55	东西/dong1 xi0/	0.0467	4464	11	80	石榴/shi2 liu0/	0.0005	50	19
56	巴西/ba1 xi1/	0.0014	132	10	81	明星/ming2 xing1/	0.0009	90	17
57	巴黎/ba1 li2/	0.0042	401	17	82	明年/ming2 nian2/	0.0024	226	14
58	巴蜀/ba1 shu3/	0.0005	50	19	83	明显/ming2 xian3/	0.0183	1751	17
59	巴望/ba1 wang4/	0.0005	50	17	84	明快/ming2 kuai4/	0.0006	62	18
60	巴掌/ba1 zhang0/	0.0009	85	16	85	明白/ming2 bai0/	0.0083	796	13
61	含羞/han2 xiu1/	0.0005	50	17	86	衣衫/yi1 shan1/	0.0007	64	14
62	含情/han2 qing2/	0.0005	50	18	87	衣着/yi1 zhuo2/	0.0008	76	22
63	含有/han2 you3/	0.0104	990	13	88	衣领/yi1 ling3/	0.0005	50	17
64	含笑/han2 xiao4/	0.0007	63	17	89	衣物/yi1 wu4/	0.0006	59	14
65	含糊/han2 hu0/	0.0009	85	22	90	衣裳/yi1 shang0/	0.0022	210	20
66	张开/zhang1 kai1/	0.0023	216	11	91	行车/xing2 che1/	0.0006	53	13
67	张弛/zhang1 chi2/	0.0005	50	13	92	行为/xing2 wei2/	0.0329	3145	10
68	张口/zhang1 kou3/	0.001	92	13	93	行驶/xing2 shi3/	0.0017	165	14
69	张力/zhang1 li4/	0.0007	68	9	94	行动/xing2 dong4/	0.015	1437	12
70	张罗/zhang1 luo0/	0.0005	50	15	95	行李/xing2 li0/	0.0017	161	13
71	裁军/cai2 jun1/	0.0005	50	18	96	念书/nian4 shu1/	0.0011	107	12
72	裁员/cai2 yuan2/	0.0005	50	19	97	念白/nian4 bai2/	0.0005	50	13
73	裁剪/cai2 jian3/	0.0005	50	23	98	念想/nian4 xiang3/	0.0005	50	21
74	裁判/cai2 pan4/	0.0009	88	19	99	念旧/nian4 jiu4/	0.0005	50	13
75	裁缝/cai2 feng0/	0.0005	50	25	100	念头/nian4 tou0/	0.0019	180	13

No.	Firstconstituent(C1)	Frequency	Occurrence Times	Stroke number	No.	Firstconstituent(C1)	Frequency	Occurrence Times	Stroke number
1	窗/chuang1/	0.0026	253	12	11	东/dong1/	0.0114	1091	5
2	笑/xiao4/	0.0415	363	10	12	巴/ba1/	0.0007	66	4
3	本/ben3/	0.0246	351	5	13	含/han2/	0.0103	982	7
4	收/shou1/	0.0086	820	6	14	张/zhang1/	0.0231	227	7
5	地/di4/	0.4127	390	6	15	裁/cai2/	0.0007	64	12
6	跳/tiao4/	0.0121	161	13	16	石/shi2/	0.0077	736	5
7	认/ren4/	0.003	286	4	17	明/ming2/	0.0039	373	8
8	骨/gu3/	0.0051	485	9	18	衣/yi1/	0.005	482	6
9	名/ming2/	0.0206	170	6	19	行/xing2/	0.019	1812	6
10	大/da4/	0.2098	250	4	20	念/nian4/	0.0076	726	8

Appendices 3-1: related pairs⁵

No.	Category	Prosody	Prime	Frequency	Target	Frequency	No.	Category	Prosody	Prime	Frequency	Target	Frequency
1	(1)Frequent_SW_C1-frequent	SW	意思 Yi4si0 Meaning	1274	意	708	21	(3)Rare_SW_C1-frequent	SW	大意 Da4yi0 Careless	50+	大	20907
2	(1)Frequent_SW_C1-frequent	SW	部分 Bu4fen4 Part	2086	部	2619	22	(3)Rare_SW_C1-frequent	SW	结巴 Jie1ba0 Stammer	50+	结	461
3	(1)Frequent_SW_C1-frequent	SW	粮食 Liang2shi0 Cereals	1076	粮	411	23	(3)Rare_SW_C1-frequent	SW	见识 jian4 shi0 Experience	96	见	4823
4	(1)Frequent_SW_C1-frequent	SW	合同 He2tong0 Contract	649	合	858	24	(3)Rare_SW_C1-frequent	SW	将就 Jiang1jiu0 Put up with	50+	将	9315
5	(1)Frequent_SW_C1-frequent	SW	衣服 Yi1fu0 Clothes	1122	衣	459	25	(3)Rare_SW_C1-frequent	SW	岁数 sui4 shu0 age	73	岁	2993
6	(1)Frequent_SW_C1-frequent	SW	时候 Shi2hou0 Time	6721	时	18257	26	(3)Rare_SW_C1-frequent	SW	含糊 Han2hu0 Vague	86	含	868
7	(1)Frequent_SW_C1-frequent	SW	事情 Shi4qing0 Thing	2163	事	5775	27	(3)Rare_SW_C1-frequent	SW	能耐 Nen2nai0 Capacity	62	能	14413
8	(1)Frequent_SW_C1-frequent	SW	耳朵 Er2duo0 Ear	506	耳	383	28	(3)Rare_SW_C1-frequent	SW	使唤 Shi2huan0 Order somebody	50+	使	16571
9	(1)Frequent_SW_C1-frequent	SW	石头 Shi2tou0 Stone	635	石	818	29	(3)Rare_SW_C1-frequent	SW	记忆 Ji4xing0 memory	50+	记	1032
10	(1)Frequent_SW_C1-frequent	SW	脑袋 Kou3dai0 Pocket	288	脑	2045	30	(3)Rare_SW_C1-frequent	SW	凑合 cou4 he0 not too bad	50+	凑	180
11	(2)Frequent_SW_C1-rare	SW	姑娘 Gu1niang0 Girl	1752	姑	74	31	(4)Rare_SW_C1-rare	SW	巴掌 Ba1zhang0 Hand	83	巴	102
12	(2)Frequent_SW_C1-rare	SW	消息 Xiao1xi0 Information	906	消	203	32	(4)Rare_SW_C1-rare	SW	裁缝 Cai2feng0 Tailor	50	裁	61
13	(2)Frequent_SW_C1-rare	SW	朋友 Peng2you0 Friend	1310	朋	50+	33	(4)Rare_SW_C1-rare	SW	溜达 Liu1da0 Hang around	50+	溜	206
14	(2)Frequent_SW_C1-rare	SW	媳妇 Xi2fu0 Wife	445	媳	50+	34	(4)Rare_SW_C1-rare	SW	奴才 Nu2cai0 Servant	52	奴	73
15	(2)Frequent_SW_C1-rare	SW	舒服 Shu1fu0 Comfortable	293	舒	76	35	(4)Rare_SW_C1-rare	SW	牲口 sheng1kou0 animal	171	牲	50+
16	(2)Frequent_SW_C1-rare	SW	模糊 Mo2hu0 Vague	341	模	107	36	(4)Rare_SW_C1-rare	SW	琢磨 Zuo2mo0 Think	132	琢	50+
17	(2)Frequent_SW_C1-rare	SW	厉害 Li4hai0 Bravo	409	厉	50+	37	(4)Rare_SW_C1-rare	SW	帐篷 Zhan2peng0 Tent	180	帐	193
18	(2)Frequent_SW_C1-rare	SW	漂亮 Piao4liang0 Beautiful	526	漂	61	38	(4)Rare_SW_C1-rare	SW	扁担 Bian3dan0 Carrying pole	79	扁	174
19	(2)Frequent_SW_C1-rare	SW	眨巴 zha3 ba 0 wink and blink	297	眨	117	39	(4)Rare_SW_C1-rare	SW	戒指 jie4 zhi0 ring	58	戒	86
20	(2)Frequent_SW_C1-rare	SW	麻烦 Ma2fan0 Trouble	323	麻	195	40	(4)Rare_SW_C1-rare	SW	闺女 Gu1nv0 Daughter	126	闺	50+

No.	Category	Prosody	Prime	Frequency	Target	Frequency	No.	Category	Prosody	Prime	Frequency	Target	Frequency
41	(1)Frequent_SS_C1-Frequent	SS	沉思 Chen2si1 Think	277	沉	289	61	(3)Rare_SS_C1-frequent	SS	无意 Wu2yi4 Accidentally	121	无	4167
42	(1)Frequent_SS_C1-Frequent	SS	十分 Shi2fen1 Very much	3199	十	3548	62	(3)Rare_SS_C1-frequent	SS	泥巴 Ni2ba1 Mud	50	泥	372
43	(1)Frequent_SS_C1-Frequent	SS	饮食 Yin3shi2 Food and drink	220	饮	217	63	(3)Rare_SS_C1-frequent	SS	结识 jie2shi2 know	57	结	461
44	(1)Frequent_SS_C1-Frequent	SS	不同 Bu4tong2 Different	7851	不	46965	64	(3)Rare_SS_C1-frequent	SS	造就 Zao4jiu4 Create	125	造	572
45	(1)Frequent_SS_C1-Frequent	SS	克服 Ke4fu2 Overcome	777	克	740	65	(3)Rare_SS_C1-frequent	SS	分数 Fen1shu4 Score	73	分	3458
46	(1)Frequent_SS_C1-Frequent	SS	气候 Qi4hou4 Weather	1081	气	1947	66	(3)Rare_SS_C1-frequent	SS	面糊 mian4hu2 batter	50	面	1884
47	(1)Frequent_SS_C1-Frequent	SS	感情 Gan3qing2 Emotion	1836	感	789	67	(3)Rare_SS_C1-frequent	SS	忍耐 Ren3nai4 Tolerate	122	忍	331
48	(1)Frequent_SS_C1-Frequent	SS	花朵 Hua1duo3 Flower	1165	花	2108	68	(3)Rare_SS_C1-frequent	SS	叫唤 Jiao4huan4 Cry	56	叫	6289
49	(1)Frequent_SS_C1-Frequent	SS	点头 Dian3tou2 Nod	346	点	5958	69	(3)Rare_SS_C1-frequent	SS	中性 zhong1xing4 neutral	136	中	40849
50	(1)Frequent_SS_C1-Frequent	SS	口袋 kou3 dai4 Head	580	口	525	70	(3)Rare_SS_C1-frequent	SS	复合 Fu2he2 'get back together'	117	复	516
51	(2)frequent_SS_C1_rare	SS	新娘 Xin1niang2 Bride	442	新	102	71	(4)Rare_SS_C1-rare	SS	鼓掌 Gu3zhang3 Applaud	75	鼓	193
52	(2)frequent_SS_C1_rare	SS	利息 Li4xi1 Interests	286	利	85	72	(4)Rare_SS_C1-rare	SS	裂缝 Lie4feng4 Break	57	裂	192
53	(2)frequent_SS_C1_rare	SS	战友 Zhan4you3 Brothers in arms	203	战	69	73	(4)Rare_SS_C1-rare	SS	雷达 Lei2da2 Radar	161	雷	180
54	(2)frequent_SS_C1_rare	SS	寡妇 Gua3fu4 Widow	353	寡	50	74	(4)Rare_SS_C1-rare	SS	蠢才 Chun3cai2 Stupid	50	蠢	78
55	(2)frequent_SS_C1_rare	SS	征服 zheng1fu2 conquer	260	征	189	75	(4)Rare_SS_C1-rare	SS	港口 Gang3kou3 Port	156	港	189
56	(2)frequent_SS_C1_rare	SS	浆糊 jiang4hu2 paste	50+	浆	119	76	(4)Rare_SS_C1-rare	SS	消磨 Xiao1mo2 Kill the time	54	消	180
57	(2)frequent_SS_C1_rare	SS	损害 Sun3hai4 Hurt	397	损	87	77	(4)Rare_SS_C1-rare	SS	斗篷 Dou3peng2 Cloak	50	斗	50
58	(2)frequent_SS_C1_rare	SS	闪亮 Shan3liang4 Shiny	478	闪	78	78	(4)Rare_SS_C1-rare	SS	承担 Cheng2dan1 Undertake	161	承	144
59	(2)frequent_SS_C1_rare	SS	嘴巴 zu1ba1 mouth	302	嘴	102	79	(4)Rare_SS_C1-rare	SS	油脂 You2zhi3 Oil	74	油	190
60	(2)frequent_SS_C1_rare	SS	厌烦 Yan4fan2 Hate	375	厌	50	80	(4)Rare_SS_C1-rare	SS	仕女 Shi4nv3 Servant (female)	50	仕	50

⁵ **Frequency:** the occurrence times of a word per 10 million in the Corpus of Contemporary Mandarin; **Prosody:** SS: Strong-strong (spondee), full tone compound; SW: Strong-weak (trochee), neutral tone compound.

Appendices 3-2: non-related pairs

No.	Category	Prosody	Prime	Frequency	Target	Frequency	No.	Category	Prosody	Prime	Frequency	Target	Frequency
1	(1).Frequent SW C1-High	SW	队伍 Dui4wu0 Army	982	意	1602	21	(3).Rare SW C1-High	SW	状元 Zhuang4yuan0 Best scholar	53	大	877
2	(1).Frequent SW C1-High	SW	明白 Ming2bai0 Understand	1477	部	660	22	(3).Rare SW C1-High	SW	算计 suan4 ji0 Calculate	50+	结	1392
3	(1).Frequent SW C1-High	SW	大夫 Dai4fu0 Doctor	450	粮	20907	23	(3).Rare SW C1-High	SW	灯笼 Deng1long0 Lantern	78	见	600
4	(1).Frequent SW C1-High	SW	实在 Shi2zai0 Honest	1197	合	552	24	(3).Rare SW C1-High	SW	亲家 Qin4jia0 Relative of marriage	58	将	500
5	(1).Frequent SW C1-High	SW	结实 Jie1 shi0 Strong	223	衣	461	25	(3).Rare SW C1-High	SW	作坊 Zuo1fang0 workshop	68	岁	5480
6	(1).Frequent SW C1-High	SW	头发 Tou2fa0 Hair	629	时	4234	26	(3).Rare SW C1-High	SW	火候 Huo3hou0 degree of cooking	50+	含	1188
7	(1).Frequent SW C1-High	SW	和尚 He2shang0 Monk	293	事	83958	27	(3).Rare SW C1-High	SW	相声 xiang4sheng0 Comic dialogue	54	能	3422
8	(1).Frequent SW C1-High	SW	家伙 Jia1huo0 Dude	406	耳	5361	28	(3).Rare SW C1-High	SW	便宜 Pian2yi0 Cheap	50+	使	6790
9	(1).Frequent SW C1-High	SW	东西 Dong1xi0 chase	4174	石	1037	29	(3).Rare SW C1-High	SW	云彩 yun2 cai0 cloud	56	记	815
10	(1).Frequent SW C1-High	SW	本事 Ben3shi0 Capacity	165	脑	4115	30	(3).Rare SW C1-High	SW	名堂 ming2 tang0 large story	58	凑	2716
11	(2).Frequent SW C1-Low	SW	脾气 Pi2qi4 Temper	303	姑	142	31	(4).Rare SW C1-Low	SW	锄头 Chu2tou0 Hoe	78	巴	72
12	(2).Frequent SW C1-Low	SW	招呼 Zhao1hu0 Salutation	373	消	205	32	(4).Rare SW C1-Low	SW	丫头 ya1tou0 girl	91	裁	50+
13	(2).Frequent SW C1-Low	SW	商量 Shang1liang0 Talk over	436	朋	267	33	(4).Rare SW C1-Low	SW	唾沫 Tuo4mo0 Spittle	70	溜	50+
14	(2).Frequent SW C1-Low	SW	休息 Xiu1xi0 Rest	812	媳	68	34	(4).Rare SW C1-Low	SW	秀才 xiu4cai0 Scholar	75	奴	87
15	(2).Frequent SW C1-Low	SW	认识 Ren4shi0 Know	4395	舒	272	35	(4).Rare SW C1-Low	SW	衙门 Ya2men0 Government	88	牲	50+
16	(2).Frequent SW C1-Low	SW	尾巴 Wei3ba0 Tale	479	模	362	36	(4).Rare SW C1-Low	SW	欺负 Qi1fu0 Bully	90	琢	50+
17	(2).Frequent SW C1-Low	SW	弟兄 Di4xiong0 Brother	217	厉	132	37	(4).Rare SW C1-Low	SW	胭脂 yan1 zhi0 blusher	50	帐	50
18	(2).Frequent SW C1-Low	SW	窗户 Chuang1hu0 Window	226	漂	224	38	(4).Rare SW C1-Low	SW	爽快 Shuang3kuai0 Balminess	57	扁	50+
19	(2).Frequent SW C1-Low	SW	胳膊 Ge1bo0 Arm	247	眨	50+	39	(4).Rare SW C1-Low	SW	妖精 Yao1jing0 Goblin	50+	戒	72
20	(2).Frequent SW C1-Low	SW	豆腐 Dou4fu0 Tofu	189	麻	125	40	(4).Rare SW C1-Low	SW	吆喝 Yao1he0 Cry	80	闹	50+

No.	Category	Prosody	Prime	Frequency	Target	Frequency	No.	Category	Prosody	Prime	Frequency	Target	Frequency
41	1).Frequent SS C1-Frequent	SS	入伍 ru4wu3 join the army	81	沉	1401	61	(3).Rare SW C1-frequent	SS	美元 Mei3yuan2 Dollar	77	无	2100
42	1).Frequent SS C1-Frequent	SS	雪白 Xue3bai2 White	1150	十	350	62	(3).Rare SW C1-frequent	SS	累计 Lei3ji4 'accumulative'	67	泥	459
43	1).Frequent SS C1-Frequent	SS	车夫 Che1fu1 Driver	220	饮	1140	63	(3).Rare SW C1-frequent	SS	鸟笼 Niao3long2 'Bird cage'	50	结	558
44	1).Frequent SS C1-Frequent	SS	现在 Xian4zai4 Now	7642	不	526	64	(3).Rare SW C1-frequent	SS	画家 Hua4jia1 'Painter'	152	造	1614
45	1).Frequent SS C1-Frequent	SS	其实 qi2shi2 actually	1301	克	10206	65	(3).Rare SW C1-frequent	SS	工厂 Gong1fang3 'Factory'	50	分	554
46	1).Frequent SS C1-Frequent	SS	开发 Kai1fa1 exploit	973	气	3105	66	(3).Rare SW C1-frequent	SS	问候 Wen4hou4 'greetings/ regards'	62	面	4505
47	1).Frequent SS C1-Frequent	SS	高尚 Gao1shang4 Nobleness	455	感	6508	67	(3).Rare SW C1-frequent	SS	歌声 Ge1sheng1 song	237	忍	608
48	1).Frequent SS C1-Frequent	SS	团伙 tuan2huo3 gang group	423	花	1182	68	(3).Rare SW C1-frequent	SS	适宜 Shi4yi2 appropriate	46	叫	120
49	1).Frequent SS C1-Frequent	SS	广西 Guang3xi1 guangxi province	291	点	576	69	(3).Rare SW C1-frequent	SS	光彩 Guang1cai3 'glamour'	161	中	2143
50	1).Frequent SS C1-Frequent	SS	军事 Jun1shi4 military	1321	口	998	70	(3).Rare SW C1-frequent	SS	教堂 Jiao4tang2 'church'	197	复	859
51	(2).frequent SS C1 rare	SS	勇气 Yong3qi4 Courage	303	新	142	71	(4).Rare SS C1 rare	SS	枕头 Zhen3tou2 Pillow	124	鼓	82
52	(2).frequent SS C1 rare	SS	欢呼 Huan1hu1 Cheer up	390	利	109	72	(4).Rare SS C1 rare	SS	额头 E2tou2 forehead	87	裂	327
53	(2).frequent SS C1 rare	SS	质量 Zhi4liang4 Quality	2429	战	71	73	(4).Rare SS C1 rare	SS	泡沫 Pao4mo4 Foam	65	雷	109
54	(2).frequent SS C1 rare	SS	叹息 Tan4xi1 Sigh	473	寡	50	74	(4).Rare SS C1 rare	SS	蠢才 chun3cai2 stupid person	50+	蠢	50+
55	(2).frequent SS C1 rare	SS	意识 Yi4shi1 Recognition	1431	征	165	75	(4).Rare SS C1 rare	SS	嗓门 Sang3men2 Voice	103	港	88
56	(2).frequent SS C1 rare	SS	嘴巴 Zui3ba1 Mouth	302	蒙	95	76	(4).Rare SS C1 rare	SS	辜负 Gu1fu4 Let somebody down	61	消	50
57	(2).frequent SS C1 rare	SS	堂兄 Tang2xiong1 Cousin	334	损	84	77	(4).Rare SS C1 rare	SS	油脂 You2zhi3 Oil	74	斗	190
58	(2).frequent SS C1 rare	SS	客户 Ke4hu4 Client	260	闪	78	78	(4).Rare SS C1 rare	SS	欢快 Huan1kuai4 happy	94	承	90
59	(2).frequent SS C1 rare	SS	赤裸 Chi3bo2 nude	50	嘴	157	79	(4).Rare SS C1 rare	SS	酒精 Jiu3jing1 Alcohol	119	油	162
60	(2).frequent SS C1 rare	SS	迂腐 Yu1fu3 pedantic	250	灰	50+	80	(4).Rare SS C1 rare	SS	吃喝 Chi1he1 food and drink	57	仕	50+

Appendices 3-3: pseudo-character pairs

No.	Category	Prosody	Prime	Frequency	C1Frequency	Target	No.	Category	Prosody	PrimeCompound	imeFrequency	C1Frequency	RelatedTarget
81	(3) Rare_SW_C1-High	SW	下场 Xia4chang0 end of life	50	14051	Pseudo-character	101	(3) rare_SS_C1_frequent	SS	农场 nong2chang3 'farm'	197	566	Pseudo-character
82	(3) Rare_SW_C1-High	SW	亲戚 Qin1qi0 Relatives	50	500	Pseudo-character	102	(3) rare_SS_C1_frequent	SS	外戚 Wai4qi4 Relatives	50	4489	Pseudo-character
83	(3) Rare_SW_C1-High	SW	对付 Du1fu0 Handle	50	30936	Pseudo-character	103	(3) rare_SS_C1_frequent	SS	支付 Zhi1fu4 Pay	257	1359	Pseudo-character
84	(3) Rare_SW_C1-High	SW	行李 Xing2li0 Baggage	50	2093	Pseudo-character	104	(3) rare_SS_C1_frequent	SS	桃李 tao2li3 student	50	110	Pseudo-character
85	(3) Rare_SW_C1-High	SW	打量 Da3 liang0 Stare up and down	50	4413	Pseudo-character	105	(3) rare_SS_C1_frequent	SS	适量 shi4liang4 correct quantity	90	120	Pseudo-character
86	(3) Rare_SW_C1-High	SW	念头 Nian4tou0 Idea	50	654	Pseudo-character	106	(3) rare_SS_C1_frequent	SS	人头 Ren2tou2 Head	77	33991	Pseudo-character
87	(3) Rare_SW_C1-High	SW	点心 Dian3 xin0 Dessert	50	5958	Pseudo-character	107	(3) rare_SS_C1_frequent	SS	用心 yong4xin1 attentively	91	17259	Pseudo-character
88	(3) Rare_SW_C1-High	SW	称呼 cheng1 hu0 call	50	1313	Pseudo-character	108	(3) rare_SS_C1_frequent	SS	高呼 gao1hu1 cry out	57	6508	Pseudo-character
89	(3) Rare_SW_C1-High	SW	提防 di1fang0 beware of	54	1295	Pseudo-character	109	(3) rare_SS_C1_frequent	SS	国防 guo2fang2 national defence	383	1203	Pseudo-character
90	(3) Rare_SW_C1-High	SW	快活 kuai4 huo0 cheerfulness	50	3321	Pseudo-character	110	(3) rare_SS_C1_frequent	SS	养活 yang3 huo0 feed	95	742	Pseudo-character
91	(3) Rare_SW_C1-High	SW	老实 lao3shi0 caracity	51	6236	Pseudo-character	111	(3) rare_SS_C1_frequent	SS	充实 chong1shi2 enrich	234	155	Pseudo-character
92	(3) Rare_SW_C1-High	SW	东家 dong1jia1 landlord	70	1037	Pseudo-character	112	(3) rare_SS_C1_frequent	SS	全家 quan2jia1 family	194	4334	Pseudo-character
93	(3) Rare_SW_C1-High	SW	怪物 guai4wu0 monster	117	739	Pseudo-character	113	(3) rare_SS_C1_frequent	SS	礼物 li3wu4 present	141	341	Pseudo-character
94	(3) Rare_SW_C1-High	SW	包涵 bao1 han0 excuser	89	903	Pseudo-character	114	(3) rare_SS_C1_frequent	SS	内涵 nei4han2 connotation	177	6025	Pseudo-character
95	(3) Rare_SW_C1-High	SW	利索 li4 suo0 neat and tidy	50*	405	Pseudo-character	115	(3) rare_SS_C1_frequent	SS	线索 xian4suo3 clue	189	1477	Pseudo-character
96	(3) Rare_SW_C1-High	SW	补丁 bu3ding0 patch	50*	471	Pseudo-character	116	(3) rare_SS_C1_frequent	SS	拉丁 lai3ding1 latine	53	1404	Pseudo-character
97	(3) Rare_SW_C1-High	SW	利落 li4 luo0 smart	50*	405	Pseudo-character	117	(3) Rare_SW_C1-frequent	SW	冷落 leng3 luo4 desolate	72	872	Pseudo-character
98	(3) Rare_SW_C1-High	SW	指甲 zhi3 jia0 nail	65	3519	Pseudo-character	118	(3) Rare_SW_C1-frequent	SW	铁甲 tie3jia3 armour	50*	1050	Pseudo-character
99	(3) Rare_SW_C1-High	SW	张罗 zhang1 luo0 take care of	52	3235	Pseudo-character	119	(3) Rare_SW_C1-frequent	SW	包罗 bao1 luo2 include	50*	903	Pseudo-character
100	(3) Rare_SW_C1-High	SW	马虎 ma3hu0 careless	57	1455	Pseudo-character	120	(3) Rare_SW_C1-frequent	SW	老虎 lao2hu3 tiger	209	6236	Pseudo-character

No.	Category	Prosody	Prime	Frequency	C1Frequency	Target	No.	Category	Prosody	Prime	Frequency	C1Frequency	Target
121	(4) Rare_SW_C1-Low	SW	糊涂 Hu2tu0 stupid	62	103	Pseudo-character	141	(4) rare_SS_C1-rare	SS	征途 zheng1tu2 journey	50	189	Pseudo-character
122	(4) Rare_SW_C1-Low	SW	答应 Da1ying0 Respond	73	371	Pseudo-character	142	(4) rare_SS_C1-rare	SS	回应 hui2ying4 respond	50*	1868	Pseudo-character
123	(4) Rare_SW_C1-Low	SW	屁股 Pi4gu0 Ass	63	57	Pseudo-character	143	(4) rare_SS_C1-rare	SS	入股 ru4gu3 Invest	50	106	Pseudo-character
124	(4) Rare_SW_C1-Low	SW	眉毛 Mei2mao0 Eyebrow	68	172	Pseudo-character	144	(4) rare_SS_C1-rare	SS	绒毛 rong2mao2 Wool/fluff	66	63	Pseudo-character
125	(4) Rare_SW_C1-Low	SW	秧歌 yang1ge0 Dance	68	147	Pseudo-character	145	(4) rare_SS_C1-rare	SS	赞歌 zan4 ge1 song of praise	50*	57	Pseudo-character
126	(4) Rare_SW_C1-Low	SW	伙计 huo3 ji0 camarade	79	88	Pseudo-character	146	(4) rare_SS_C1-rare	SS	审计 shen3ji4 audit	50*	78	Pseudo-character
127	(4) Rare_SW_C1-Low	SW	规矩 gui1 ju0 rule	85	69	Pseudo-character	147	(4) rare_SS_C1-rare	SS	逾越 yu3ju3 overstep	50	50	Pseudo-character
128	(4) Rare_SW_C1-Low	SW	柴火 chai2 huo0 firewood	50*	170	Pseudo-character	148	(4) rare_SS_C1-rare	SS	烈火 lie4huo3 raging fire	114	57	Pseudo-character
129	(4) Rare_SW_C1-Low	SW	暖和 nuan3 he0 warm	50*	287	Pseudo-character	149	(4) rare_SS_C1-rare	SS	附和 fu4he2 echo	50*	214	Pseudo-character
130	(4) Rare_SW_C1-Low	SW	招牌 zhao1 pai0 brand	50*	205	Pseudo-character	150	(4) rare_SS_C1-rare	SS	铜牌 tong2pai2 bronze medal	50*	455	Pseudo-character
131	(4) Rare_SW_C1-Low	SW	栅栏 zha4lan0 fence	50*	50*	Pseudo-character	151	(4) rare_SS_C1-rare	SS	围栏 wei2lan2 rail	50*	495	Pseudo-character
132	(4) Rare_SW_C1-Low	SW	芝麻 zhi1ma0 sesame	50*	50*	Pseudo-character	152	(4) rare_SS_C1-rare	SS	棉麻 mian2ma2 cotton and linen	50*	257	Pseudo-character
133	(4) Rare_SW_C1-Low	SW	稳当 wen3dang0 stable	50*	220	Pseudo-character	153	(4) rare_SS_C1-rare	SS	充当 chong1dang1 serve as	147	155	Pseudo-character
134	(4) Rare_SW_C1-Low	SW	舒适 shu1tan0 comfortable	50*	76	Pseudo-character	154	(4) rare_SS_C1-rare	SS	平坦 ping2 tan3 flatness	123	650	Pseudo-character
135	(4) Rare_SW_C1-Low	SW	特务 te4 wu0 spy	68	263	Pseudo-character	155	(4) rare_SS_C1-rare	SS	债务 zhai4wu4 debt	120	137	Pseudo-character
136	(4) Rare_SW_C1-Low	SW	脊梁 ji2liang0 backbone	50*	129	Pseudo-character	156	(4) rare_SS_C1-rare	SS	鼻梁 bi2liang2 nose	66	193	Pseudo-character
137	(4) Rare_SW_C1-Low	SW	勤快 qin2kuai0 hardworking	50*	140	Pseudo-character	157	(4) rare_SS_C1-rare	SS	痛快 tong4kuai4 delighted	50*	411	Pseudo-character
138	(4) Rare_SW_C1-Low	SW	世故 shi4gu0 manipulative	50*	311	Pseudo-character	158	(4) Rare_SS_C1_rare	SS	典故 dian3gu4 literary quotation	50*	67	Pseudo-character
139	(4) Rare_SW_C1-Low	SW	思量 si1liang0 think	50*	255	Pseudo-character	159	(4) Rare_SS_C1_rare	SS	储量 chu3liang4 reserves	109	70	Pseudo-character
140	(4) Rare_SW_C1-Low	SW	凉快 liang2kuai0 cool	50*	243	Pseudo-character	160	(4) Rare_SS_C1_rare	SS	飞快 fei1kuai4 very fast	122	50*	Pseudo-character

No.	Category	Prosody	Prime	Frequency	C1Frequency	Target	No.	Category	Prosody	PrimeCompound	Frequency	C1Frequency	RelatedTarget
81	(3).Rare_SW_C1-High	SW	下场 Xia4chang0 end of life	50	14051	Pseudo-character	101	(3).rare_SS_C1.frequent	SS	农场 nong2chang3 'farm'	197	566	Pseudo-character
82	(3).Rare_SW_C1-High	SW	亲戚 Qin1qi0 Relatives	50	500	Pseudo-character	102	(3).rare_SS_C1.frequent	SS	外戚 Wai4qi4 Relatives	50	4489	Pseudo-character
83	(3).Rare_SW_C1-High	SW	对付 Dui4fu0 Handle	50	30936	Pseudo-character	103	(3).rare_SS_C1.frequent	SS	支付 Zhi1fu4 Pay	257	1359	Pseudo-character
84	(3).Rare_SW_C1-High	SW	行李 Xing2li0 Baggage	50	2093	Pseudo-character	104	(3).rare_SS_C1.frequent	SS	桃李 tao2li3 student	50	110	Pseudo-character
85	(3).Rare_SW_C1-High	SW	打量 Da3 liang0 Stare up and down	50	4413	Pseudo-character	105	(3).rare_SS_C1.frequent	SS	打量 shi4liang4 correct quantity	90	120	Pseudo-character
86	(3).Rare_SW_C1-High	SW	念头 Nian4tou0 Idea	50	654	Pseudo-character	106	(3).rare_SS_C1.frequent	SS	人头 Ren2tou2 Head	77	33991	Pseudo-character
87	(3).Rare_SW_C1-High	SW	点心 Dian3 xin0 Dessert	50	5958	Pseudo-character	107	(3).rare_SS_C1.frequent	SS	用心 yong4xin1 attentively	91	17259	Pseudo-character
88	(3).Rare_SW_C1-High	SW	称呼 cheng1 hu0 call	50	1313	Pseudo-character	108	(3).rare_SS_C1.frequent	SS	高呼 gao1hu1 cry out	57	6508	Pseudo-character
89	(3).Rare_SW_C1-High	SW	提防 di1fang0 beware of	54	1295	Pseudo-character	109	(3).rare_SS_C1.frequent	SS	国防 guo2fang2 national defence	383	1203	Pseudo-character
90	(3).Rare_SW_C1-High	SW	快活 kuai4 hu0 cheerfulness	50	3321	Pseudo-character	110	(3).rare_SS_C1.frequent	SS	养活 yang3 hu0 feed	95	742	Pseudo-character
91	(3).Rare_SW_C1-High	SW	老实 lao3shi0 caracity	51	6236	Pseudo-character	111	(3).rare_SS_C1.frequent	SS	充实 chong1shi2 enrich	234	155	Pseudo-character
92	(3).Rare_SW_C1-High	SW	东家 dong1jia0 landlord	70	1037	Pseudo-character	112	(3).rare_SS_C1.frequent	SS	全家 quan2jia1 family	194	4334	Pseudo-character
93	(3).Rare_SW_C1-High	SW	怪物 gua1wu0 monster	117	739	Pseudo-character	113	(3).rare_SS_C1.frequent	SS	礼物 li3wu4 present	141	341	Pseudo-character
94	(3).Rare_SW_C1-High	SW	包涵 bao1 han0 excuser	89	903	Pseudo-character	114	(3).rare_SS_C1.frequent	SS	内涵 nei4han2 connotation	177	6025	Pseudo-character
95	(3).Rare_SW_C1-High	SW	利索 li4 suo0 neat and tidy	50*	405	Pseudo-character	115	(3).rare_SS_C1.frequent	SS	线索 xian4suo3 clue	189	1477	Pseudo-character
96	(3).Rare_SW_C1-High	SW	补丁 bu3ding0 patch	50*	471	Pseudo-character	116	(3).rare_SS_C1.frequent	SS	拉丁 la1ding1 latina	53	1404	Pseudo-character
97	(3).Rare_SW_C1-High	SW	利落 li4 luo0 smart	50*	405	Pseudo-character	117	(3).Rare_SW_C1-frequent	SW	冷落 leng3 luo4 desolate	72	872	Pseudo-character
98	(3).Rare_SW_C1-High	SW	指甲 zhi3 jia0 nail	65	3519	Pseudo-character	118	(3).Rare_SW_C1-frequent	SW	铁甲 tie3jia3 armour	50*	1050	Pseudo-character
99	(3).Rare_SW_C1-High	SW	张罗 zhang1luo0 take care of	52	3235	Pseudo-character	119	(3).Rare_SW_C1-frequent	SW	包罗 bao1luo2 include	50*	903	Pseudo-character
100	(3).Rare_SW_C1-High	SW	马虎 ma3hu0 careless	57	1455	Pseudo-character	120	(3).Rare_SW_C1-frequent	SW	老虎 lao2hu3 tiger	209	6236	Pseudo-character

No.	Category	Prosody	Prime	Frequency	C1Frequency	Target	No.	Category	Prosody	Prime	Frequency	C1Frequency	Target
121	(4).Rare_SW_C1-Low	SW	糊涂 Hu2tu0 stupid	62	103	Pseudo-character	141	(4).rare_SS_C1-rare	SS	征徭 zheng1tu2 journey	50	189	Pseudo-character
122	(4).Rare_SW_C1-Low	SW	答应 Da1 ying0 Respond	73	371	Pseudo-character	142	(4).rare_SS_C1-rare	SS	回应 hui2ying4 respond	50*	1868	Pseudo-character
123	(4).Rare_SW_C1-Low	SW	屁股 Pi4gu0 Ass	63	57	Pseudo-character	143	(4).rare_SS_C1-rare	SS	入股 Ru4gu3 invest	50	106	Pseudo-character
124	(4).Rare_SW_C1-Low	SW	眉毛 Mei2mao0 Eyebrow	68	172	Pseudo-character	144	(4).rare_SS_C1-rare	SS	绒毛 Rong2mao2 Wool/fluff	66	63	Pseudo-character
125	(4).Rare_SW_C1-Low	SW	秧歌 Yang1ge0 Dance	68	147	Pseudo-character	145	(4).rare_SS_C1-rare	SS	赞歌 zan4 ge1 song of praise	50*	57	Pseudo-character
126	(4).Rare_SW_C1-Low	SW	伙计 huo3 ji0 camarade	79	88	Pseudo-character	146	(4).rare_SS_C1-rare	SS	审计 shen3ji4 audit	50*	78	Pseudo-character
127	(4).Rare_SW_C1-Low	SW	规矩 gui1 ju0 rule	85	69	Pseudo-character	147	(4).rare_SS_C1-rare	SS	逾越 yu3yu3 overstep	50	50	Pseudo-character
128	(4).Rare_SW_C1-Low	SW	柴火 chai2 hu0 firewood	50*	170	Pseudo-character	148	(4).rare_SS_C1-rare	SS	烈火 lie4huo3 raging fire	114	57	Pseudo-character
129	(4).Rare_SW_C1-Low	SW	暖和 nuan3 ha0 warm	50*	287	Pseudo-character	149	(4).rare_SS_C1-rare	SS	附和 fu4he2 echo	50*	214	Pseudo-character
130	(4).Rare_SW_C1-Low	SW	招牌 zhao1 pai0 brand	50*	205	Pseudo-character	150	(4).rare_SS_C1-rare	SS	铜牌 tong2pai2 bronze medal	50*	455	Pseudo-character
131	(4).Rare_SW_C1-Low	SW	栅栏 zha4lan0 fence	50*	50*	Pseudo-character	151	(4).rare_SS_C1-rare	SS	围栏 wei2lan2 rail	50*	495	Pseudo-character
132	(4).Rare_SW_C1-Low	SW	芝麻 zhi1ma0 sesame	50*	50*	Pseudo-character	152	(4).rare_SS_C1-rare	SS	棉麻 mian2ma2 cotton and linen	50*	257	Pseudo-character
133	(4).Rare_SW_C1-Low	SW	稳当 wen3dang0 stable	50*	220	Pseudo-character	153	(4).rare_SS_C1-rare	SS	充当 chong1dang1 serve as	147	155	Pseudo-character
134	(4).Rare_SW_C1-Low	SW	舒适 shu1tan0 comfortable	50*	76	Pseudo-character	154	(4).rare_SS_C1-rare	SS	平坦 ping2 tan3 flatness	123	650	Pseudo-character
135	(4).Rare_SW_C1-Low	SW	特务 te4 wu0 spy	68	263	Pseudo-character	155	(4).rare_SS_C1-rare	SS	债务 zhai4wu4 debt	120	137	Pseudo-character
136	(4).Rare_SW_C1-Low	SW	脊梁 ji2liang0 backbone	50*	129	Pseudo-character	156	(4).rare_SS_C1-rare	SS	鼻梁 bi2liang2 nose	66	193	Pseudo-character
137	(4).Rare_SW_C1-Low	SW	勤快 qin2kuai0 hardworking	50*	140	Pseudo-character	157	(4).rare_SS_C1-rare	SS	痛快 tong4kuai4 delighted	50*	411	Pseudo-character
138	(4).Rare_SW_C1-Low	SW	世故 shi4gu0 manipulative	50*	311	Pseudo-character	158	(4).Rare_SS_C1.rare	SS	典故 dian3gu4 literary quotation	50*	67	Pseudo-character
139	(4).Rare_SW_C1-Low	SW	思量 si1liang0 think	50*	255	Pseudo-character	159	(4).Rare_SS_C1.rare	SS	储量 chu3liang4 reserves	109	70	Pseudo-character
140	(4).Rare_SW_C1-Low	SW	凉快 liang2kuai0 cool	50*	243	Pseudo-character	160	(4).Rare_SS_C1.rare	SS	飞快 fei1kuai4 very fast	122	50*	Pseudo-character

Appendices 4⁶

No.	Neutral tone compound	Occurrence Times	Stroke number	No.	Neutral tone compound	Occurrence Times	Stroke number	No.	Neutral tone compound	Occurrence Times	Stroke number
1	折腾 zhe1teng0 toss about	89	20	21	裁缝 cai2feng0 tailor	50*	25	41	咳嗽 ke2sou0 cough	226	23
2	狐狸 hu2li0 fox	87	18	22	琢磨 zhuo2 mo0 think about	132	28	42	快活 kuai4huo0 happy	145	16
3	稀罕 xi1han0 like	55	19	23	明白 ming2 bai0 understand	477	13	43	称呼 cheng1hu0 name	160	18
4	吆喝 yao1 he0 cry out	80	18	24	丫头 ya1 tou0 little girl	91	8	44	官司 guan1si0 lawsuit	50*	13
5	豁达 hu1da0 hang out	50*	19	25	胭脂 yan1zhi0 blush	50*	20	45	规矩 gu1ju0 rule	169	17
6	胳膊 ge1bo0 arm	247	24	26	闺女 gui1 nv0 daughter	126	12	46	在乎 zai4hu0 care about	74	11
7	豆腐 dou4 fu0 tofu	189	21	27	戒指 jie2 zhi0 ring	58	16	47	应酬 ying1chou0 dinner party	52	20
8	张罗 zhang1 luo0 get together	52	15	28	妖精 yao1 jing0 elf	50*	21	48	累赘 lei2zhu0 problem	50*	25
9	苍蝇 cang1ying0 fly	108	21	29	窟窿 ku1 long0 hole	62	29	49	叫唤 jiao4huan0 cry out	56	15
10	挖苦 wa1ku0 sarcastical	50*	17	30	冤枉 yuan1 wang0 treat unjustly	76	18	50	眨巴 zha1ba0 blink	50*	13
11	糊涂 hu2 tu0 stupid	204	25	31	嫁妆 jia4 zhuang0 dowry	50*	19	51	马虎 ma3hu0 careless	57	11
12	行李 xing2 li0 luggage	155	13	32	罐头 guan4 tou0 can	73	28	52	行当 hang2dang0 profession	68	12
13	东西 dong1xi0 thing	474	11	33	队伍 dui4 wu0 troops	582	10	53	打发 da4fa0 send and dispatch	97	10
14	大夫 dai4 fu0 doctor	450	7	34	记号 ji4hao0 marker	106	10	54	凑合 cou1 he0/ not too bad	50*	17
15	姑娘 gu1 niang0 girl	752	18	35	麻烦 ma2fan0 trouble	323	21	55	牲口 sheng1 kou0/ animal	171	12
16	将就 jiang1 jiu0 put up with	50*	21	36	窗户 chuang1 hu0 window	226	16	56	下场 xia4 chang0/ end of life	104	9
17	云彩 yun2 cai0 cloud	56	15	37	媳妇 xi2 fu0 wife	445	19	57	暖和 nuan3 he0/ warm	68	21
18	巴掌 ba1 zhang0 hand palm	83	16	38	衣裳 yi1 shang0 clothes	202	20	58	大意 da4 yi0/ careless	50*	16
19	名堂 ming2 tang0	58	17	39	耳朵 er3 duo0 ear	506	12	59	念头 nian4 tou0/ idea	179	13
20	动弹 dong4 tan0 movement	63	17	40	萝卜 luo2 bo0 radish	87	13	60	清楚 qing1 chu0/ clear	615	24

⁶ **Frequency:** the word's frequency is calculated by counting the occurrence times per 10 million in the Corpus of Contemporary Mandarin; **Stroke number:** the number of strokes contained in the two characters of a compound.

Reference

- Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, 16(1-2), 285-311.
- Ashby, J. (2010). Phonology is fundamental in skilled reading: Evidence from ERPs. *Psychonomic bulletin & review*, 17(1), 95-100.
- Ashby, J., & Martin, A. E. (2008). Prosodic phonological representations early in visual word recognition. *Journal of Experimental Psychology: Human perception and performance*, 34(1), 224.
- Ashby, J., Sanders, L. D., & Kingston, J. (2009). Skilled readers begin processing sub-phonemic features by 80 ms during visual word recognition: Evidence from ERPs. *Biological psychology*, 80(1), 84-94.
- Baayen, H., Burani, C., & Schreuder, R. (1997). Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In *Yearbook of morphology 1996* (pp. 13-33). Springer.
- Baayen, H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of memory and language*, 37(1), 94-117.
- Badecker, W. (2001). Lexical composition and the production of compounds: Evidence from errors in naming. *Language and Cognitive Processes*, 16(4), 337-366.
- Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: evidence from ERPs. *NeuroReport*, 15(3), 545-548.
- Beckman, M. E. (2012). Stress and non-stress accent. In *Stress and Non-Stress Accent*. De Gruyter Mouton.
- Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology*, 3, 255-309.
- Bisetto, A., & Scalise, S. (2005). The classification of compounds. *Lingue e linguaggio*, 4(2), 319-310.
- Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding In Gernsbacher MA (Ed.), *Handbook of Psycholinguistics* (pp. 945–984). In. San Diego: Academic Press.
- Böcker, K. B., Bastiaansen, M. C., Vroomen, J., Brunia, C. H., & De Gelder, B. (1999). An ERP correlate of metrical stress in spoken word recognition. *Psychophysiology*, 36(6), 706-720.
- Boersma, P. (2006). *Praat: doing phonetics by computer*. <http://www.praat.org/>
- Boersma, P., & Weenink, D. (2006). Praat (version 4.5)[Computer software]. *Amsterdam: Institute of Phonetic Sciences*.
- Bohn, K., Knaus, J., Wiese, R., & Domahs, U. (2013). The influence of rhythmic (ir) regularities on speech processing: Evidence from an ERP study on German phrases. *Neuropsychologia*, 51(4), 760-771.
- Bolinger, D. L. (1958). A theory of pitch accent in English. *Word*, 14(2-3), 109-149.
- Bolinger, D. L. (1961). Contrastive accent and contrastive stress. *Language*, 37(1), 83-96.
- Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5(1), 34-44.
- Brown, J. S., & Carr, T. H. (1993). Limits on perceptual abstraction in reading: Asymmetric transfer between surface forms differing in typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(6), 1277.
- Brunellière, A., Sánchez-García, C., Ikumi, N., & Soto-Faraco, S. (2013). Visual information constrains early and late stages of spoken-word recognition in sentence context. *International Journal of Psychophysiology*, 89(1), 136-147.
- Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes*, 2(3-4), 217-227.

- Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language*, 18(4), 342-352.
- Cao, J. (1986). A study on neutral tone in Standard Chinese. *Appl. Acoust*, 4, 1-6.
- Cao, J. (1992). On neutral-tone syllables in Mandarin Chinese. *Canadian Acoustics*, 20(3), 49-50.
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28(3), 297-332.
- Caramazza, A., Miceli, G., Silveri, M. C., & Laudanna, A. (1985). Reading mechanisms and the organisation of the lexicon: Evidence from acquired dyslexia. *Cognitive neuropsychology*, 2(1), 81-114.
- Ceccagno, A., & Scalise, S. (2006). Classification, structure and headedness of Chinese compounds. *Lingue e linguaggio*(2), 233-260.
- Chao, Y. R. (1968). *Language and symbolic systems* (Vol. 260). Cambridge University Press Cambridge.
- Chen, C.-Y. (1984). Neutral tone in Mandarin: phonotactic description and the issue of the norm.
- Chen, M.-J., Weekes, B. S., Peng, D., & Lei, Q. (2006). Effects of semantic radical consistency and combinability on Chinese character processing. In *Handbook of East Asian Psycholinguistics*. Cambridge University Press.
- Chen, M. J. (2003). Component skills for reading Chinese in primary school children. *Reading development in Chinese children*, 19-33.
- Chen, Y., & Xu, Y. (2006a). Production of Weak Elements in Speech – Evidence from F₀ Patterns of Neutral Tone in Standard Chinese. *Phonetica*, 63(1), 47-75.
<https://doi.org/10.1159/000091406>
- Chen, Y., & Xu, Y. (2006b). Production of weak elements in speech—evidence from f₀ patterns of neutral tone in Standard Chinese. *Phonetica*, 63(1), 47-75.
- Cheng, C.-C. (1973). *A synchronic phonology of Mandarin Chinese*. Mouton.
- Chomsky, N., & Halle, M. (1968). The sound pattern of English.
- Clark, J., & Yallop, C. (1990). *An Introduction to Phonetics & Phonology* Basil Blackwell. In: Oxford and Cambridge, MA.
- Colé, P., Segui, J., & Taft, M. (1997). Words and morphemes as units for lexical access. *Journal of memory and language*, 37(3), 312-330.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, 82(6), 407.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological review*, 108(1), 204.
- Comrie, B. (1987). *The world's major languages*. Routledge London, UK.
- Connine, C. M., Clifton Jr, C., & Cutler, A. (1987). Effects of lexical stress on phonetic categorization. *Phonetica*, 44(3), 133-146.
- Cooper, F. S., Liberman, A. M., Harris, K. S., & Grubb, P. M. (1960). *Some input-output relations observed in experiments on the perception of speech*. Association internationale de cybernétique.
- Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech*, 45(3), 207-228.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13(1), 21-58.
- Cruttenden, A. (1997). *Intonation*. Cambridge University Press.
- Cui, L., Häikiö, T., Zhang, W., Zheng, Y., & Hyönä, J. (2017). Reading monomorphemic and compound words in Chinese. *The Mental Lexicon*, 12(1), 1-20.

- Cui, L., Wang, J., Zhang, Y., Cong, F., Zhang, W., & Hyönä, J. (2021). Compound word frequency modifies the effect of character frequency in reading Chinese. *Quarterly Journal of Experimental Psychology*, 74(4), 610-633.
- Cutler, A. (1984). B. Butterworth (Ed.), *Language Production, Vol. 2: Development, Writing and other Language Processes*.
- Cutler, A. (1995). Spoken word recognition and production.
- Cutler, A. (2005). Lexical stress.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of memory and language*, 31(2), 218-236.
- Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, 2(3-4), 133-142.
- Cutler, A., & Clifton, C. (1984). The use of prosodic information in word recognition.
- Cutler, A., & Clifton, C. (1999). Comprehending spoken language: A blueprint of the listener. *The neurocognition of language*, 123-166.
- Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), 141-201.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human perception and performance*, 14(1), 113.
- Cutler, A., & Van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech*, 44(2), 171-195.
- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, 8(4), 919-924.
- Delattre, P. (1966). Les dix intonations de base du français. *French review*, 1-14.
- Desimone, R., Miller, E. K., Chelazzi, L., & Lueschow, A. (1995). Multiple memory systems in the visual cortex.
- Desroches, A. S., Newman, R. L., & Joanisse, M. F. (2009). Investigating the time course of spoken word recognition: Electrophysiological evidence for the influences of phonological similarity. *Journal of Cognitive Neuroscience*, 21(10), 1893-1906.
- Domahs, U., Genc, S., Knaus, J., Wiese, R., & Kabak, B. (2013). Processing (un-) predictable word stress: ERP evidence from Turkish. *Language and Cognitive Processes*, 28(3), 335-354.
- Dressler, W. U. (2006). Compound Types. The Representation and Processing of Compound Words. G. Libben and G. Jarema (eds.), 15, 242.
- Duanmu, S. (2007). *The phonology of standard Chinese*. OUP Oxford.
- Duanmu, S. (2009). *Syllable structure: The limits of variation*. Oxford University Press.
- Duanmu, S. (2014). Syllable structure and stress. *The handbook of Chinese linguistics*, 422-442.
- Dufour, S., Brunellière, A., & Frauenfelder, U. H. (2013). Tracking the time course of word-frequency effects in auditory word recognition with event-related potentials. *Cognitive science*, 37(3), 489-507.
- Eckstein, K., & Friederici, A. D. (2005). Late interaction of syntactic and prosodic processes in sentence comprehension as revealed by ERPs. *Cognitive Brain Research*, 25(1), 130-143.
- Eisner, F., & McQueen, J. M. (2018). *Speech perception* (Vol. 3). Wiley Hoboken.
- Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, 10(5), 590-604.
- Eysenck, M. W., & Keane, M. T. (2015). *Cognitive psychology: A student handbook*. Psychology press.
- Feldman, L. B., & Siok, W. W. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of memory and language*, 40(4), 559-576.
- Feng, S. (2017). *Prosodic Morphology in Mandarin Chinese*. Routledge.
- Fiorentino, R., & Poeppel, D. (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes*, 22(7), 953-1000.
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of verbal learning and verbal behavior*, 12(6), 627-635.

- Fox, R. A., & Unkefer, J. (1985). THE EFFECT OF LEXICAL STATUS ON THE PERCEPTION OF TONE. *Journal of Chinese Linguistics*, 69-90.
- Frauenfelder, U. H., & Peeters, G. (1990). Lexical segmentation in TRACE: An exercise in simulation. *Cognitive models of speech processing: Psycholinguistic and computational perspectives*, 552.
- Frauenfelder, U. H., & Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In *Yearbook of morphology 1991* (pp. 165-183). Springer.
- Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1219.
- Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of psycholinguistic research*, 31(1), 45-63.
- Friedrich, C. K., Alter, K., & Kotz, S. A. (2001). An electrophysiological response to different pitch contours in words. *NeuroReport*, 12(15), 3189-3191.
- Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Alter, K. (2004). Pitch modulates lexical identification in spoken word recognition: ERP and behavioral evidence. *Cognitive Brain Research*, 20(2), 300-308.
- Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Gunter, T. C. (2004). ERPs reflect lexical identification in word fragment priming. *Journal of Cognitive Neuroscience*, 16(4), 541-552.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1(2), 126-152.
- Fuhrhop, J.-H., & Endisch, C. (2000). *Molecular and supramolecular chemistry of natural products and their model compounds*. CRC Press.
- Fujisaki, H. (1983). Dynamic characteristics of voice fundamental frequency in speech and singing. In *The production of speech* (pp. 39-55). Springer.
- Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of phonetics*, 11(2), 149-175.
- Gao, Y.-z. (1980). Běijīng huà de qīngshēng wèntí, [The neutral tone in the Beijing dialect]. *Yuyan jiaoxue yu yanjiu*, 2, 82-98.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human perception and performance*, 22(1), 144.
- Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. *Psychophysiology*, 27(6), 627-640.
- Gonsalvez, C. J., & Polich, J. (2002). P300 amplitude is determined by target-to-target interval. *Psychophysiology*, 39(3), 388-396.
- Goodglass, H., & Wingfield, A. (1997). *Anomia: Neuroanatomical and cognitive correlates*. Academic Press.
- Gow, D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: evidence from priming. *Journal of Experimental Psychology: Human perception and performance*, 21(2), 344.
- Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of memory and language*, 29(2), 228-244.
- Grainger, J., & Holcomb, P. J. (2009). An ERP investigation of orthographic priming with relative-position and absolute-position primes. *Brain Research*, 1270, 45-53.
- Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychological science*, 17(12), 1021-1026.
- Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1992). Neighborhood frequency effects and letter visibility in visual word recognition. *Perception & Psychophysics*, 51(1), 49-56.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24(2), 95-112.

- Grosjean, F., & Gee, J. P. (1987). Prosodic structure and spoken word recognition. *Cognition*, 25(1-2), 135-155.
- Gussenhoven, C. (2004). The phonology of tone and intonation.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8(4), 439-483.
- Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, 38(11), 1531-1549.
- Hagoort, P., Wassenaar, M., & Brown, C. (2003). Real-time semantic compensation in patients with agrammatic comprehension: Electrophysiological evidence for multiple-route plasticity. *Proceedings of the National Academy of Sciences*, 100(7), 4340-4345.
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.
- Hayes, B. P. (1980). *A metrical theory of stress rules* [Massachusetts Institute of Technology].
- Henrich, K., Alter, K., Wiese, R., & Domahs, U. (2014). The relevance of rhythmical alternation in language processing: An ERP study on English compounds. *Brain and Language*, 136, 19-30.
- Hillyard, S. A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual review of psychology*, 34(1), 33-61.
- Hirst, D., & Espesser, R. (1993). Automatic modelling of fundamental frequency using a quadratic spline function.
- Hoa, M. (1983). La réciprocité en chinois contemporain. *Cahiers de Linguistique-Asie Orientale*, 12(2), 21-64.
- Hockett, C. F. (1958). A course in modern linguistics.
- Holcomb, P. J., Grainger, J., & O'rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*, 14(6), 938-950.
- Howie. (1976). *Acoustical studies of Mandarin vowels and tones* (Vol. 18). Cambridge University Press.
- Howie, J. M. (1974). On the domain of tone in Mandarin. *Phonetica*, 30(3), 129-148.
- Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of child language*, 27(1), 3-42.
- Huang, C. J., Li, Y. A., & Simpson, A. (2018). *The handbook of Chinese linguistics*. John Wiley & Sons.
- Hung, D. L., Tzeng, O. J., & Ho, C.-y. (1999). Word superiority effect in the visual processing of Chinese. *Journal of Chinese Linguistics Monograph Series*, 61-95.
- Hyman, L. M., & Tadaejou, M. (1976). Floating tones in mbam-nkam. *Studies in Bantu tonology*, 3, 59-51.
- Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human perception and performance*, 24(6), 1612.
- Inhoff, A. W., Starr, M. S., Solomon, M., & Placke, L. (2008). Eye movements during the reading of compound words and the influence of lexeme meaning. *Memory & Cognition*, 36(3), 675-687.
- Isel, F., & Bacri, N. (1999). Spoken-word recognition: The access to embedded words. *Brain and Language*, 68(1-2), 61-67.
- Isel, F., Baumgaertner, A., Thrän, J., Meisel, J. M., & Büchel, C. (2010). Neural circuitry of the bilingual mental lexicon: Effect of age of second language acquisition. *Brain and cognition*, 72(2), 169-180.
- Isel, F., Gunter, T. C., & Friederici, A. D. (2003). Prosody-assisted head-driven access to spoken German compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(2), 277.

- Isel, F., Hahne, A., Maess, B., & Friederici, A. D. (2007). Neurodynamics of sentence interpretation: ERP evidence from French. *Biological psychology*, 74(3), 337-346.
- Isel, F., & Shen, W. (2020). Perception of lexical neutral tones in mandarin compounds: Electroencephalographic evidence from an oddball paradigm. *Neuropsychologia*, 147, 107557. <https://doi.org/10.1016/j.neuropsychologia.2020.107557>
- Janssen, N., Bi, Y., & Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access for Mandarin Chinese and English compounds. *Language and Cognitive Processes*, 23(7-8), 1191-1223.
- Jarema, G. (2006). Compound representation and processing: A cross-language perspective. *The representation and processing of compound words*, 45-70.
- Jarema, G., Busson, C., Nikolova, R., Tsapkini, K., & Libben, G. (1999). Processing compounds: A cross-linguistic study. *Brain and Language*, 68(1-2), 362-369.
- Jassem, W., Morton, J., & Steffen-Batog, M. (1968). The perception of stress in synthetic speech-like stimuli by Polish listeners. *Speech analysis and synthesis*, 1, 289-308.
- Jones, A. (1962). Tables of stresses in three-layer elastic systems. *Highway research board bulletin*(342).
- Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological review*, 83(5), 323.
- Jones, M. R. (2010). Attending to sound patterns and the role of entrainment. *Attention and time*, 317, 330.
- Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, 94(2), 223-244.
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15(1), 98-110.
- Kaseda, Y., Miyazato, Y., Ogura, C., Nakamoto, H., Uema, T., Yamamoto, K., & Ohta, I. (1994). Correlation between Event-Related Potentials and MR Measurements in Chronic Alcoholic Patients. *Psychiatry and Clinical Neurosciences*, 48(1), 23-32.
- Kaukoranta, E., Sams, M., Hari, R., Hämäläinen, M., & Näätänen, R. (1989). Reactions of human auditory cortex to a change in tone duration. *Hearing research*, 41(1), 15-21.
- Kehayia, E., Jarema, G., Tsapkini, K., Perlak, D., Ralli, A., & Kadzielawa, D. (1999). The role of morphological structure in the processing of compounds: The interface between linguistics and psycholinguistics. *Brain and Language*, 68(1-2), 370-377.
- Kilborn, K. W., & Friederici, A. D. (1994). Cognitive penetrability of syntactic priming in Broca's aphasia. *Neuropsychology*, 8(1), 83.
- Knaus, J., Wiese, R., & Janßen, U. (2007). The processing of word stress: EEG studies on task-related components. Proceedings of the 16th International Congress of Phonetic Sciences,
- Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, 383(6597), 256-259.
- Kochanski, G., Shih, C., & Jing, H. (2003). Hierarchical structure and word strength prediction of Mandarin prosody. *International Journal of Speech Technology*, 6(1), 33-43.
- Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human perception and performance*, 35(3), 876.
- Kutas, M., & Dale, A. (2013). Electrical and magnetic readings of mental functions. In *Cognitive neuroscience* (pp. 197-242). Psychology press.
- Kutas, M., & Federmeier, K. D. (2009). N400. *Scholarpedia*, 4(10), 7790.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual review of psychology*, 62, 621.

- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.
- Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.
- Ladefoged, P., & Johnson, K. (2014). *A course in phonetics*. Cengage learning.
- Lai, C., Sui, Y., & Yuan, J. (2010). A corpus study of the prosody of polysyllabic words in mandarin chinese. *Speech prosody 2010-fifth international conference*,
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological review*, 106(1), 119.
- Lee, L., & Nusbaum, H. C. (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception & Psychophysics*, 53(2), 157-165.
- Lee, W.-S., & Zee, E. (2008). Prosodic characteristics of the neutral tone in Beijing Mandarin. *Journal of Chinese Linguistics*, 36(1), 1-29.
- Lehiste, I. (1970a). Suprasegmentals.
- Lehiste, I. (1970b). Suprasegmentals, Cambridge, Massachusetts & London, UK. In: MIT Press.
- Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *The Journal of the acoustical society of America*, 51(6B), 2018-2024.
- Lehiste, I. (1973). Rhythmic units and syntactic units in production and perception. *The Journal of the acoustical society of America*, 54(5), 1228-1234.
- Lehiste, I., & Peterson, G. E. (1959). Vowel amplitude and phonemic stress in American English. *The Journal of the acoustical society of America*, 31(4), 428-435.
- Leinonen, A., Grönholm-Nyman, P., Järvenpää, M., Söderholm, C., Lappi, O., Laine, M., & Krause, C. M. (2009). Neurocognitive processing of auditorily and visually presented inflected words and pseudowords: Evidence from a morphologically rich language. *Brain Research*, 1275, 54-66.
- Leitman, D. I., Sehatpour, P., Shpaner, M., Foxe, J. J., & Javitt, D. C. (2009). Mismatch negativity to tonal contours suggests preattentive perception of prosodic content. *Brain imaging and behavior*, 3(3), 284-291.
- Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42(1-3), 1-22.
- Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar* (Vol. 3). Univ of California Press.
- Li, X., Yang, Y., & Hagoort, P. (2008). Pitch accent and lexical tone processing in Chinese discourse comprehension: An ERP study. *Brain Research*, 1222, 192-200.
- Li, Z. (2003). *The phonetics and phonology of tone mapping in a constraint-based approach* [Massachusetts Institute of Technology].
- Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. *Brain and Language*, 61(1), 30-44.
- Libben, G. (2006). Why study compound processing? An overview of the issues. *The representation and processing of compound words*, 1-22.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology*, 54(5), 358.
- Lieberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic inquiry*, 8(2), 249-336.
- Lieberman, M. Y. (1975). *The intonational system of English* [Massachusetts Institute of Technology].
- Lieber, R. (2021). *Introducing morphology*. Cambridge University Press.
- Lin, M., & Yan, J. (1990). Neutral tone and stress in Mandarin Chinese. *Language Teaching and Linguistic Studies*, 3, 88-104.
- Lin, T. (1962). Xiandai Hanyu qingyin he jufa jigou de guanxi. (The relationship between the neutral tone and the syntax in modern Chinese). *Zhongguo Yuwen*, 7, 301-311.

- Lin, Y. (1972). *Chinese-English dictionary of modern usage*.
- Liu, S., & Samuel, A. G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Language and Speech*, 47(2), 109-138.
- Lu, B. (1981). Conversional words of verbs and nouns. *Lexicographical Studies*(1).
- Lu, S., & Ding, S. (2008). The Contemporary Chinese Dictionary. In
- Luck, S. J. (2005). Ten simple rules for designing and interpreting ERP experiments. *Event-related potentials: A methods handbook*, 4.
- Luo, H., Ni, J.-T., Li, Z.-H., Li, X.-O., Zhang, D.-R., Zeng, F.-G., & Chen, L. (2006). Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. *Proceedings of the National Academy of Sciences*, 103(51), 19558-19563.
- Luo, Y., Yan, M., Yan, S., Zhou, X., & Inhoff, A. W. (2016). Syllabic tone articulation influences the identification and use of words during Chinese sentence reading: evidence from ERP and eye movement recordings. *Cognitive, Affective, & Behavioral Neuroscience*, 16(1), 72-92.
- Luo, Y., & Zhou, X. (2010). ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading. *NeuroImage*, 49(3), 2836-2849.
- Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: behavioral and electrophysiological evidence. *Cerebral cortex*, 17(11), 2659-2668.
- Magne, C., Gordon, R. L., & Midha, S. (2010). Influence of metrical expectancy on reading words: An ERP study. *Speech Prosody 2010-Fifth International Conference*,
- Mäkisalo, J., Niemi, J., & Laine, M. (1999). Finnish compound structure: Experiments with a morphologically impaired patient. *Brain and Language*, 68(1-2), 249-253.
- Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, 23(2), 294-305.
- Marr, D., & Vaina, L. (1982). Representation and recognition of the movements of shapes. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 214(1197), 501-524.
- Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8(1), 1-71.
- Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human perception and performance*, 15(3), 576.
- Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition: A tutorial review. *Attention and performance: Control of language processes*, 125-150.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.
- Marslen-Wilson, W. D., & Tyler, L. K. (1981). Central processes in speech understanding. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 295(1077), 317-332.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10(1), 29-63.
- Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., & Lleras, A. (2010). Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pre-target entrainment. *Cognition*, 115(1), 186-191.
- Matthews, P. H. (1991). *Morphology*. Cambridge university press.
- Mattys, S. L. (2000). The perception of primary and secondary stress in English. *Perception & Psychophysics*, 62(2), 253-265.
- McCauley, S. M., Hestvik, A., & Vogel, I. (2013). Perception and bias in the processing of compound versus phrasal stress: Evidence from event-related brain potentials. *Language and Speech*, 56(1), 23-44.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.

- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, 88(5), 375.
- McQueen, J. M., & Cutler, A. (2010). Cognitive processes in speech perception. In *The handbook of phonetic sciences* (pp. 489-520). Blackwell.
- Melvold, J. L. (1989). *Structure and stress in the phonology of Russian* [Massachusetts Institute of Technology].
- Mertens, R., & Polich, J. (1997). P300 from a single-stimulus paradigm: passive versus active tasks and stimulus modality. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials Section*, 104(6), 488-497.
- Mikkola, K., Kushnerenko, E., Partanen, E., Serenius-Sirve, S., Leipälä, J., Huottilainen, M., & Fellman, V. (2007). Auditory event-related potentials and cognitive function of preterm children at five years of age. *Clinical Neurophysiology*, 118(7), 1494-1502.
- Mohanan, K. P. (2012). *The theory of lexical phonology* (Vol. 6). Springer Science & Business Media.
- Morton, J., & Jassem, W. (1965). Acoustic correlates of stress. *Language and Speech*, 8(3), 159-181.
- Munte, T. F., Heinze, H.-J., & Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience*, 5(3), 335-344.
- Myers, J. (2006). 1Processing Chinese Compounds: A Survey of the Literature.
- Myers, J., Libben, G., & Derwing, B. (2004). The nature of transparency effects in Chinese compound processing. Poster presented at the Fourth International Conference on the Mental Lexicon, Windsor, Canada,
- Näätänen, R. (1982). Processing negativity: An evoked-potential reflection. *Psychological bulletin*, 92(3), 605.
- Näätänen, R. (1991). Mismatch negativity outside strong attentional focus: A commentary on Woldorff et al.(1991). *Psychophysiology*, 28(4), 478-484.
- Näätänen, R., & Alho, K. (1997). Mismatch negativity—the measure for central sound representation accuracy. *Audiology and Neurotology*, 2(5), 341-353.
- Näätänen, R., & Gaillard, A. (1983). 5 The orienting reflex and the N2 deflection of the event-related potential (ERP). In *Advances in psychology* (Vol. 10, pp. 119-141). Elsevier.
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica*, 42(4), 313-329.
- Näätänen, R., & Michie, P. T. (1979). Early selective-attention effects on the evoked potential: a critical review and reinterpretation. *Biological psychology*, 8(2), 81-136.
- Näätänen, R., Paavilainen, P., & Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neuroscience Letters*, 107(1-3), 347-352.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, 118(12), 2544-2590.
- Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology*, 115(1), 140-144.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological bulletin*, 125(6), 826.
- Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3(2), 151-165.
- Niemi, J., Laine, M., & Tuominen, J. (1994). Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes*, 9(3), 423-446.
- Nordby, H., Hammerborg, D., Roth, W. T., & Hugdahl, K. (1994). ERPs for infrequent omissions and inclusions of stimulus elements. *Psychophysiology*, 31(6), 544-552.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189-234.

- Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of memory and language*, 31(6), 785-806.
- Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1991). Right hemisphere dominance of different mismatch negativities. *Electroencephalography and clinical neurophysiology*, 78(6), 466-479.
- Paavilainen, P., Karlsson, M.-L., Reinikainen, K., & Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography and clinical neurophysiology*, 73(2), 129-141.
- Packard, J. L. (1999). Lexical access in Chinese speech comprehension and production. *Brain and Language*, 68(1-2), 89-94.
- Peng, D., & Wang, C. (1997). Basic processing unit of Chinese character recognition: Evidence from stroke number effect and radical number effect. *Acta Psychologica Sinica*.
- Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human perception and performance*, 24(1), 134.
- Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1), 24.
- Pierrehumbert, J. (1981). Synthesizing intonation. *The Journal of the acoustical society of America*, 70(4), 985-995.
- Pisoni, D. B., & Remez, R. E. (2008). *The handbook of speech perception*. John Wiley & Sons.
- Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human perception and performance*, 16(3), 564.
- Polich, J. (2003). Theoretical overview of P3a and P3b. *Detection of change*, 83-98.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128-2148.
- Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological psychology*, 41(2), 103-146.
- Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*, 20(1-2), 261-290.
- Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human perception and performance*, 26(2), 820.
- Pulvermüller, F. (1996). Hebb's concept of cell assemblies and the psychophysiology of word processing. *Psychophysiology*, 33(4), 317-333.
- Pulvermüller, F., & Schumann, J. H. (1994). Neurobiological mechanisms of language acquisition. *Language learning*, 44(4), 681-734.
- Quené, H., & Kager, R. (1992). The derivation of prosody for text-to-speech from prosodic sentence structure. *Computer Speech & Language*, 6(1), 77-98.
- Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading*. Psychology Press.
- Reinisch, E., Jesse, A., & McQueen, J. M. (2010). Early use of phonetic information in spoken word recognition: Lexical stress drives eye movements immediately. *Quarterly Journal of Experimental Psychology*, 63(4), 772-783.
- Ren, G.-Q., Yang, Y., & Li, X. (2009). Early cortical processing of linguistic pitch patterns as revealed by the mismatch negativity. *Neuroscience*, 162(1), 87-95.
- Repp, B. H., & Liberman, A. M. (1987). Phonetic category boundaries are flexible.
- Rietveld, A. C., & Koopmans-van Beinum, F. J. (1987). Vowel reduction and stress. *Speech communication*, 6(3), 217-229.

- Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *NeuroImage*, 12(1), 14-19.
- Rodrigues Silva, D. M., & Rothe-Neves, R. (2020). Context-dependent categorisation of vowels: a mismatch negativity study of positional neutralisation. *Language, Cognition and Neuroscience*, 35(2), 163-178.
- Rosner, B. S., & Pickering, J. B. (1994). *Vowel perception and production* (Vol. 23). OUP Oxford.
- Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia*, 50(2), 232-244.
- Rothermich, K., Schmidt-Kassow, M., Schwartz, M., & Kotz, S. A. (2010). Event-related potential responses to metric violations: rules versus meaning. *NeuroReport*, 21(8), 580-584.
- Rushby, J. A., Barry, R. J., & Doherty, R. J. (2005). Separation of the components of the late positive complex in an ERP dishabituation paradigm. *Clinical Neurophysiology*, 116(10), 2363-2380.
- Rüsseler, J., Altenmüller, E., Nager, W., Kohlmetz, C., & Münte, T. F. (2001). Event-related brain potentials to sound omissions differ in musicians and non-musicians. *Neuroscience Letters*, 308(1), 33-36.
- Sabri, M., & Campbell, K. B. (2001). Effects of sequential and temporal probability of deviant occurrence on mismatch negativity. *Cognitive Brain Research*, 12(1), 171-180.
- Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, 90(1), 51-89.
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials Section*, 62(6), 437-448.
- Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *The quarterly journal of Experimental Psychology*, 42(3), 529-567.
- Sato, Y., Yabe, H., Hiruma, T., Sutoh, T., Shinozaki, N., Nashida, T., & Kaneko, S. (2000). The effect of deviant stimulus probability on the human mismatch process. *NeuroReport*, 11(17), 3703-3708.
- Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, 21(9), 1693-1708.
- Schreuder, R., & Baayen, H. (1995). Modeling morphological processing. *Morphological aspects of language processing*, 2, 257-294.
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in neurosciences*, 32(1), 9-18.
- Segui, J. (1992). Les composantes cognitives de la lecture. *La Lecture: processus, apprentissage, troubles*, 43-53.
- Segui, J. (2015). Évolution du concept de lexique mental. *Revue de neuropsychologie*, 7(1), 21-26.
- Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: effects of relative prime-target frequency. *Journal of Experimental Psychology: Human perception and performance*, 16(1), 65.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological review*, 96(4), 523.
- Selkirk, E. (1984). On the major class features and syllable theory. *Language sound structure*.
- Semenza, C., & Mondini, S. (2010). Compound words in neuropsychology. *New Impulses in Word-Formation*, 331-348.

- Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *The Journal of the acoustical society of America*, 106(2), 1078-1083.
- Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *The Journal of the acoustical society of America*, 107(5), 2697-2703.
- Shatzman, K. B., & McQueen, J. M. (2006). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception & Psychophysics*, 68(1), 1-16.
- Shei, C. (2014). Understanding the Chinese Language. In: Routledge.
- Shen, X.-n. S. (1990). *the prosody of Mandarin Chinese* (Vol. 118). Univ of California Press.
- Shen, X. S. (1992). Mandarin neutral tone revisited. *Acta linguistica hankniensia*, 24(1), 131-152.
- Shen, X. S. (1993). Relative duration as a perceptual cue to stress in Mandarin. *Language and Speech*, 36(4), 415-433.
- Shen, X. S., & Lin, M. (1991). Concept of tone in Mandarin revisited: A perceptual study on tonal coarticulation. *Language Sciences*, 13(3-4), 421-432.
- Shi, R. (1988). Shuo qingsheng (On the neutral tone). *Yuyan Yanjiu (Language Studies)*(1).
- Shillcock, R. (1990). Lexical hypotheses in continuous speech.
- Shoolman, N., & Andrews, S. (2003). Racehorses, reindeer, and sparrows. *Masked priming: The state of the art*, 241-278.
- Silva, D. M., Melges, D. B., & Rothe-Neves, R. (2017). N1 response attenuation and the mismatch negativity (MMN) to within-and across-category phonetic contrasts. *Psychophysiology*, 54(4), 591-600.
- Slowiacek, L. M., Nuabbaum, H., & Pisoni, D. (1984). Acoustic-phonetic priming in auditory word recognition. *Cognitive Psychology*.
- Sommer, W., Matt, J., & Leuthold, H. (1990). Consciousness of attention and expectancy as reflected in event-related potentials and reaction times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(5), 902.
- Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of memory and language*, 45, 412-432.
- Spinelli, E., McQueen, J. M., & Cutler, A. (2003). Processing resyllabified words in French. *Journal of memory and language*, 48(2), 233-254.
- Sproat, R., & Shih, C. (1987). Prenominal adjectival ordering in English and Mandarin. North East Linguistics Society,
- Squire, L. R., & Kandel, E. R. (1999). Brain system from declarative memory. *Scientific American Library, editor. Memory: From Mind to Molecules. New York*, 83-107.
- Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature neuroscience*, 2(2), 191-196.
- Su, Y.-C. (1998). The representation of compounds and phrases in the mental lexicon: Evidence from Chinese. *University of Maryland Working Papers in Linguistics*, 6, 179-199.
- Swinney, D. A., Onifer, W., Prather, P., & Hirshkowitz, M. (1979). Semantic facilitation across sensory modalities in the processing of individual words and sentences. *Memory & Cognition*, 7(3), 159-165.
- Tabossi, P., Collina, S., Mazzetti, M., & Zoppello, M. (2000). Syllables in the processing of spoken Italian. *Journal of Experimental Psychology: Human perception and performance*, 26(2), 758.
- Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7(4), 263-272.
- Taft, M. (1981). Prefix stripping revisited. *Journal of verbal learning and verbal behavior*, 20(3), 289-297.
- Taft, M. (1985). The decoding of words in lexical access: A review of the morphographic approach. *Reading research: Advances in theory and practice*, 5, 83-123.
- Taft, M. (1994). The influence of character frequency on word recognition responses in Chinese. *Advances in the study of Chinese language processing*, 1, 59-73.

- Taft, M. (2003). Morphological representation as a correlation between form and meaning. In *Reading complex words* (pp. 113-137). Springer.
- Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A*, 57(4), 745-765.
- Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of verbal learning and verbal behavior*, 14(6), 638-647.
- Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of verbal learning and verbal behavior*, 15(6), 607-620.
- Taft, M., & Hambly, G. (1986). Exploring the cohort model of spoken word recognition. *Cognition*, 22(3), 259-282.
- Taft, M., Li, S., & Beyersmann, E. (2018). What cross-morphemic letter transposition in derived nonwords tells us about lexical processing. *Journal of Cognition*, 1(1).
- Taft, M., Liu, Y., & Zhu, X. (1999). Reading Chinese script: A cognitive analysis. In (pp. pp. 91-114): Lawrence Erlbaum Associates Hillsdale, NJ.
- Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(3), 761.
- Taft, M., Zhu, X., & Peng, D. (1999). Positional specificity of radicals in Chinese character recognition. *Journal of memory and language*, 40(4), 498-519.
- Tagliapietra, L., & Tabossi, P. (2005). Lexical stress effects in Italian spoken word recognition. *Proceedings of the XXVII annual conference of the cognitive science society*.
- Tang, P. (2014). A study of prosodic errors of Chinese neutral tone by advanced Japanese students (Chinese version). *TCSOL Studies*, 56(4), 39-47.
- Taylor, P. (2000). Analysis and synthesis of intonation using the tilt model. *The Journal of the acoustical society of America*, 107(3), 1697-1714.
- Tervaniemi, M., Jacobsen, T., Röttger, S., Kujala, T., Widmann, A., Vainio, M., Näätänen, R., & Schröger, E. (2006). Selective tuning of cortical sound-feature processing by language experience. *European Journal of Neuroscience*, 23(9), 2538-2541.
- Toivainen, P., Tervaniemi, M., Louhivuori, J., Saher, M., Huotilainen, M., & Näätänen, R. (1998). Timbre similarity: Convergence of neural, behavioral, and computational approaches. *Music Perception*, 16(2), 223-241.
- Tsai, C.-H. (1994). Effects of semantic transparency on the recognition of Chinese two-character words: Evidence for a dual-process model. *Unpublished master's thesis, National Chung Cheng University, Chia-Yi, Taiwan*.
- Tsang, Y.-K., Jia, S., Huang, J., & Chen, H.-C. (2011). ERP correlates of pre-attentive processing of Cantonese lexical tones: The effects of pitch contour and pitch height. *Neuroscience Letters*, 487(3), 268-272.
- Van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology Section A*, 58(2), 251-273.
- van Heuven, V. J., & Menert, L. (1996). Why stress position bias? *The Journal of the acoustical society of America*, 100(4), 2439-2451.
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, 18(4), 380-393.
- Vroomen, J., & De Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human perception and performance*, 21(1), 98.
- Vroomen, J., & De Gelder, B. (1997). Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human perception and performance*, 23(3), 710.
- Wai-Sum, L., & Eric, Z. (2014). Chinese phonetics. In *The handbook of Chinese linguistics* (pp. 369-399). Wiley-Blackwell.
- Wang, J., Friedman, D., Ritter, W., & Bersick, M. (2005). ERP correlates of involuntary attention capture by prosodic salience in speech. *Psychophysiology*, 42(1), 43-55.

- Wang, W., Lu, A., He, D., Zhang, B., & Zhang, J. X. (2017). ERP evidence for Chinese compound word recognition: does morpheme work all the time? *NeuroQuantology*, 15(3).
- Wang, Y. (1996). Discussing the pitch of neutral syllables. *Chinese Teaching in the world*, 3, 50-53.
- Wang, Y. (2004). The effects of pitch and duration on the perception of the neutral tone in standard Chinese. *Acta Acustica*.
- Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., Aaltonen, O., Raimo, I., Alho, K., & Lang, H. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, 7(3), 357-369.
- Winkler, I., Reinikainen, K., & Näätänen, R. (1993). Event-related brain potentials reflect traces of echoic memory in humans. *Perception & Psychophysics*, 53(4), 443-449.
- Wong, K. F. E., & Chen, H.-C. (1999). Orthographic and phonological processing in reading Chinese text: Evidence from eye fixations. *Language and Cognitive Processes*, 14(5-6), 461-480.
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of phonetics*, 25(1), 61-83.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of phonetics*, 27(1), 55-105.
- Xu, Y. (2004). The PENTA model of speech melody: Transmitting multiple communicative functions in parallel. *Proceedings of from sound to sense*, 50, 91-96.
- Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech communication*, 46(3-4), 220-251.
- Xu, Y., Lee, A., Prom-on, S., & Liu, F. (2015). Explaining the PENTA model: a reply to Arvaniti and Ladd. *Phonology*, 32(3), 505-535.
- Xu, Y., Santitham, P.-o., & Liu, F. (2022). The PENTA model: Concepts, use and implications. *Prosodic theory and practice. The MIT Press. Search in*.
- Xu, Y., & Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech communication*, 33(4), 319-337.
- Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport*, 8(8), 1971-1974.
- Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology*, 97(2), 259-268.
- Yan, M., Luo, Y., & Inhoff, A. W. (2014). Syllable articulation influences foveal and parafoveal processing of words during the silent reading of Chinese sentences. *Journal of memory and language*, 75, 93-103.
- Yang, S. (1991). The synthetic rules for neutral tone in standard Chinese. *J. Applied Acoustics*, 10(1), 12-18.
- Yip, M. J. (1980). *The tonal phonology of Chinese* [Massachusetts Institute of Technology].
- Yip, M. J. (2002). *Tone*. Cambridge University Press.
- Yu, B., & Cao, H. (1992). An investigation of the effect of stroke number on the identification of Chinese characters and a discussion of the effect of stroke frequency. *Acta Psychologica Sinica*.
- Zhang, B., & Peng, D. (1992). Decomposed storage in the Chinese lexicon. In *Advances in psychology* (Vol. 90, pp. 131-149). Elsevier.
- Zhang, J.-S., & Kawanami, H. (1999). Modeling carryover and anticipation effects for Chinese tone recognition. Sixth European Conference on Speech Communication and Technology.
- Zhang, L., Xi, J., Wu, H., Shu, H., & Li, P. (2012). Electrophysiological evidence of categorical perception of Chinese lexical tones in attentive condition. *NeuroReport*, 23(1), 35-39.
- Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. i. (2005). Effects of language experience: neural commitment to language-specific auditory patterns. *NeuroImage*, 26(3), 703-720.
- Zhao, G. (2015). *The contemporary Chinese dictionary*. In: Oxford University Press.

- Zhou, & Marslen-Wilson. (1999a). The nature of sublexical processing in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 819.
- Zhou, & Marslen-Wilson. (1999b). Pseudohomophone effects and lexical processing in reading Chinese. *Manuscript submitted for publication*.
- Zhou, Marslen-Wilson, Wang, J., Inhoff, A., & Chen, H. (1999). Reading Chinese script: A cognitive analysis. In: Lawrence Erlbaum Associates Hillsdale, NJ.
- Zhou, X., & Marslen-Wilson, W. (1994). Words, morphemes and syllables in the Chinese mental lexicon. *Language and Cognitive Processes*, 9(3), 393-422.
- Zhou, X., & Marslen-Wilson, W. (1995). Morphological structure in the Chinese mental lexicon. *Language and Cognitive Processes*, 10(6), 545-600.
- Zhou, X., & Marslen-Wilson, W. (1997). The abstractness of phonological representation in the Chinese mental lexicon. *Cognitive Processing of Chinese and other Asian languages*, 3-26.
- Zhou, X., & Marslen-Wilson, W. (2000). Lexical representation of compound words: Cross-linguistic evidence. *Psychologia*, 43(1), 47-66.
- Zhou, X., & Marslen-Wilson, W. D. (2009). Pseudohomophone effects in processing Chinese compound words. *Language and Cognitive Processes*, 24(7-8), 1009-1038.
- Zhou, X., Taft, M., & Shu, H. (1995). Morphological, orthographic, and phonological structure in the Chinese mental lexicon. seventh international conference on the cognitive processing of Chinese and other Asian languages. Hong Kong.
- Zhu, X., & Taft, M. (1994). The influence of perceptual experience on Chinese character processing. *Advances in the study of Chinese language processing*, 1, 85-99.
- Zsiga, N. (2012). *Indirect boosting for turbocharged SI engines* [ETH-Zürich].
- Zwitserslood, P. (1996). Form priming. *Language and Cognitive Processes*, 11(6), 589-596.