Chapter 1 Introduction

In the world we live, many sounds surround us and help us perceive the world. Among these sounds, the voices are not individual sounds but complex combinations of different sources. Though there are various sounds in the world, only a small part of them are used in languages and we call them speech sounds. Speech sounds are regarded as a continuous flow of sounds, and they relate the complexity and organization of articulators. The difference between speech sound and noise is that speech sounds are meaningful and important elements to form languages whereas noises are not.

In languages, pronunciation differences are crucially distinctive. Speakers organize the differences or contrasts as 'making different words.' The minimal triplets 'think', 'sink', and 'pink' and 'thick', 'sick', and 'pick' in English prove that the /θ/, /s/ and /p/ sounds are contrastive or distinctive in English. Therefore, the three sounds are phonemes in English but do not need to be contrastive phonemes in other languages. For instance, the /l/ and /r/ sounds which are contrastive in English and Mandarin Chinese are not contrastive in Japanese. In Japanese, /r/ and /l/ sounds are allophones of the same phoneme.

1.1 Study Background and Motivation

Our study aims at figuring out the importance of the two phonetic levels, segmental and supra-segmental elements, to the lexical meaning recognition process. Intuition knowledge tells us that lexicon recognition is an effortless process in speech-- we hear the speech sounds and then access the lexical memory directly. However, there are many detailed and complex auditory cues which affect the process of word recognition. Theoretically, speech sounds can be subcategorized into two dimensions: segmental elements (small units) and supra-segmental elements (large units). Segmental elements fit an array of timing slots, each supporting a cluster of distinctive features (Roca and Johnson, 1999). Segmental features usually relate to the place and manner of articulations. The place and manner of articulations in vowels and consonants are segmental features in languages whereas supra-segmental elements involve pitch variations to form tones and intonations distinctions. They require the model with a tone level independent of the segmental levels, or cluster of planes (Roca and Johnson, 1999). Acoustically, the most important cue to distinguish supra-segmental features is the fundamental frequency, or F0. Tone and intonation can also refer to as the prosodic features. Prosody functions linguistic and nonlinguistic levels. For the later one, prosodies are features of voice quality, but for the linguistic level they are elements in specific language systems, such as stresses and tones.

1.2 Literatures of Segmental and Suprasegemntal Perceiving

Previous studies indicated that segmental elements are the crucial and determining cues for word recognition process, such as Cutler's research in English (1986). The supra-segmental elements like stress did not facilitate to the word identification. In these kinds of languages, the stress patterns are already included in the segmental unit. That is, the unstressed syllable usually co-occurs with reduced vowel; hence, the segmental cues are sufficient information for the language users recognizing lexical items. However, the result is not the same as the Spanish (Soto-Faraco, 2001) and

German (Freidrich et al, 2004), which are also subbranches of Indo-European languages. In Spanish and German, the stress mismatched primes showed inhibitions to lexical recognition. Another study relating to Cantonese, a dialect of Chinese and of tone languages, also revealed the fact that both segmental and tonal cues are computable information for the process of semantic retrieval (Schirmer et al, 2005). The previous researches indicated the fact that the contribution of segmental and supra-segmental elements was language dependent. The previous studies indicated that people use the cues differently based on their inner language pattern. Based on these findings, we were interested in discussing the effect of inner phonetic factors (language background and language difference) to the semantic recognition.

There are six chapters in this study and two experiments were included. The research was examined by behavioral experiment and ERP experiments in Chapter 4 and Chapter 5, and we introduced the usage and data analysis methods of the ERP tool in Chapter3. The general discussions were in Chapter 6. Our expectation was that the supra-segmental information, the tonal features in our study, was equally contributed to the Mandarin users.

Chapter 2 Literatures Review

2.1 Language and the Brain

2.1.1 Functional Divisions of the Cerebral Cortex

According to Gazzaniga and his colleges' (1998) search, the cerebral cortex consists of two symmetrical hemispheres, and each of them contains large sheets of layered neurons. In higher mammals and humans, the cerebral cortex consists of many infoldings which are further defined as sulci (the enfolded regions) and gyri (the crowns of the folded tissue). The sections of cortex have a variety of function in neural processing. Basically, the divisions of cerebral cortex can be briefly categorized as five areas. The five areas are frontal lobe, parietal lobe, occipital lobe, temporal lobe, and central sulcus. The frontal lobe plays the major role of planning movements. It is located in front of the central sulcus. The parietal lobe is in charge of the sense and feeling of the bodies, and it is the somatosensory area. The regions receive somatosensory information and represent information about touch, pain, temperature sense, and so forth. The visual processing areas are in the occipital lobe, which locates in the post area of the brains. The primary visual cortex receives visual inputs from the lateral nucleus of the thalamus. The auditory processing areas lie in the superior part of the temporal lobe, which places on the lateral sulci. The specific position of different functions in the cortex was expressed in Figure 1.

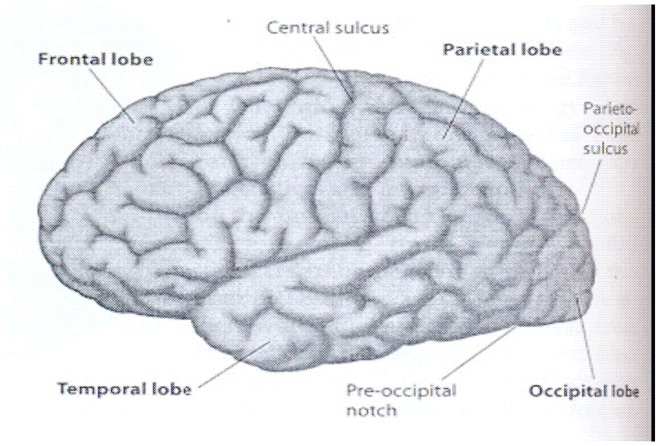


Figure 1: Four lobes of the cerebral cortex, in lateral view of the left hemisphere (Gazzaniga et al, 1998)

2.1.2 The Relationship of Brain Activities and Language

There are some brain waves which relate to language processing and are very important index to help observe the language processing activity in the brain. They are N100, N400, MMNm, and SPS, to name just a few. Steven Hillyard and his colleagues (1973) discovered the brain wave related to auditory activities. The N400 wave was found by Marta Kutas and Steven Hillyard (1980). They found that auditory ERPs were enlarged in amplitude when stimuli were attended comparing to when they were ignored. This effect in the ERPs was known as the auditory N1 potential because it is the first large, negative deflections in signal-averaged waveform. The subsequent study by Marty Woldorff and Steven Hillyard's (1993) combination of ERPs and ERFs was the earliest modulations of auditory inputs take place in the auditory cortex in Heschl's gyri.

The other important brain wave, N400 response, is semantic specific. Kutas and Hillyard (1980) found the N400 effect by comparing the processing of last word of sentences in three situations: the normal sentence, anomalous sentence, and physically deviant sentence. The sentence with anomalous word in ending show obvious increase during 400 ms and the difference of the amplitude is called N400 effect.

In addition, there is a special brain wave ‘MMN’ which occurs when the context is beyond expectation like N400. But different from the N400, MMN is phonological unique. The critical feature differs between the two brain waves is the brain responses. According to Connolly and Phillips’ study (1994), MMN was isolated from N400. MMN was elicited in the phonological mismatch condition but not the semantic violation. In contrast, N400 occurs in semantically incongruent word as we mentioned previously but MMN doesn’t. When the sentences contained both phonological and semantic violation in the sentence-ending word, both N400 and MMN were detectable in different point of time.

Finally, there are less electrophysiological researches correlating to other linguistic analysis such as syntax. Dutch psychologists Peter Hagoort, Colin Brown, and their colleagues (1993) discovered increasing amplitude in positive polarity when grammatical violation occurs in reading. The positive component was regarded as the syntactic positive shift (SPS).

2.2 The ERP Tool

Language is one kind of cognitive activity, and it relates to numerous brain neurons activation. In order to examine the relationship between language and brain activities, there are many researches using brain detective tools to examine their studies, such fMRI, PET, MEG, and ERP. Among these neuroscience equipments, ERP (event-related potentials) is a very convenient tool to help record the brain activities. The advantage of ERP is its sensitivity of time window. Its sensitivity can record brain activities in 1 msec. Since our study relates to negativity, and it is also a time-sensitivity experimental design, we chose the ERP tool to record the brain wave in our study. Before introducing the ERP, we have to know what the EEG (electroencephalogram) is. The term 'EEG' represents the electrical activity of the brain, and it is popularly known as *brain waves*. The scalp EEG is an average of multifarious activities of many small zones of the cortical surface beneath the electrodes (Rangayyan, 2002). The EEG patterns are associated with various physiological and mental processes, and they include spikes, transients, or other waves and patterns associated with various nervous disorder.

ERP (event-related potential) is an extension of EEG. By the recording or marking on specific events or stimuli, we can time-lock the events and then get the results of brain activations after filtering out the noises. The term *event-related potential* is more general than the term *evoked potential*. Short-latency ERPs are dependent on the physical characteristics of the stimuli, but long-latency ERPs are predominantly influenced by the conditions of stimuli presentation.

2.5 Researches of Neuro-linguistics

For many years, the interhemispheric researches have been noticed because the functional specialization of cortical areas relates to different mental activities and behaviors. Figuring out these different correlations in our brains helps people discover the inner mental activities of human beings. Like languages, music also involves high level of mental activities and is highly relevant to auditory information.

There was research indicating that music stimuli and linguistic stimuli cause different hemispheric activation. For example, Tervaniemi et al (1999) investigated the functional specialization of auditory cortex in processing phonetic and musical sounds. The study detected the reaction of MMN response in the condition of frequent and infrequent phonemes (/e/ and /o/ in Finish) or chords (A major and A minor) by MEG. The result represented that chord changes are processed more strongly than phoneme changes within right hemisphere while in the left hemisphere there is no such difference. Moreover, this study further confirms the importance of the posterior areas of the left temporal lobe in processing phonetic information.

As we mentioned in the previous paragraph, language process can be regarded as an activity which involves functional specification of cortex; moreover, language experiences also show differences in brain processing. Zhang and Kuhl and their colleagues (2005) compared the brain activities between adult American and Japanese listeners' phonetics processing with MEG. This study showed that processing non-native speech sounds recruited more neural populations in both hemispheres and required longer period of brain activation in two areas: the superior temporal area and the inferior parietal areas. They further claimed that early language experiences produce a 'neural commitment' to the acoustic properties of that language, and this neural commitment became less efficient by interferences of the unfamiliar or foreign languages.

In addition, the following researches represented more details of relationships between language processing and cortical areas. Recently, different vowels have been hypothesized to elicit distinct spatio-temporal patterns of cortical activity.

Makela et al (2003) used MEG to elicit N1m response of vowel /i/, /u/, and /o/. The results showed that the change in speech F0 is not accompanied by location shifts of the N1m source. The data confirm that the categorical perception of vowels might be explained by the mapping of the vowels F1 and F2 in organized cortical areas.

There are still some researches that discussed the recognition of lexical access from the behavioral and neuropsychological evidences. In Cutler's research (1986) the supra-segmental information is largely redundant for word discrimination in English because the stress pattern differences between words involve vocalic differences. That is, the stressed syllable always co-occurs with full vowels but unstressed syllable co-occurred with reduced vowel, and thus, supra-segmental features facilitate little to those listeners. However, the opposite result was found from other languages. Soto-Faraco et al (2001) used four cross-modal priming experiments to addressed the role of segmental and supra-segmental information in the activation of spoken words. The responses were recorded by comparing the prime fully matching with the target and the prime partially mismatching with the target at the single element (stress pattern, one vowel, and one consonant, respectively). The results indicate that full matching primes facilitate lexical decision responses while mismatching primes produced inhibition. The different contributions between supra-segmental elements (stress) and segmental elements (vowels and consonants) do not reach statistic differences and the fact shows that supra-segmental elements do contribute during word recognition process even in the intonation languages like Spanish. Furthermore, the study also showed that pitch is used for lexical identification in spoken word recognition in German (Freidrich et al, 2004).

The opposite findings to English were not only found in the intonation languages. Related study from Cantonese, a dialect of Chinese and a tone language, also showed the difference from English. Schirmer et al (2005) examined the role of tone and segmental information in Cantonese word processing by the event-related potential (ERP). There were differences between completely matched words and the other two violation conditions (either in tone or in segmental level) with respect to the latency and amplitude of the ERP effects. But, the tonal and segmental elements that induced semantic violations were comparable. The findings suggested tonal and segmental information are accessed at a similar point and both type of information play the same role during Cantonese word-recognition.

In sum, we have learned that language experiences affect the learning to a foreign language. However, there was no sufficient discussion in the issue of the interaction of acoustical information and speech comprehension in Mandarin. Besides, previous studies examined the difference of acoustical effect in speech perception through the results of behavioral experiments and there were little studies discussing the brain reaction to phonetic information in Mandarin. That is, the previous studies told us that the two levels of phonetic cues contributed differently when semantic violations occurred—some languages showed asymmetry between the segmental and supra-segmental contribution, such as English (Culter, 1986), but other languages showed equivalent contribution between the two levels of phonetic cues, such as Spanish and Cantonese.

In our study, we discussed the interaction of speech signal and brain activities in semantic word recognition process. We followed the method of Schirmer et al (2005) but there were still some differences between their previous research and our study. In Schirmer's study, the role of tone and segmental information in Cantonese word processing were examined by the ERP method. The participants were asked to listen to sentences that were either semantically correct or semantically violated words. There were four sentence conditions in the study, one of semantically correct and three of semantic incorrect. Semantically incorrect words differed from the expected words at the tonal level (the TV group), the segmental level (the SV group), and at both levels (the CV group).

Similar to the Cantonese research, there were four sentence conditions in our study, but we further distinguish the segmental violations into two categories: the consonant violation and vowel violation. Furthermore, in order to examine the specific relationship between speech sounds and word recognition, we allowed only one violation in every incorrect sentence; therefore, there was no complete violation (a trial with more than one mismatching) in our study.

Chapter 3 Methodology

In this chapter, we introduced the method that our thesis based on and we divided the chapter into two sections. They were Experimental Procedure and Data Analysis. In the section of Experimental Procedure, we illustrated the process of the whole experiment; in addition, the equipments and software we used to record the data were introduced in the section of Data Analysis.

3.1 Experimental procedure

There were two tasks that the participants were asked to do during Experiment 1. They were asked to concentrate on listening to the sentence in the step 1 and then make judgment in step 2. All participants were introduced and seated in a comfortable chair in a sound-proofed chamber during the whole experiment. Auditory sentences were presented over headphones during the experiment. Every trial included two steps, the first is the brain waves detection and the second was the behavioral reaction. In the first step of the experiment, the participants saw a fixation, the white-eye icon on the black-background screen, and the auditory trial came out through the headphone at the same time. Then, after the sentence finished, the experiment went to the judgment section.

The second icon was with an instruction 請按鈕, meaning "*please press the button.*" The participants were asked to make a judgment when the participants saw the instruction on whether the sentence they heard made sense or not. If they thought it was right, they pressed the right button on the controller; if not, they pressed the left button.

They were also asked to make the judgment as soon and as correct as possible. Next trial started after the button was pressed or 2 seconds after. The ISI (inter stimuli interval) lasted for 2 seconds. The whole experiment lasted for 40 minutes and the experiment paused every 15 minutes and the subjects can take a break. Before the experiment started again, the screen shows the icon 實驗即將開始 expressing the information "The experiment is going to start again."

The subjects were informed the preparation details before joining this activity such as sufficient sleep and hair-washing at the day before they attended the experiment. The instruction had been provided to the subjects on the paper before the experiment started. Because the brainwave is minor to most of biological signals, we have to avoid this confounding as well. Therefore, the subjects were asked not to move their body or blink when the auditory sentences were presented. The flowchart of the experiment is shown in Figure 2 below.

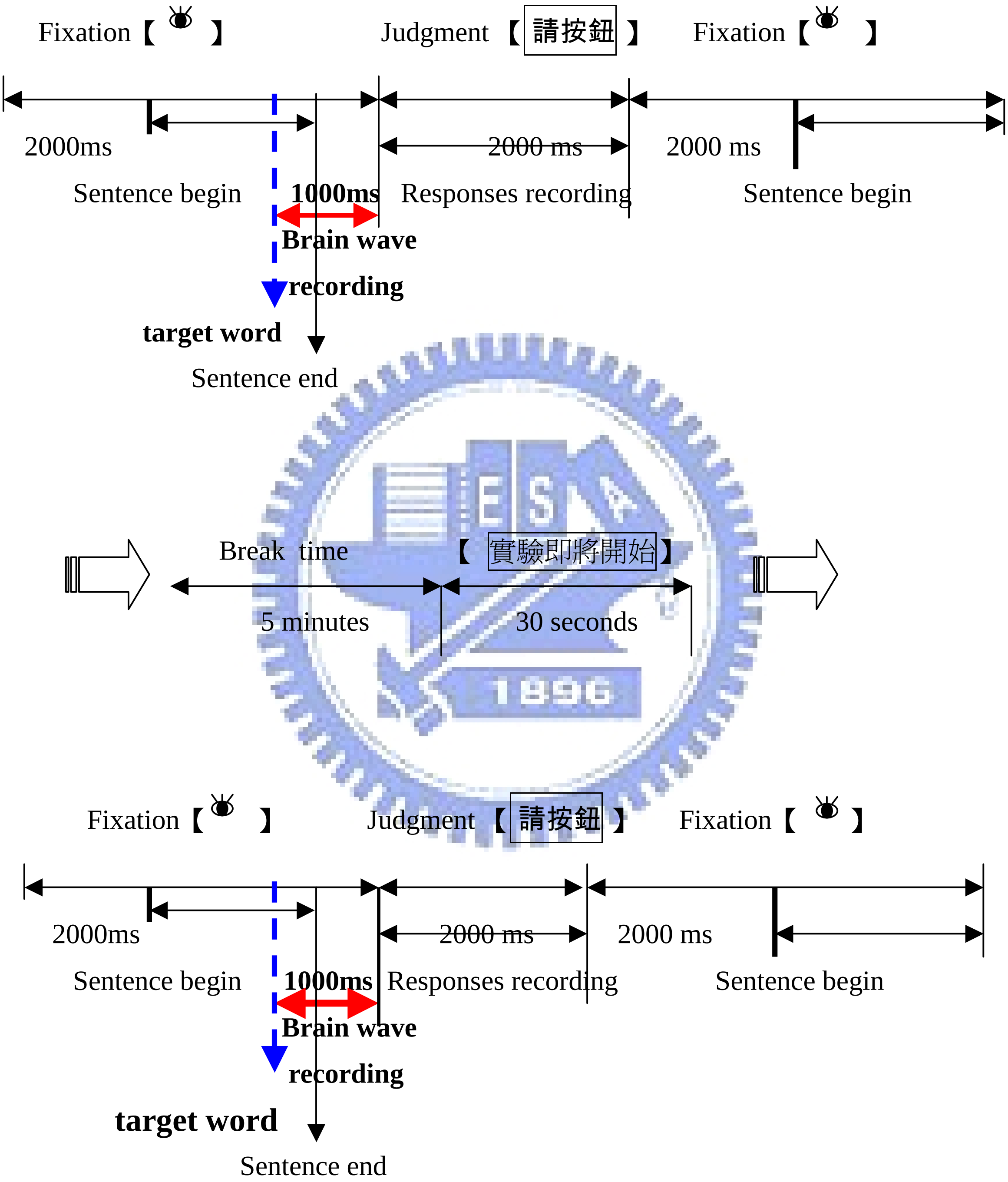


Figure 2: The procedure of the experiment

3.2 Data analysis

3.2.1 Behavioral results

Subjects who participated this experiment were asked to make a judgment after they listened to the previous sentence. Their responses were recorded and we calculated the rate of matching the target and rate of failure among the four conditions.

3.2.2 ERP data analysis

After data collection, we had to further select the EEG signals. As we mentioned previously, there are many environmental noises or other biological signals which were regarded as interferences during data analysis. Those unexpected interferences would affect the correctness and the result of this experiment. To avoid the interference, we need to refine the recorded data further. We used the software Neuro Scan to conduct the data screening and data analysis process. The whole process can be divided into 8 stages: (Delorem et al, 2004)

There are different events in our experiment. In this experiment, we designed different conditions to observe the brain wave difference. Also, we asked subjects to make judgment after the auditory trial finished and recorded the events and response simultaneously. These different events of ERP experiment and the responses of the behavior experiment were coded. Then, we further excluded the invalid data in order to avoid the incorrect results. We name those selected data as 'event' file.

3.2.2.1 Epoching

We selected certain events from the event files at the stage and tagged the represented code of these event types before the experiment and then extracted the different events according to the tags. The criteria for extracting the events were to average time window of -100 to 1000 msec from word onset.

3.2.2.2 Baseline Correction

The step of baseline correction is to equate the effects across electrode sites. It is important to select an appropriate baseline because any noise in the baseline would add noise to our measures. Because of the pre-stimulus activities, the carrier sentences, are not different across experiment conditions (only the stimuli themselves cause condition differences), we assume that the brain waves during this period is unaffected by the stimulus. We chose pre-stimulus brain wave as the baseline (100 ms before stimulus onset) according to our experimental design.

3.2.2.3 Artifact Rejection

We removed the potentially confounding by the step of artifact rejection. First, we made the blinking correction to the channel HEOG and VEOG. If the amplitude was beyond the range of ±100μV, we regarded it as blinking and excluded the signal. Second, we modified other electrodes. All electrodes except for HEOG and VEOG were selected and we also conducted the artifact rejection. At the stage, amplitudes beyond ±80μV were regarded as noise and excluded.

3.2.2.4 Filtering

After the artifact rejection, we further reduced noise with filters. Most of the relevant parts of the ERP waveform in cognitive neuroscience experiments consist of frequencies between 0.01Hz and 30 Hz. We set the low pass filter at 30 Hz and 12dB/oct.

3.2.2.5 Average

We averaged EEG files to AVG files at the stage. We selected and averaged the same event types of each subject to see the brain wave data. The averaging is a measure of central tendency for signals. It minimized the EEG noise, or increased the signal-to-noise ratio (SNR).

3.2.2.6 Group Average (Grand Average)

At this step, data from each of the four experimental conditions were averaged across subjects and formed the grand average files.

So far we have introduced the details and the procedure of our study. Because the process of ERP experiments is complicated and the method we chose affects the results and reliability of the study, we have to take care of every detail and the experiment design. The above stages were the processes we followed when we conducted the experiments and two experiments of the study will be introduced in Chapter 4 and Chapter5.

Chapter 4 Experiment 1

4.1 Subject

Ten native Mandarin speakers participated this experiment (four females and six males). They speak fluent Taiwan Mandarin and are skilled in Mandarin reading, writing, listening, and writing. Some of them can also use Taiwanese, which is one of the most dominant dialects in Taiwan and also a tone language. All of them grew up in Taiwan and accepted college education. They used Mandarin as the daily language, too. To ensure the correct result for this auditory experiment, participants should have normal hearing ability and none of them have the problem of brain impairment. They did not have a self-reported history of speech or hearing problems. All participants were voluntary to participate this experiment and were paid after the experiment finished.

4.2 Materials

In the experiment, four conditions of materials were manipulated. Each condition contained 72 sentences. The sentences were presented in Mandarin and the target word was presented at the end of each sentence. The differences between four conditions were of the last word of each sentence; they may either match or violate the correct word in the auditory way. There were three kinds of violations—the tonal, vowel, and consonantal violations. We further categorized the violations into segmental violation and supra-segmental violations. Tonal violation was regarded as supra-segmental violation; besides, the vowel and consonantal violations were included in the segmental violation. The examples were shown in Table 1.

Table 1: The sentence 'Xiao3-ming2 yin1 wei4 gan4 mao4 liu2 bi2 ti4',

***Xiao3-ming2 sniveled because of the flu*, in four conditions**

Conditions	Sentences
1. Semantic Congruence (SC)	Xiao3-ming2 yin1 wei4 gan4 mao4 liu2 bi2 **ti4**
2. Tonal Violation (TV)	Xiao3-ming2 yin1 wei4 gan4 mao4 liu2 bi2 **ti2**
3. Vowel Violation (VV)	Xiao3-ming2 yin1 wei4 gan4 mao4 liu2 bi2 **tu4**
4. Consonant Violation (CV)	Xiao3-ming2 yin1 wei4 gan4 mao4 liu2 bi2 **ki4**

Condition 1, the semantic congruous (SC) condition comprised with 72 sentences. They were consisted of four tones, eighteen consonants, and four kinds of vowels, and totally 72 sentences were presented. The four tones were high-level, rising, fall-rising and high falling, marking as Tone 1, Tone 2, Tone 3, and Tone 4, respectively. According to the articulation manner, the eighteen consonants can be further subcategorized as including stops /p/, /p^h/, /t/, /t^h/, /k/ and /k^h/, affricatives /ts/, /ts^h/, /tɕ/ and /tɕ^h/, fricatives /f/, /s/, /ɕ/ and /x/, nasals/m/ and /n/, liquid /l/, and retroflex /r/. Vowels included three cardinal vowels /i/, /u/, /a/ and a combination of diphthong or complex vowel, such as /ai/ or /iao/ according to the real situation in which they co-occur with the onset consonant without violating the syllable formation rules. The Condition 1 was the controlled group and all the sentences in this group did not result in any incongruous meaning because of phonetics mismatch.

Target syllable of condition 2, the tonal violated (TV) condition, differed from correct targets of the SC condition on tonal mismatch in last syllable. Sentences in this group violated semantic congruency on the supra-segmental level. For example, the word 估 with high-level tone 1 and pronounced as /gu/ with Tone 1 in sentence ‘到了年終許多企業要提出績效評估’ (*many companies have to evaluate their profit at the end of a year)* changed into high-falling tone 4 /gu/. Each of the four tones had their variants of the other three tones. For example, tone 1, the high-level tone, may be changed into tone 2, tone 3, or tone 4, respectively.

Target words in condition 3 violated the vowel correctness in condition 1 and the condition 3 was designed as vowel violated (VV) condition. The same as the SC condition, the VV condition included 72 sentences. But, the last word of each sentence did not match the correct word with vowels. For instance, the word 估 with the back lip-round vowel /u/ and pronounced as /gu/ (Tone1) in the sentence ‘到了年終許多企業要提出績效評估’ (*many companies have to evaluate their profit at the end of a year)* was changed into front lip-split [i] and pronounced as Tone 1 /gi/.

In condition 4, the consonant violated (CV) condition, we again presented 72 sentences but the target word at the end of the sentence mismatched semantic congruence in consonants. The features that differentiated consonant were manipulated as place of articulation, manner of articulation, voicing, and so on. In this experiment, we designed materials in CV condition as mismatched condition of consonant feature, but we did not focus on the difference of minimal difference. That is, we focus on both mono and multiple consonantal mismatches. Therefore, the trials in CV condition consisted of mono-feature and multi-feature mismatch sentences. The examples of four conditions were already shown in Table 1. (c.f. Appendix I for more details)

4.3 Behavioral Results

The results of the behavior experiment were accordant with our expectation. As our expectation, subjects misjudged more targets when listening to violation conditions comparing to the semantic congruent situation. The results showed that consonantal violation made subjects make most mistakes (wrong judgment: 57%). It was speculated that the participants used guessing strategies when they listened to words of consonantal violations. Besides, tonal violations also caused more mistakes (wrong judgment: 32%) comparing to the semantic congruous and tonal violated conditions. In the three mismatch conditions, vowel violations triggered least wrong judgments than the other two violations. From the results of behavioral experiment, we saw all violations triggered more incorrect judgments than in the semantic congruous sentences. The behavioral results for each condition were illustrated in Table 2.

Table2: Behavioral results of four conditions (The trial number was presented in upper row in every cell, and the percentage the lower row.)

Conditions	SC: Semantic Congruous (No Violation)	TV: Tonal Violation	VV: Vowel Violation	CV: Consonant Violation
target-hit	1149 88.6%	797 61.5%	1011 78%	493 38%
wrong judgment	95 7%	417 32%	181 14%	742 57%
failure	52 4%	82 6%	104 8%	61 5%

There were several possibilities to explain the consequence of accuracy. The first possibility is that the condition with lower accuracy (TV and CV) revealed that the fact they were not the determining cues in word recognition process. Participants did not rely on these cues very much in speech recognition process and these cues were easily detected. The second possibility was opposite to the first assumption. Participants may rely on these cues very much during the speech recognition process but the cues were not easily detected. We hardly figured out which hypothesis was more possible according to the results from the behavior experiment; therefore, we also conducted the ERP experiment to see brain wave reaction among the violation conditions trying to figure out the specific differences and the determining elements the participants relied on during speech recognition process.

4.4 ERP results

The behavioral results expressed the phenomenon that subjects misjudged more trials in mismatch conditions. In order to figure out the role of different acoustic cues, we further conducted the ERP experiment to see whether the brain waves showed the distinctions when different mismatch conditions occurred. There were total 30 electrodes placed on three parts of the brain: (a) left hemisphere, which are end with odd numbers, such as T3; (b) the central part, ended with the latter z, such as CZ; (c) right hemisphere, ended with even number, such as T4. Or, we can divide these electrodes according to their cortical location: frontal, central, temporal, parietal, and occipital. The followings are the 5 sets of 30 electrodes we recorded the brain waves: F3, F4, F7, F8, FC3, FC4, FCZ, FP1, FP2, FT7, FT8, and FZ; C3, C4, CP3, CP4, CPZ, and CZ; T3, T4, T5, T6, TP7, and TP8; P3, P4 and Pz; O1, O2, and OZ. The distributions of all electrodes were shown in Figure 3.

According to previous literatures, the N400 brain wave occurs when semantic-violated sentences were presented. We extracted the epochs in the duration of 100 ms before the target onset to 1000 ms since the target onset.

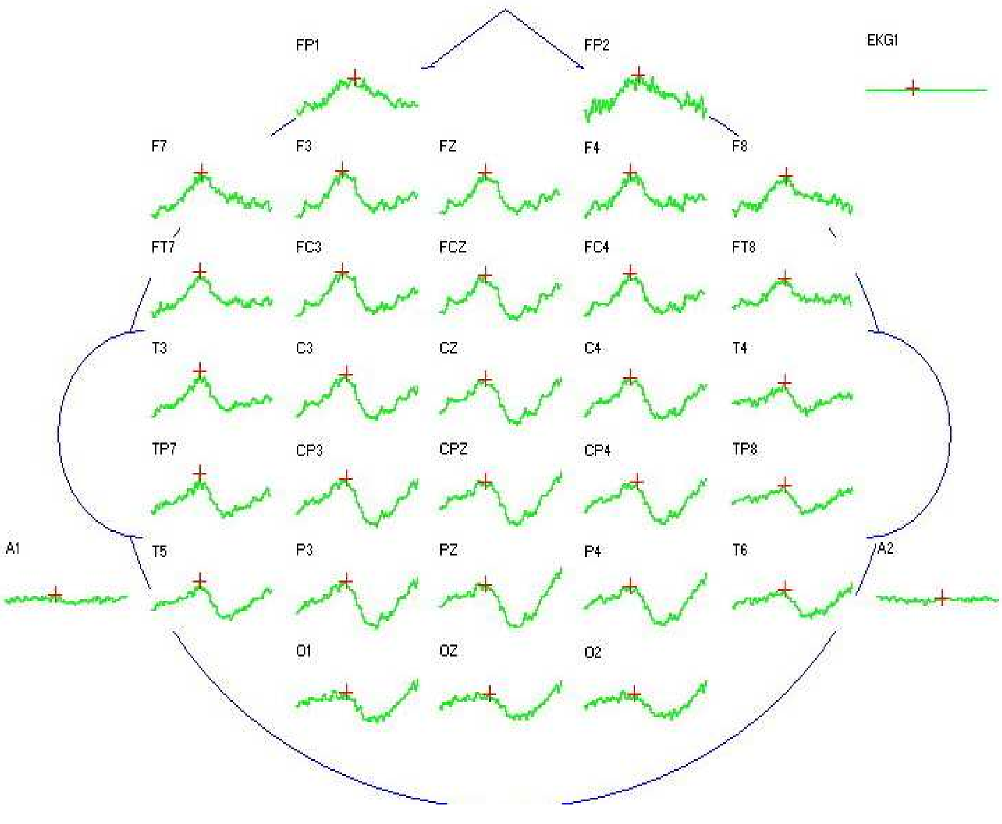


Figure 3: The distributions of 30 electrodes

4.4.1 N400 effect among four conditions

Conventionally, the negative values were represented at the upside and positive at the opposite side; therefore, the N400 wave occurred at the upside from the horizon. We average the minimal value at the time window 350 ms~500 ms to extract the N400 wave form. The values of largest N400 wave at the time 350 ms~500 ms window were extract and marked as in Figure 4. Figure 4 shows the result of four conditions in this experiment.

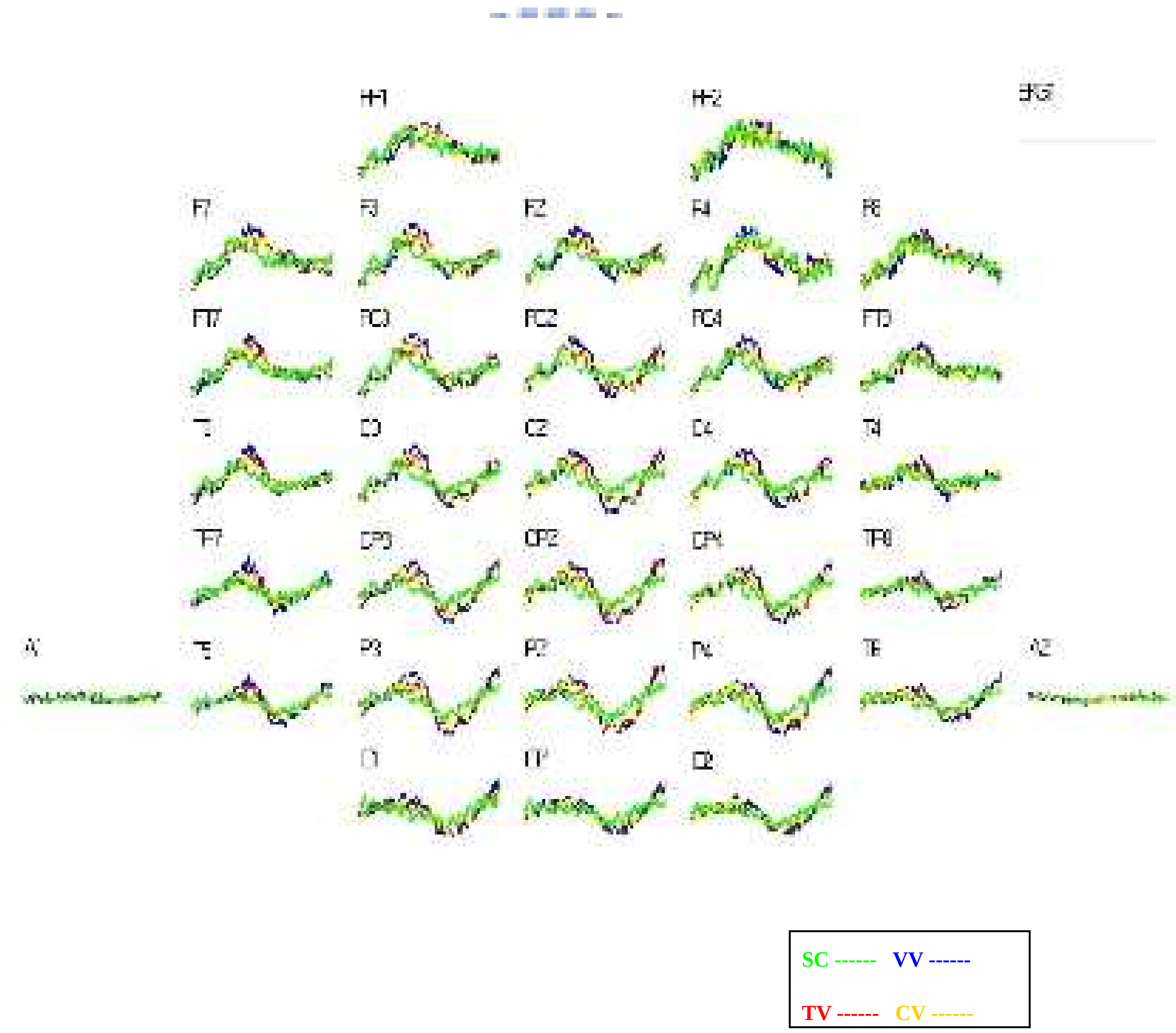


Figure 4: Group Averages of Four Conditions.

After examining the effects of brain distributions on N400 peak latency and peak amplitude through the one-way ANOVA analysis, the N400 peak latency occurred earliest on central and parietal cortex and the N400 peak amplitude were strongest on central and frontal areas. Therefore, we extracted the electrodes from the central and parietal areas to examine the effect on N400 peak latency and extracted the electrodes on the central and frontal areas. Comparing with the semantically congruous condition, all violation conditions elicited the N400 negativity in the time window of 350 and 500 ms after the word onset. Examining the N400 negativity of all violation conditions, we found N400 reaction occurred obviously especially in condition 3, the vowel violation (see Figure 3). The statistic results of ANOVA revealed the main effect of the factor Word Condition was significant on variances of latency and amplitude. The results were shown in the Table 3 and Table 4 below.

Table 3: The Significance of Word Condition on Amplitude

變異數分析

amplitude

			平方和	自由度	平均平方和	F 檢定	顯著性
組間	(組合)		221.060	3	73.687	51.077	.000
	一次項	對比	92.342	1	92.342	64.008	.000
		離差	128.718	2	64.359	44.611	.000
組內			929.078	644	1.443		
總和			1150.138	647			

The main effect of Word Condition on amplitude was significant (F = (3, 644), $p < .005$). The average amplitude is -1.701µV in SC condition, -2.157µV in TV condition, -3.303µV in VV condition, and -2.445µV in CV condition. We further examined the post hoc report, and the results expressed that all violated conditions were elicited larger waveform during 300 ms and 500 ms from the onset of the word compared to the semantic congruous condition, and the distinctions between violated conditions also reached significance (SC vs. TV: $p < .005$; SC vs. VV: $p < .005$; SC vs. CC: $p < .005$). The difference among the three violated conditions was obvious. The vowel violation condition showed the most significant difference among all violated conditions. The N400 negativity was stronger in vowel-violated words and then followed by the consonantal violated (CV) and the tonal violated (TV) conditions.

The main effect of Word Condition on latency was also significant and results were in Table 4 ($F = (3, 320)$, $p < .005$). The highest peak during 350 ms to 500 ms time window occurred during 396 ms from the word onset in SV, 359 ms in TV, 350 ms in VV, and 350 in CV, respectively. Post hoc comparison revealed the distinction between the semantic congruous condition and violated conditions were significant (SC vs. TV: $p < .005$; SC vs. VV: $p < .005$; SC vs. CC: $p < .005$); however, there were no significant difference between all violated conditions (TV vs. VV: $p > .05$; TV vs. VV: $p > .05$; TV vs. CV: $p > .05$; VV vs. CV: $p > .05$).

Table 4: The Significance of Word Condition on Latency

變異數分析

latency

			平方和	自由度	平均平方和	F 檢定	顯著性
組間	(組合)		118066.864	3	39355.621	15.017	.000
	一次項	對比	89363.595	1	89363.595	34.099	.000
		離差	28703.269	2	14351.635	5.476	.005
組內			838638.123	320	2620.744		
總和			956704.988	323			

To summarize our findings of N400 waveform between the semantically congruous condition and the mismatched conditions, we found that N400 negativity showed distinctions in the violated conditions both on latency and amplitude domain. Moreover, vowel violation showed the most obvious differences compared with semantic congruous (condition 1) and semantic incongruous word conditions (condition 2 and condition 4).

4.4.2 Interaction of Word Condition and Hemisphere

After reporting the overall result of different word conditions, we presented the interaction between the factors of Word Condition and Hemispheres as below. We divided Hemisphere into three parts: they were left hemisphere, right hemisphere, and the central area. Electrodes disturbed at the left hemisphere were F7, FT7, T3, TP7, T5, FP1, F3, FC3, C3, CP3, P3, and O1. F8, FT8, T4, TP8, T6, FP2, F4, FC4, C4, CP4, P4 and O2 were distributed at the right hemisphere; the central part was FZ, FCZ, CZ, CPZ, PZ, and OZ. The specific distributions of all the electrodes were shown in Figure 3 above. The Word × Hemisphere interaction failed to reach the significance on latency [$F(6, 1080)=0.921$, $p > .05$], but the Word×Hemisphere interaction was not significant on amplitude [$F(6, 1080)=1.387$, $p < .05$]. The results were presented in Table 5 and Table 6.

Table 5: Word Condition×Hemisphere on latency

受試者間效應項的檢定

依變數: latency

來源	型 III 平方和	自由度	平均平方和	F 檢定	顯著性	淨相關 Eta 平方	Noncent. 參數	觀察的檢定能力[a]
校正後的模式	109104.867[b]	11	9918.624	3.121	.000	.031	34.334	.991
Intercept	2176962.613	1	32176962.6	41594.096	.000	.975	41594.096	1.000
COND	88748.192	3	29582.731	9.309	.000	.025	27.928	.997
HEMI	3332.561	2	1666.281	.524	.592	.001	1.049	.137
COND * HEMI	17560.498	6	2926.750	.921	.479	.005	5.526	.369
誤差	3393871.000	1068	3177.782					
總和	0565712.000	1080						
校正後的總數	3502975.867	1079						

a. 使用 alpha = .05 計算

Table 6: Word Condition×Hemisphere on amplitude

受試者間效應項的檢定

依變數: amplitude

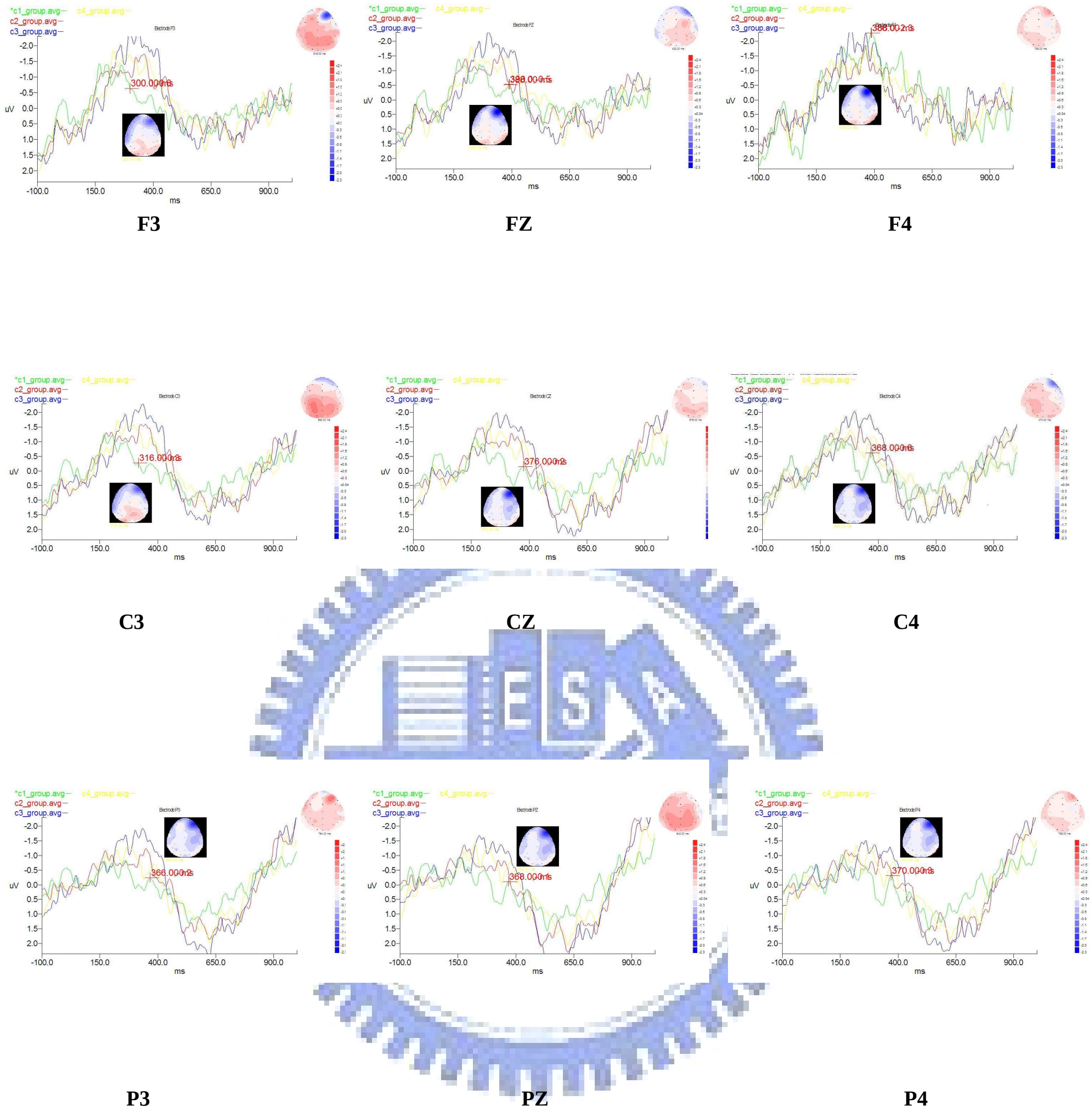
來源	型 III 平方和	自由度	平均平方和	F 檢定	顯著性	淨相關 Eta 平方	Noncent. 參數	觀察的檢定能力[a]
校正後的模式	315.762[b]	11	28.706	21.391	.000	.181	235.300	1.000
Intercept	4147.546	1	4147.546	3090.666	.000	.743	3090.666	1.000
COND	263.751	3	87.917	65.514	.000	.155	196.542	1.000
HEMI	14.226	2	7.113	5.301	.005	.010	10.601	.838
COND * HEMI	11.166	6	1.861	1.387	.217	.008	8.321	.547
誤差	1433.212	1068	1.342					
總和	6493.150	1080						
校正後的總數	1748.974	1079						

a. 使用 alpha = .05 計算

Post hoc analysis revealed that main effect of Word Condition reached significance on amplitude [F = (3, 1080), p < .001] and on latency [F = (3, 1080), p < .001]. The post hoc comparison reported that relative to the semantic congruous condition, the N400 effect occurred in all violation situations (SC vs. TV, p < .001; SC vs.

VV, $p < .001$; SC vs. CV, $p < .001$), and vowel violation was also different from other violations (VV vs. TV, $p < .001$; VV vs. CV, $p < .001$). Tonal violation and consonantal violations are not different in the N400 amplitude (TV vs. CV, $p > .05$). The occurring time of SC was also different from VV and CV (SC vs. VV, $p < .001$; SC vs. CV, $p < .05$), but not different from TV ($p > .05$). N400 effect occurred earlier in vowel violations and consonantal violations but not in the tonal violated word condition.

Hemisphere differences were relevant to the N400 negativity on amplitude [F = (3, 1068), $p < .001$], but not on latency ([F = (3, 1068), $p > .05$]). The N400 negativity detected from the central part was different from left hemisphere and right hemisphere (both p values < .05), and the amplitude was weaker in the central part than the left and right hemispheres. However, there was no significant difference between left and right hemispheres. We compared N400 waveforms from different hemispherical areas in Figure 5, and the electrodes F3, C3, P3, FZ, CZ, PZ, F4, C4, and P4 were chosen. The left column referenced to electrodes at left hemisphere, central column to central part, and right column to right hemisphere. The X-axis represented the occurring time from the onset of the word, measured with millisecond, and the Y-axis represented the elicited strength of brain wave signals.



------ Congruous Words

------ Tone Violation

------ Vowel Violation

------ Consonant Violation

Figure 5: N400 effect of four conditions among left, central, and right hemispheres

4.2.3 Interaction of Word Condition and Scalp distribution on N400 effect

So far we have discussed the interaction between hemisphere and acoustical violation types, and then we will discuss the N400 effect among different violated conditions and among different functions of cortex. We can briefly divide the brain into five parts based on their function to work and we have already introduced their functions in Chapter 2. The factor Site included five levels: frontal, temporal, central, parietal and occipital. We examined the interaction between the factor Word Condition and Site by ANOVA test. Table 7 and Table 8 indicated the result that Word Condition×Site interaction failed to reach significant difference on amplitude ([F(12, 1060)= 1.238, $p > .05$], but reached significance on latency ([F(12, 1060)= 2.36, $p < .05$].

Table 7: Word Condition×Site on amplitude

受試者間效應項的檢定

依變數: amplitude

來源	型 III 平方和	自由度	平均平方和	F 檢定	顯著性	淨相關 Eta 平方	Noncent. 參數	觀察的檢定能力[a]
校正後的模式	522.555[b]	19	27.503	23.771	.000	.299	451.646	1.000
Intercept	2945.764	1	2945.764	2546.037	.000	.706	2546.037	1.000
COND	191.629	3	63.876	55.209	.000	.135	165.626	1.000
SITE	215.000	4	53.750	46.456	.000	.149	185.825	1.000
COND * SITE	17.185	12	1.432	1.238	.251	.014	14.853	.713
誤差	1226.420	1060	1.157					
總和	6493.150	1080						
校正後的總數	1748.974	1079						

a. 使用 alpha = .050 計算

b. R 平方 = .299 (調過後的 R 平方 = .286)

Table 8: Word Condition×Site on latency

受試者間效應項的檢定

依變數: latency

來源	型 III 平方和	自由度	平均平方和	F 檢定	顯著性	淨相關 Eta 平方	Noncent. 參數	觀察的檢定能力[a]
校正後的模式	202390.126[b]	19	10652.112	3.421	.000	.058	64.999	1.000
Intercept	11618422.479	1	111618422.5	35846.828	.000	.971	35846.828	1.000
COND	96134.190	3	32044.730	10.291	.000	.028	30.874	.999
SITE	25985.524	4	6496.381	2.086	.081	.008	8.345	.623
COND * SITE	88192.794	12	7349.400	2.360	.005	.026	28.324	.967
誤差	3300585.741	1060	3113.760					
總和	50565712.000	1080						
校正後的總數	3502975.867	1079						

a. 使用 alpha = .05 計算

b. R 平方 = .058 (調過後的 R 平方 = .041)

Post hoc results revealed that the strongest N400 was elicited in the frontal area (Front vs. Temp, $p < .005$; Front vs. Cent, $p < .005$; Front vs. Parietal, $p < .005$; Front vs. Occipital, $p < .005$) and the weakest N400 was elicited in the occipital area (Occipital vs. Fontal, $p < .005$; Occipital vs. Temp, $p = .005$; Occipital vs. Cent, $p < .005$; Occipital vs. Parietal, $p < .05$). However, the multiple comparisons did not reveal the N400 strength distinction between the three cortical locations: temporal, central, and parietal ($p > .05$).

The significant interaction of Word Condition×Site on latency can be seen from the areas of temporal, central, and parietal (Frontal: SC vs. CV, $p < .05$; Central: SC vs. TV, $p < .05$; SC vs. VV, $p < .001$; SC vs. CV, $p < .05$; Parietal: SC vs. VV, $p < .05$; SC vs. CV, $p < .05$). The result expressed that CV condition elicited earlier N400 peak than the SC condition in temporal lobe. Moreover, all violation situations elicited earlier N400 effect than the semantic congruous in central lobe, and both vowel and consonantal violations occurred earlier N400 effect in parietal.

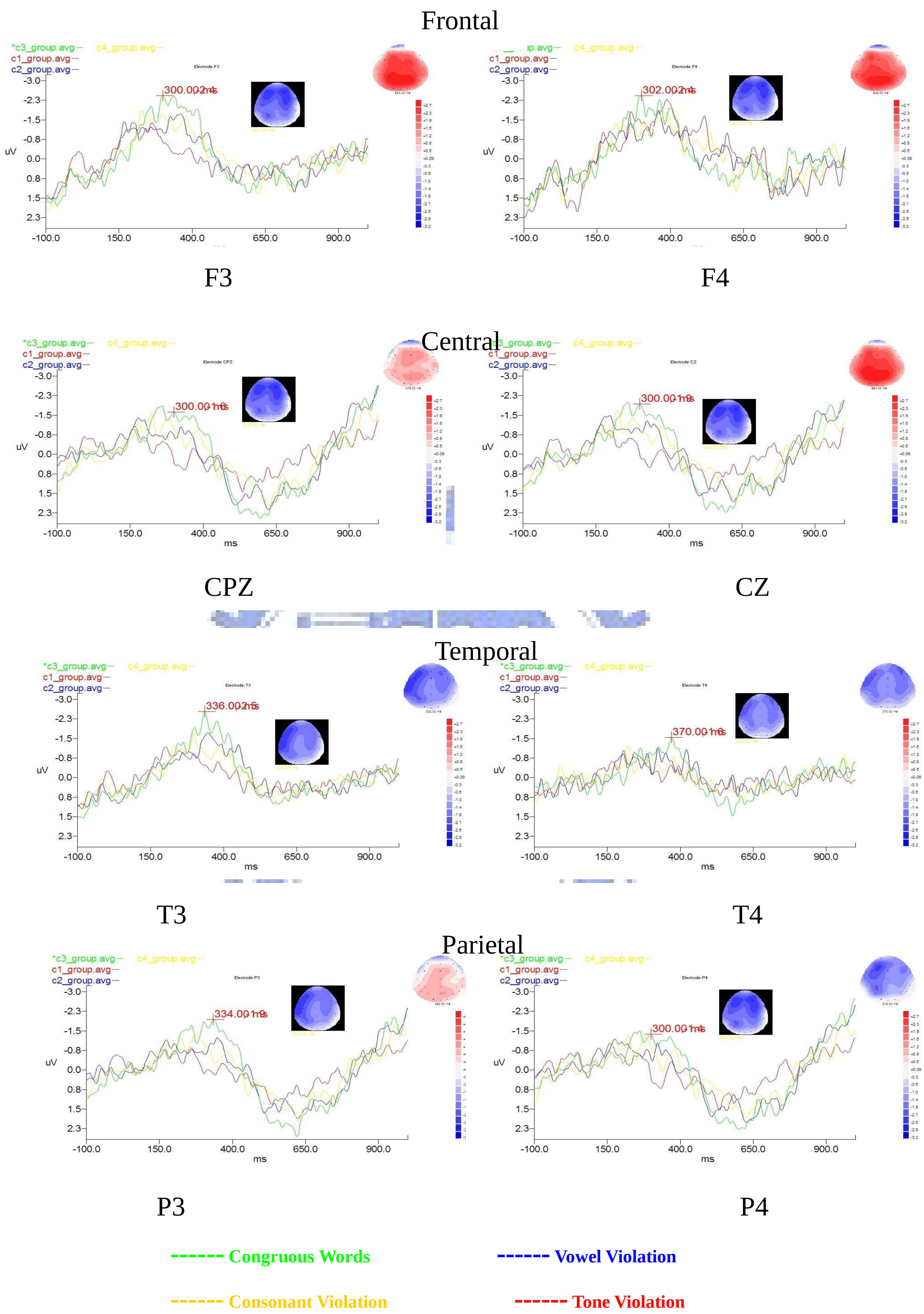


Figure 6: N400 effect of four conditions among frontal, central, temporal, and parietal cortex

4.2.4 Summary of Experiment 1

The findings of the Experiment 1 can be summarized as follow. First, the behavioral experiment indicated that participants used both tonal and segmental cues during the word recognition process. Moreover, consonantal mismatched sentences triggered the most wrong judgments, and tonal violations also caused the high rate of wrong judgment. Finally, participants made the least wrong judgments in the semantically congruous condition.

The ERP results further showed the different reactions between the semantic congruous condition and mismatched conditions. All the mismatched conditions were elicited the N400 negativity and the N400 difference reached significance among all semantic violated conditions. The N400 peak amplitude were strongest when vowel violations occurred and smallest in tonal mismatched conditions. However, we cannot find significantly difference of the N400 negativity reactions between all violations from the performance of peak latency or of peak amplitude. It was possible that the N400 peak amplitude was the more reliable index to the auditorily induced semantic violations. Otherwise, it was also possible that the consonant violations consisted of many heterogeneous combinations and the effects were canceled each other. In order to figure out the effects of tonally mismatched conditions to reactions of N400, we further conducted the Experiment 2. In Experiment 2, the violations resulted from consonantal elements were recorded and the reactions of the N400 waveform were also recorded and labeled according to the classes of the consonantal elements. There were 12 kinds of consonants included in Exp 2.

Chapter 5 Experiment 2

In order to figure out the role all acoustically violated cues played during the semantic recognition process, we designed Experiment 2. As in experiment 1, we examined the response between different semantic violation sentences caused by auditorily mismatched sentences. Because of the more sufficient explanations from EFP experiments, only ERP experiment was conducted in Exp 2 to examine the study.

5.1 Subjects

There were 8 (6 males, 2 females) undergraduates of NCTU participating the experiment voluntarily. They were paid after the experiment finished. All participants are right-handed and speak fluent Mandarin. None of them had lived or studied abroad before 18 years old. Mandarin is their daily language for all participants; moreover, they are skilled in Mandarin reading, writing, listening, and writing. They do not have self-reported hearing and speaking problems.

5.2 Materials

Four types of sentence conditions were included, as in Exp 1, including one semantic congruous condition and three semantic violation conditions. The four conditions were semantic congruous (SC), tonally violated (TV), vowel violated (VV) and consonantally violated (CV) conditions. Each condition contained 53 sentences. The combination of target syllables included four tones of Mandarin, which are high-level, rising, falling-rising, and high-falling tones. The vowels in the target syllables were /i/, /u/, /a/, /ɔ/ and /æ/.

According to the manner of articulation, the consonants of target syllables were divided into 4 groups. Each group consisted of three different consonants and the three consonants of each group were only different in their place of articulation. The four groups are stops /p/, /t/, /k/, aspirated stops /p^h/, /t^h/, /k^h/, fricatives /f/, /x/, /s/, and sonorants /l/, /m/, /n/. The deviant syllables violated one cue of syllable structure in Mandarin on tone, vowel, or onset consonant. That is, there were no syllables with more than one violated element in the experiment, such as the syllable with the violations of both tone and vowel or with the violations of both vowel and onset. Moreover, compared to the study with first-syllable mismatched in disyllabic phrases, all trials in our study belonged to the highly constrained conditions which were final- syllabic mismatched, and this kind of highly constrained conditions were less difficult for the participants during the process of sentence-meaning integration. There were 212 trials included in the Exp 2 in total.

5.3 Procedures

All participants were seated in a comfortable chair in a sound-proof room. The procedure was the same as Exp1, and the only difference is that we exclude the behavior experiment from Exp2. Subjects were asked to stare at the fixation icon on the screen while the auditory sentences were presented through headphones, and they had to press the bottom to decide whether the auditory sentence is semantically correct or not. Because the request was to encourage participants to concentrate on the experiment, the results of subjects' judgments were not recorded.

5.4 Results

We extracted the N400 negativity in time window of 350ms~500ms from the onset of the last syllable in every sentence. The peak amplitude and peak latency of N400 waveform were extracted in our study.

5.4.1 N400 effect among 4 conditions

The statistic results showing the effect of Word Condition on latency and amplitude were shown in Table 8 and Table 9. We also choose the electrodes on the central and parietal areas to examine the effects on the N400 peak latency and choose the electrodes on central and parietal areas to examine the effects on the N400 peak amplitude. The main effect of Word Condition on latency reached significance ([F = (3, 1148)], $p < .05$). The peak of N400 waveform occurred earliest in tonal violation condition, and it occurred latest in consonant violation condition. The means of the N400 peak latency between the time window of 350~500ms from the onset of last syllable in sentences were as follow: the semantic congruous condition (SC) 385.19 ms, the tonally violated condition (TV) 365.98 ms, the vowel violated condition (VV) 378.98 ms, and consonantally violated condition (CV) 400.17 ms. According to the post hoc report, N400 peaks occurred earlier in tonal violation and vowel violation condition comparing to the SC but the CV did not show the significantly statistic difference from the SC group. The distinction between TV and VV did not reach significance. In sum, the earlier N400 negativity was extracted when subjects listened to the tonally and vowel violated sentences while in the consonant violated sentences, the earlier N400 peak did not occur.

The main effect of Word Condition on amplitude also reached significance ([F = (3, 2268)], $p < .05$). The means of peak amplitude between the time window of 300~500ms from the sentence-final syllables were –4.175μ in SC group, -4.681μ in TV group, -4.595μ in VV group, and -4.458μ in CV group (the details were shown in Table 9 below). Post hoc analysis reported that there was neither significant difference between SC and CV group nor significant difference between TV and VV group.

In brief, the one-way ANOVA analysis indicated the effect of the factor Word Condition on the latency and amplitude of N400 peak, and the results showed the consistence of latency and amplitude. Based on the multiple comparisons between the four kinds of sentence conditions, they revealed that the N400 negativity occurred earlier when the subjects listened to the sentences with last tonal or vowel deviant. Also, the N400 waveform reacted stronger while the tonal and vowel violated sentences were presented through auditory way to the participants. On the other hand, the violation of consonant, the CV group, did not show significant difference compared to the controled sentences, the SC group.

Table 9: The main effect of Word Condition on N400 peak latency

變異數分析

latency

			平方和	自由度	平均平方和	F 檢定	顯著性
組間	(組合)		174193.316	3	58064.439	11.857	.000
	一次項	對比	48383.617	1	48383.617	9.880	.002
		離差	125809.699	2	62904.849	12.845	.000
組內			5622029.0	1148	4897.238		
總和			5796222.3	1151			

Table 10: The main effect of Word Condition on N400 peak amplitude

變異數分析

amplitude

			平方和	自由度	平均平方和	F 檢定	顯著性
組間	(組合)		83.558	3	27.853	4.472	.004
	一次項	對比	16.520	1	16.520	2.652	.104
		離差	67.038	2	33.519	5.382	.005
組內			14126.011	2268	6.228		
總和			14209.569	2271			

5.4.2 The interaction of Word Condition and Hemisphere

Beside Word Condition, there were other possibilities which may correlate with the differences of the brain waves in different word condition, such as the brain lateralization or cortical functional difference. Therefore, we further examined the interaction between the two factors of Word Condition and Hemisphere. Like in Exp 1, we divided the brain into three parts—the left part, the right part, and the central part. The interaction between Hemisphere and Word Condition on latency reached significance ($p < .05$), and details were shown in Table 10. Post hoc report revealed that in TV and SC conditions, the sequence of the N400 peak latency occurred earlier in the left side than in the right and it occurred latest in the central part. In VV and CV conditions the N400 waveform occurred earliest in left side, and then in central side and in right side. However, the N400 peak latency extracted from the left, right and central part did not reach significance in SC, TV, and CV condition. The hemispheric difference reached to statistic significance only in the vowel violated (VV) condition. N400 negativities occurred earlier in the left hemisphere than in the right hemisphere when the subjects listened to the sentence with last vowel deviant.

Table 11: Two-way ANOVA of Hemisphere * Word Condition on latency

Tests of Between-Subjects Effects

Dependent Variable: latency

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	505113.395[a]	11	45919.400	8.762	.000
Intercept	528378897.9	1	528378897.9	100820.796	.000
COND	371104.617	3	123701.539	23.604	.000
HEMI	27661.513	2	13830.756	2.639	.072
COND * HEMI	94297.975	6	15716.329	2.999	.006
Error	19726269.232	3764	5240.773		
Total	605602184.0	3776			
Corrected Total	20231382.627	3775			

a. R Squared = .025 (Adjusted R Squared = .022)

Different from the result of latency, the N400 peak amplitude did not show differences among left, right and central areas. The interaction between Word Condition and Hemisphere on amplitude was not significant ([F = (2, 3773)], p > .05). We cannot find the correlation of N400 peak amplitude between the three brain areas and word conditions. Since we already found that there was no correlation between Hemisphere and Condition on amplitude, then we only focused on the effect of Hemisphere on amplitude. One-way ANOVA analysis showed that the effect of Hemisphere on amplitude was significant (F= (2, 3773), p < .05). The mean of the left, right and central hemispheres were –4.15μ, -3.86μ, and –3.93μ, respectively. The post hoc analysis reported that the difference between left and right hemispheres reached significance. The report expressed that stronger N400 negativity was detected in the left hemisphere than in the right hemisphere.

Table 12: The average of peak latency and peak amplitude of N400 among three parts of brains in Exp 2

描述性統計量

		個數	平均數	標準差	標準誤	平均數的 95% 信賴區間		最小值	最大值
						下界	上界		
latency	left	1488	396.7930	74.6552	1.9353	392.9967	400.5893	300.00	548.00
	right	1520	392.7513	70.5746	1.8102	389.2006	396.3021	300.00	548.00
	central	768	389.7370	75.3114	2.7176	384.4022	395.0717	300.00	548.00
	總和	3776	393.7309	73.2073	1.1913	391.3952	396.0667	300.00	548.00
amplitude	left	1488	-4.1492	2.3741	6.155E-02	-4.2699	-4.0284	-26.48	3.92
	right	1520	-3.8610	2.4088	6.178E-02	-3.9822	-3.7398	-18.43	6.57
	central	768	-3.9284	2.3714	8.557E-02	-4.0963	-3.7604	-14.34	2.40
	總和	3776	-3.9882	2.3906	3.890E-02	-4.0645	-3.9120	-26.48	6.57

Table 13: One-way ANOVA: the main effect of Hemisphere on latency and amplitude of N400 negativity in Exp2

變異數分析

		平方和	自由度	平均平方和	F 檢定	顯著性
latency	組間	27661.513	2	13830.756	2.583	.076
	組內	20203721	3773	5354.816		
	總和	20231383	3775			
amplitude	組間	65.908	2	32.954	5.781	.003
	組內	21508.401	3773	5.701		
	總和	21574.308	3775			

*5.4.3 Two-way ANOVA-- Site*Condition*

After discussing the N400 reaction between hemispheres and different types of violated sentences, we further discussed the interaction of cortical areas and sentence conditions on N400 waveform. We use the method of 2-way ANOVA to exam the interaction of the two factors—Site*Condition. The statistic analysis showed that the relation between Site and Condition did not reach significance ($p = .310 > .05$). The results revealed the word condition effect did not show difference in the five cortical areas.

After figuring out the interaction between cortical areas and violated conditions, we further examined the effect of cortical areas on N400 peak latency and amplitude. The one-way ANOVA indicated that the effect of Site on latency was significant ([F=(4, 3771)], $p < .05$). The mean of the N400 peak latency among the frontal, temporal, central, parietal, and occipital areas were 400ms, 392ms, 380ms, 387ms, 402ms, and 393ms, respectively. The earliest N400 negativities were detected from the central area; in contrast, the latest N400 waveforms occurred in the occipital area. The multiple comparisons showed that the N400 significantly occurred earlier in central cortical than in frontal, temporal, and occipital area, but there was no significant difference between the central area and the parietal area. The statistic result also showed that the main effect of Site on amplitude was significant ([F=(4, 3771)], $p < .05$), see also Table 14. Multiple comparisons revealed that N400 component observed in frontal and central areas had larger amplitude than in other areas of the scalp, but statistic results did not show the significant strength between frontal and central areas.

Table 14: The average of peak latency and peak amplitude of N400 negativity among five cortical areas in Exp 2

描述性統計量

		個數	平均數	標準差	標準誤	平均數的 95% 信賴區間		最小值	最大值
						下界	上界		
latency	frontal	1504	400.5638	74.3808	1.9179	396.8017	404.3260	300.00	548.00
	temporal	736	392.8397	71.2910	2.6278	387.6807	397.9986	300.00	548.00
	central	768	380.2083	69.9596	2.5244	375.2527	385.1640	300.00	548.00
	parietal	384	387.3281	72.7901	3.7146	380.0247	394.6316	300.00	548.00
	occipital	384	402.1250	74.8048	3.8174	394.6194	409.6306	300.00	548.00
	總和	3776	393.7309	73.2073	1.1913	391.3952	396.0667	300.00	548.00
amplitude	frontal	1504	-4.7776	2.6526	6.840E-02	-4.9118	-4.6434	-26.48	6.57
	temporal	736	-3.2485	1.8379	6.774E-02	-3.3815	-3.1155	-14.30	1.65
	central	768	-3.8888	2.0529	7.408E-02	-4.0342	-3.7434	-11.81	1.11
	parietal	384	-3.4891	2.1581	.1101	-3.7056	-3.2725	-14.06	.93
	occipital	384	-3.0125	2.1013	.1072	-3.2234	-2.8017	-13.70	.46
	總和	3776	-3.9882	2.3906	3.890E-02	-4.0645	-3.9120	-26.48	6.57

Table 15: One-way ANOVA: the main effect of Site on latency and amplitude of N400 negativity in Exp2

變異數分析

		平方和	自由度	平均平方和	F 檢定	顯著性
latency	組間	254040.350	4	63510.088	11.988	.000
	組內	19977342	3771	5297.625		
	總和	20231383	3775			
amplitude	組間	1808.765	4	452.191	86.272	.000
	組內	19765.544	3771	5.241		
	總和	21574.308	3775			

Chapter 6 General Discussion

Our study aimed at specifying the role of tone and segmental information for word processing in Mandarin. In order to answer the question, we conducted a behavior experiment and ERP experiments. The behavioral results in Exp1 revealed that both tonal elements and segmental elements played important roles on word processing, and tone and consonant mismatched words involved more wrong judgments than the vowel violation words did. The behavior results may also tell us the possibility that the listeners endured the tonal and consonantal information so that they still regarded the mismatched conditions were correct. In sum, we hardly saw the clear and specific effect between different speech information from the behavior result. In order to answer the question, we conducted the ERP experiment and compare the N400 effect between word conditions.

6.1 The comparisons of behavioral experiment and ERP experiments

From results of experiment 1, we discovered that different acoustic cues triggered different N400 reaction on peak latency and amplitude. The semantically violated sentences caused by the mismatching of vowels, tones and consonants involved earlier N400 peak latency and stronger amplitude compared to the semantically congruous sentences. The ERP results of experiment 1 provided explanations of the findings in the behavior experiment. From the behavior experiment, we found that the accuracy was highest when participants listened to the vowel violated sentences than listening to other two violated sentences. Moreover, we found that the violations of vowels triggered earlier N400 waveform and stronger N400 amplitude as well. The findings showed that not only participants perceived vowel cues precisely but also they can detect the mismatch of vowel well during speech recognition process. The participants' sensitivity to vowel elements in behavioral experiment was reconfirmed from the findings of the ERP experiment 1.

Beside vowels, the behavior experiment also showed the different reactions between tonal violation and consonantal violated conditions. The consonantal violations resulted in more incorrect judgments than the tonal violations did. Therefore, we expected the earlier N400 peak latency and stronger amplitude in the tonal violated conditions. However, we did not find the specific difference of N400 peak latency when subjects listened to the tonal and consonantal mismatch sentences in Experiment 1 and the N400 peak amplitudes in consonantal violation were stronger than the tonal violated conditions. The insignificant difference between tonal and consonantal violations might be resulted from the heterogeneous combinations of consonants so that the distinction between the two conditions became unclear. In order to figure out the role the tonal and consonantal cues played during the recognition process, we divided the consonantal elements into several subcategories and compared them with the tonal elements in experiment 2. The results of Exp 2 showed that tonally violated conditions caused earlier and stronger N400 waveform than consonantal violated conditions did. The distinction reached significance in Exp 2. Furthermore, looking into details and comparing different subcategories of consonantal violation with tonal violation, we found that the latency differentiation between CV and TV groups was most salient in the fricative group.

The tonal violations triggered significantly earlier N400 waveform than the fricative violation in CV group. In addition, the N400 of TV group was significantly stronger compared to the category of sonorant violations in CV group.

We further compared the N400 waveform reaction among the consonant violated group, and the within group comparison indicated the difference as well. The means of N400 peak latency occurred in the time window of 413ms in unaspirated stops, 381ms in aspirated stops, 428ms in fricatives and 406ms in sonorants. Besides, the means of the peak amplitude were - 3.57 μ in unaspirated stops, -4.58 μ in aspirated stops, -4.33 μ in fricatives and -2.95 μ in sonorants. The table 16 showed the multiple comparisons between different subcategories of consonantal violated condition. There were two p-values included in each cell with the upper row of latency and the lower row of amplitude.

Table 16: The multiple comparisons of latency and amplitude among the violations of unaspirated stops, aspirated stops, fricatives and sonorants

p-values	Unaspirated stops	Aspirated stops	Fricatives	Sonorants
Unaspirated stops **Lat: 413ms** **Amp: -3.57μ**		p= 0* p= 0*	p= .98 p= .639	p= 1 p= .044*
Aspirated stops **Lat: 381 ms** **Amp: -4.58μ**	p= 0* p= 0*		p= 0* p= 1	p= .006* p=0*
Fricatives **Lat: 428 ms** **Amp: -4.33μ**	p= .98 p= .639	p= 0* p= 1		p= .2 p= 0*
Sonorants **Lat: 406 ms** **Amp: -2.95μ**	p= 1 p= .044*	p=.006* p= 0*	p= .2 p= 0*	

The multiple comparisons indicated that the N400 waveforms showed significant difference in aspirated stops, such as /p^h/, /t^h/ and /k^h/ in our study, compared with other groups. The N400 negativity occurred earlier and waved more strongly in the condition of aspirated stops. In addition, the peak amplitude of N400 waveform was also strong if the expected fricatives were mismatched. The findings indicated the fact that aspirated stops and fricatives triggered more distinctive N400 reactions than other groups of consonants did, and there are two possibilities to explain the phenomenon. First, the similarity of aspirated stops and fricatives is the longer duration when people pronounce them. Second, both aspirated stops and fricatives have the feature of aspiration. The same, there are also two possibilities to explain the distinctive reaction of N400 negativity during the mismatching of tones and vowels. It could be the longer duration of the two elements or the stronger sonority of the tones and vowels compared to the consonantal elements. If the lasting of duration is the crucial element in speech recognition process, it can also explain the result that tonal and vowel violation triggered significant N400 difference compared to the consonant violations, since the vowel and tonal elements involved longer duration, and that explains why aspirated and fricatives also triggered significant N400 distinction than other classes of consonants as well. And the explanation implies that it is hard for people to detect the mismatch or errors during speech if the violation doesn't last long enough. That is to say, people need enough information to find the deviation of speech, and the longer duration provides the more time for people to integrate the speech sounds. Otherwise, people cannot detect the mismatching if the duration of cues is too short.

In order to figure out whether duration is the critical element to help people detects the mismatch well; we further compared the N400 waveforms between vowel violation and tonal violation in the sonorant group. For the reason that the duration of tones is longer in words with sonorant onsets than the words with unvoiced consonant onsets, we expected that tonal violation triggered salient N400 than vowel violation in this sonorant group if assumption 1 was sustained. Therefore, we conducted the comparison between tonal violation and vowel violation whose target words started with the sonorant onsets. If duration were the determining element, then the results would show that the tonal violations triggered earlier N400 peak latency and stronger amplitude compared to the vowel violated sentences. However, the T-test analysis did not support this assumption directly. The means of latency in TV and VV conditions were 368.8ms and 379.68ms, respectively. The effect of duration on N400 peak latency was marginal (p= .59, p >.05). The means of N400 amplitude was –4.05 in TV group and was –4.63 in VV group. The difference was significant (p= .037, p< .05) and it indicated that violation of tones did not trigger more salient N400 negativity than VV group.

Since the duration assumption was not sufficient to explain the robust N400 waveform reaction among tonal and vowel violations nor can it explain the significant N400 waveform among aspirated stops and fricatives of consonant violation, we had to take the assumption 1 into more consideration. It was not sustained that the possibility of duration was a critical element to cause stronger N400 reactions during speech process.

Now we have to consider other possibilities that triggered salient peak latency and amplitude of the N400 waveforms. The statistic results from the comparison of tonal and vowel violated conditions in the sonorant group did not support the assumption that the intrinsic longer duration of vowels and tones than consonants was the determining elements for participants to detect the violation well. However, we don't have to deny the possibility of duration as an important cue in speech process. The important concept we may ignore is that the complexity and correlation of a speech sounds and the clues that people rely on to detect the unexpected or violated speech sounds can be multiple. For example, the sonority of tones and vowels are stronger than consonants, and the stronger sonority of tones and vowels in Mandarin could be also an important clue for people to perceive the deviation. Moreover, people may detect the mismatch in tonal and vowel violated conditions by both the longer duration and stronger sonority. That is, the determining elements may not be only one, and that reflects the complex process of speech integration.

Furthermore, we also found that although consonantal violations did not trigger salient N400 waveform compared with the tonal and vowel violations, there were still some differences of N400 negativity among different classes of consonants. The mismatch of aspirated stops triggered distinctive N400 waveform—the earlier peak latency and stronger peak amplitude, and the mismatch of fricatives also triggered stronger peak amplitude (but not earlier peak latency). We have to notice that the violations of the consonants in our study were to change the place of articulation but not to change the manner of articulation. And, we found that the violations in the consonants with aspiration were elicited salient N400 reaction. This finding revealed that consonantal violations also triggered salient N400 waveforms if the consonant came with aspiration.

6.2 The comparison and analysis among Mandarin and Cantonese

Comparing the studies of Cantonese (Schirmer et al, 2005) and our study, we found a few distinctions between our study and the Cantonese study on the word condition design. First, in our study we compared three acoustic signals: tones, consonants, and vowels while in the Cantonese study the compared information was tones and vowels only. In the study of Cantonese, consonantal element was ignored and the focus was on the core part of a syllable/word— the nucleus which included vowel and tonal elements. Furthermore, in Experiment 1 our ERP results indicated that listeners were most sensitive to the vowel information so that the N400 effect occurred earliest and had most severe waves in the vowel violation condition. The salient N400 reaction of the vowel-violation-only phenomenon did not occur in Experiment 2. The results from Experiment 2 showed that both tonal and vowel effects were comparable on the N400 peak latency and amplitude. The inconsistent results of Experiment 1 and Experiment 2 indicated that the intrinsic characteristic of tones and vowels in Mandarin played the important roles for people to perceive meanings and helped the subjects detect semantic violations in speech.

In the study of Gottfried and Suiter (1997), both native Mandarin speakers and non-native Mandarin speakers (native speakers of English and the mean of Mandarin learning experience was 2.75 years) had good performance and high accuracy in the center-only condition (with the initial and final syllable removed) and made lots of errors in the initial-only condition (with the rhyme part removed) during the word level identification of Mandarin vowels and tones. The non-native speaker group even had higher accuracy in the center-only condition than in the intact condition. Our study reconfirmed the significant role of vowels and tones not only on word identification level (Gottfried and Suiter, 1997) but also in sentence processing.

Finally, our study is different from the Cantonese study in the types of violated conditions. The consonantal violated condition was not included in the Cantonese study and it did not cause salient N400 negativity compared with the vowel and tonal violate conditions in our study. However, we may lose the information from the consonantal element if we ignore the effect of consonantal violations. Previous studies (Kutas &

Hillard, 1980) also indicated that N400 was a reflection of the interruption of sentence processing by semantic inappropriate words, and it also responded the degree of similarity or goodness of consonants in the input to the expected spoken words (Liu et al, 2006). The present study further indicated that although the consonantal cue played a minor role during the speech recognition process compared with the tonal and vowel elements, the aspiration feature may be an important cues in speech. The consonantal violations with the feature of aspiration triggered specifically earlier and stronger N400 waveform than the other types of consonantal violations.

Chapter 7 Conclusion

7.1 Contributions

In our study, we focused on the issue of the effect between acoustical signal and speech comprehension process through the ERP experiment. The result revealed that vowel and tonal violations triggered the most distinctive difference among all violated conditions on the reaction of N400 negativity. And the sensitivity of vowel detection helped listeners make less error during speech comprehension. We also found the convincing language-specified evidence showing that tone-language users make stronger use of supra-segmental information on a lexical level than non-tone language users. And, the nuclear part of a syllable play an important role during the ongoing sentence recognition process in tone language like Mandarin and Cantonese.

Different from the Cantonese study (Schirmer et al, 2005), we not only examined the effect of vowel and tonal violations to the process of speech comprehension but also discussed the effect of consonantal violations. The results revealed that although the consonantal elements played the minor role during the semantically comprehension process the aspirated consonants also caused salient N400 waveforms.

Furthermore, our study tried to use an experimental method to solve the complicated linguistic questions. The studies about language comprehensions have been discussed for many decades and the findings of the previous studies contributed a lot to help the following researches picture the whole concept of languages; however researchers felt difficult in figuring out the relationship between performance of language and brain processing. With the adventure of high technology, we can use the sophisticated equipments to obtain more reliable data, and in our study we discussed our study questions by the help of ERP experiment and had some findings.

7.2 Limitation and Further Study

From the present results, there are still unsolved questions and limitations. For example, we knew that segmental and tonal elements affected the semantic comprehensions when people listened to the sentence with last syllable deviations in Cantonese study (Schirmer et al, 2005) and we further designed the experiments in our study to discuss the issue. But, there was one thing we may ignore in our study and in the previous study that the acoustic elements correlated to the semantic perception may be more complicated. That is, in the Cantonese and our studies the study questions were on the issue of segmental or tonal elements to the effect of speech comprehensions; however, the results from our study showed that the determining elements were probably across the level of segmental or tonal classes. For example, the elements with stronger sonority and aspiration triggered earlier N400 negativity and larger amplitudes when the violations occurred in the sentence final positions.

If we do not restrict the elements on the segmental or tonal elements, it helps us to clear out the relation between auditory information and speech comprehensions. For the future study, the subcategory of vowels is worthy of a discussion and we may find more clues from the follow-up studies. Besides, the neuro-biological signals can be a great help for the language recognition researches. In addition to the peak latency and amplitude of the N400 negativity, the duration, frequency or the contour of the N400 negativity can be used to as a reference as well. Furthermore, we extract N400 negativity as a reference of analyzing the linguistic issue in the present study, but there are still other neuro-signals which reflect the processing of auditory and semantic information during speech, such as N1m and LPC, to name just a few. In sum, the present findings expressed the interaction between acoustic cues and semantic processing and indicated the sharing model and language-specific model in ongoing speech recognition process.

References

1. Chao, Yuen-ren. 1930. A system of tone letters, *Le Maitre phonetique* 30: 24-27.

2. Clark, J., and Yallop, C. (1995). *Phonetics and Phonology*. Massachusetts: Blackwell publishing.

3. Connolly, J.F., & Phillips, N.A. (1994) Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6, 256-266.

4. Cutler, A. (1980). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and speech*, 29, 201-220.

5. Delorem, A., Serby, H., and Makeig, S. (2004). *EEGLAB Tutorial*. Swartz Center for Computational Neuroscience.

6. Friedrich, C.K., Kotz, S.A., Friederici, A. F., & Alter, K. (2004). Pitch modulates lexical identification in spoken word recognition: ERP and behavioral evidence. *Cognitive Brain Research*, 20, 300-308.

7. Gotfried, T. and Suiter, T. (1997). Effect of linguistic experience on the indentification of Mandarin Chinese vowels and tones. *Journal of Phonetics*. 25, 207-231.

8. Gazzaniga, M., Ivry, R., Mangun, G. (1998). The substrates of cognition. *Cognitive Neuroscience*. New York: W.W. Norton & Company.

9. Hgaoort, P., Brown, C., and Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. Special Issue: Event-related

brain potentials in the study of language. *Lang. Cong. Process*. 8:439-483.

10. Hillyard, S.A., Hink, R.F., Schwent, V.L., and Picton, T.W. (1973). Electrical signs of selective attention in the human brain. *Science* 182: 177-180.

11. Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203-205.

12. Roca, I. & Johnson, W. (1999). *A Course in Phonology*. Massachusetts: Blackwell Publishers Inc.

13. Schirmer, A., Tan, S. L., Penney, T. B., Gunter. T. C., & Chen, H.C. (2005). Brain Responses to Segmentaly and Tonally induced semantic violations in Cantonese. Journal of Cognitive Neuroscience, 17:1, 1-12.

14. Soto-Faraco, S., Sebastian-Galles, N., & Cutler, A. (2001). Segmental and supra-segmental mismatch in lexical access. *Journal of Memory and Language*, 45, 412-432.

15. Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R. J. & Näätänen, R. (1999). Functional Specialization of the Human Auditory Cortex in Processing Phonetic and Musical Sounds: A Magnetoencephalographic (MEG) Study. *Neuroimage*, Vol 9, 330-336.

16. Woldorff, M., & Hillyard, S. et al (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proc. Natl. Acad. Sci. U.S.A. 90:8872-8726.

17. Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns.

Neuroimage, 26, 703-720.

18. Liu, Y., Shu, H., Wei, J. (2006). Spoken word recognition in the context: evidence from Chinese ERP analysis. *Brain and Language*, 6, 37-48.

Appendix 1 Materials in Experiment 1

Condition 1: Semantic Congruous Condition
青春痘亂擠會留下傷疤 (ba 1)
信義商圈在跨年夜不關閉 (bi 4)
警方把歹徒逮捕 (bu 3)
上海小籠包最出名的就是餡多皮薄(bou 2)
西方人吃飯習慣用刀叉 (Cha 1)
馬戲團裡有小丑 (Chou 3)
防止禽流感要避免和鳥類接觸 (Chu 4)
在昆蟲館有許多稀奇的昆蟲（Chong 2)
小明喜歡看書和閱讀（ du 2）
三打啤酒喝掉一打剩兩打(da 3)
雲南成爲投資的神奇之地 (di 4)
一年四季是春夏秋冬(dong 1)
新式身分證今天起開始換發 （fa 1）
勸人改過的成語爲：放下屠刀立地成佛（fo2)
總統希望建立一個民主的政府(fu 3)
這面牆壁有道明顯的裂縫(fong 4)
行人要遵守交通規則不能隨便穿越鐵軌(gue3)
說錯話的發言人表情顯的十分尷尬(ga 4)
到了年終許多企業要提出績效評估(gu 1)
逗點後面要空一格(ger 2)
許多詐騙集團犯罪手法十分狡猾 (hua 2)
因爲沒有路燈，這條巷子夜晚十分漆黑（hei 1)
國王的妻子是皇后（hou 4)
戰爭過後許多人的家園被摧毀（hue3)
料理麻油雞的重要的材料是米酒(jiou 3)
最安全的避風港是家（jia 1）
老師要求學生上課時要抄筆記(ji 4）
因爲長期的失眠使得他產生錯覺(jue 2)
弟弟因爲跌倒而嚎啕大哭（ku 1)
年輕人常常不當使用信用卡(ka3)
這座花園種滿了向日葵(kue 2)
他對人生有完美的計畫和掌控(kong 4)
中國人在除夕夜要圍爐(lu 2)
這家的麻辣鍋爽口而不會太辣(la 4)

搭乘飛機前要先到櫃台托運行李(li 3)
東西掉到池塘裡工人忙著打撈（lou 1)
許多父母因爲工作而爲小孩請保姆(mu 3)
每個人多少都有自己的祕密(mi4)
希望頭髮美麗可以多吃芝麻(ma 2)
有位愛貓人士養了二十隻流浪貓(mao 1)
飛行員必須接受長時間的飛行模擬(ni 3)
這個水壺的設計不當因此很難拿(na 2)
示威抗議的人民對市政府表達了他們的憤怒(nu 4)
爲了做出美味的拉麵，麵團要反覆揉捏(nie 1)
姊姊最愛的水果是枇杷(pa 2)
在演奏前歌手要反覆的練習樂譜(pu 3)
在汽車發明前人類的主要交通工具是馬匹(pi 1)
九號開始民眾可以上網訂購火車票(piao 4)
某家商店遭到暴力討債集團在門口潑油漆(qi 1)
人在壓力大的時候會變得容易生氣(qi 4)
家長應該從小就教導小朋友玩具玩完要收整齊(qi 2)
爲了身體的健康，應該養成習慣早睡早起(qi 3)
台北信義商圈在跨年夜約有四十萬人湧入（ru 4)
彰化衛生局整合人力資源照護獨居的老人(ren 2)
能得到全國冠軍，來自南投的小選手感到很光榮（rong 2)
明星常常因爲記者追新聞而受到騷擾(rao 3)
爲了防止蟲害，農藥必須定期噴灑 (sa 3)
這次的國際汽車展中有許多車子造形特殊(su 1)
爲了確保衛生，食物應該煮熟(sou 2)
今年體壇盛事爲六月份的世界盃足球比賽 (sai 4)
因爲地震使得許多房屋倒塌(ta 1)
天氣冷了容易感冒流鼻涕(ti 4)
要種植美味的水果需要肥沃的泥土(tu 3)
公園裡頭有座涼亭(ting 2)
新開幕的海鮮餐廳提供了最新鮮的螃蟹和龍蝦(xia 1)
暴露在潮溼的空氣中容易使鐵生鏽(xiou 4)
專家建議枕頭棉被應該時常清洗(xi 3)
人們假日最常做的休閒是看電視休息(xi 2)
許多家長反應小學生的數學題目太複雜(za 2)
在墾丁的觀光景點有許多民宿提供機車出租(zu 1)
新竹文化中心 18 日晚間將有樂團來演奏(zou4)
寵物貓狗要定時梳洗才能避免長跳蚤(zao 3)

Condition 2: Tonal Violated Condition

開刀過後小白的腳留下傷疤(ba 2)
誠品書局在跨年夜不關閉 (bi 1)
歹徒終於被逮捕 (bu 4)
上海水餃最出名的是餡多皮薄(bou 3)
東方人吃飯不習慣用刀叉 (Cha 2)
弟弟喜歡馬戲團可愛的小丑 (Chou 1)
感染禽流感可能是和鳥類接觸 (Chu 1)
大自然中有許多昆蟲（Chong 3)
哥哥喜歡在圖書館裡閱讀（ du 3）
三打牛奶喝掉一打剩兩打(da 4)
上海成爲投資的神奇之地 (di 1)
今年冬天是個暖冬(dong 2)
里長提醒大家身分證今天起開始換發 （fa 2）
放下屠刀立地成佛（fo 3)
副總統希望建立一個進步的政府(fu 4)
這個花瓶有道明顯的裂縫(fong 1)
老師教小朋友不能隨便穿越鐵軌(gue 4)
說錯話的老師表情顯的十分尷尬(ga 1)
到了年終許多單位要提出績效評估(gu 2)
逗點後面要空一格(ger 3)
許多小偷犯罪手法十分狡猾 (hua 3)
因爲停電，這條巷子夜晚十分漆黑（hei 2)
國王的太太是皇后（hou 1)
戰爭過後許多房屋被摧毀（hue 4)
很多人吃飯時要配啤酒(jiou 1)
最安全的港口是家（ia 3）
老師要求學生上課時要抄筆記(ji 2）
長期的失眠使得老王產生錯覺(jue 4)
妹妹因爲跌倒而嚎啕大哭（ku 3)
民眾要知道如何正確使用信用卡(ka 1)
學校花園種滿了向日葵(kue 4)
他對考試有良好的計畫和掌控(kong 2)
中國人在除夕夜時全家一起吃飯稱作圍爐(lu 1)
這家餐廳的麻辣鍋非常辣(la 2)
旅行的時候要注意手邊的行李(li 1)
戒指掉到池塘裡工人忙著打撈（lou 3)

許多大人因爲工作而爲小孩請保姆(mu 1)
小明對小寶說了自己的祕密(mi 2)
希望身體健康可以多吃芝麻(ma 4)
有位獸醫養了二十隻流浪貓(mao 3)
空軍必須接受長時間的飛行模擬(ni 1)
這個杯子的設計不當因此很難拿(na 4)
示威抗議的人民對警察表達了他們的憤怒(nu 2)
爲了做出美味的包子，麵團要反覆揉捏(nie 3)
爸爸最愛的水果是枇杷(pa 4)
在演奏前歌手要記熟樂譜(pu 1)
在中古世紀人類的主要交通工具是馬匹(pi 3)
民眾常利用網路訂購火車票(piao 2)
某家商店遭到不良少年在門口潑油漆(qi 4)
上班族在壓力大的時候會變得容易生氣(qi 3)
老師應該教導小朋友玩具玩完要收整齊(qi 1)
小朋友應該養成習慣早睡早起(qi 2)
這場演唱會約有四十萬人湧入（ru 3)
台北衛生局整合人力資源照護獨居的老人(ren 1)
能得到全國冠軍，來自台東的小選手感到很光榮（rong 1)
名人常常因爲記者追新聞而受到騷擾(rao 2)
爲了種出美麗的花，農藥必須定期噴灑 (sa 2)
這次的國際玩具展中有許多玩具造形特殊(su 4)
爲了確保衛生，肉類應該煮熟(sou 1)
去年體壇盛事爲全國棒球比賽 (sai 3)
因爲偷工減料使得許多房屋倒塌(ta 2)
寒流來了容易感冒流鼻涕(ti 3)
要種植美麗的花朵需要肥沃的泥土(tu 2)
花園裡頭有座涼亭(ting 1)
哥哥最喜歡的海鮮是螃蟹和龍蝦(xia 4)
暴露在潮溼的空氣中容易使金屬生鏽(xiou 1)
碗盤應該仔細清洗(xi 2)
小朋友假日最常做的休閒是看電視休息(xi 1)
許多家長反應小學生的自然科題目太複雜(za 1)
在墾丁的觀光景點有許多民宿提供房間出租(zu 4)
台北文化中心 18 日晚間將有樂團來演奏(zou 3)
寵物要定時梳洗才能避免長跳蚤(zao 2)

Condition 3: Vowel Violated Condition

小華因爲跌倒頭上留下傷疤 (bu 1)

許多商店在春節不關閉 (ba 4)
搶匪終於被警察逮捕 (bi 3)
這家餛飩的餡多皮薄(bu 2)
美國人吃飯習慣用刀叉 (Chu 1)
曉華喜歡馬戲團可愛的小丑 (Chu 3)
防止禽流感不要和鳥類接觸 (Chi 4)
夏天常常有許多蚊蟲（Chng 2)
圖書館有報紙可以閱讀（ di 2）
兩打雞蛋打破一打剩一打(du 3)
香港成爲國際貿易的神奇之地 (da 4)
春夏秋冬中最冷是的冬(dng 1)
新健保卡從今天起換發 （fu 1）
俗語說：放下屠刀立地成佛（fr2)
大家要團結建立良好的政府(fi 3)
這個杯子有道明顯的裂縫(fng 4)
父母告訴小朋友不能隨便穿越鐵軌(ge3)
說錯話的小明表情顯的十分尷尬(gu 4)
到了期末老師們要做教學評估(gi 1)
逗點後面要空一格(ga 2)
許多學生作弊的手法十分狡猾 (ha 2)
因爲路燈壞了，這條巷子夜晚十分漆黑（hi 1)
毒死白雪公主的是壞皇后（hu 4)
颱風過後許多人的家園被摧毀（he3)
埔里酒廠生產良好米酒(jiu 3)
最安全的地方是自己的家（iu 1)
學生們上課時認真地抄筆記(ja 4）
因爲生病使他產生錯覺(je 2)
小華因爲跌倒而嚎啕大哭（ki 1)
看病要記得帶健保卡(ku3)
他家後院種滿了向日葵(ke 2)
小明對未來有完美的計畫和掌控(kng 4)
中國人在除夕夜會一起吃飯圍爐(li 2)
這家的麻婆豆腐爽口而不會太辣(lu 4)
小明在旅行時遺失了行李(la 3)
戒指掉到池塘裡小明忙著打撈（lu 1)
許多父母因爲忙碌而爲小孩請保姆(ma 3)
小孩也會有自己的祕密(mu 4)
姊姊希望頭髮美麗常吃芝麻(mi 2)

隔壁的老奶奶養了二十隻貓(mo 1)
新的模擬機器讓飛行員練習飛行模擬(nu 3)
這個水壺的設計不當因此很難拿(ni 2)
抗議的學生對學校表達了他們的憤怒(na 4)
爲了做出美味的饅頭，麵團要反覆揉捏(ne 1)
哥哥最不喜歡的水果是枇杷(pi 2)
爲了有好的演出，他反覆的練習樂譜(pa 3)
工業革命前人類的主要交通工具是馬匹(pu 1)
民眾抱怨無法上網訂購火車票(pao 4)
調皮的小孩在門口亂潑油漆(qu 1)
他因爲工作太多而變得容易生氣(qu 4)
小新從小就知道玩具玩完要收整齊(qu 2)
爺爺每天早睡早起(qu 3)
新開幕的水族館約有十萬人湧入（ra 4)
大家要關心獨居的老人(rin 2)
能得到全國冠軍，來自花蓮的小選手感到很光榮（rng 2)
明星不喜歡私生活受到騷擾(ro 3)
爲了蔬菜長得好，農藥必須定期噴灑 (si 3)
這次的服裝展出中有許多服裝設計特殊(sa 1)
爲了確保衛生，開水應該煮熟(su 2)
經過無數練習弟弟贏得這次游泳比賽 (si 4)
因爲設計不良使得這棟房屋倒塌(ti 1)
小明因爲感冒流鼻涕(tu 4)
要種植良好的茶葉需要適當的泥土(ta 3)
隔壁院子裡有頭有座涼亭(tng 2)
小明最討厭的海鮮是螃蟹和龍蝦(xi 1)
避免暴露在潮溼的空氣，鐵才不會生鏽(xiu 4)
衣服應該時常清洗(xu 3)
姊姊假日最常做的休閒是看電視休息(xu2)
許多家長反應小學生的寒假作業太複雜(zi 2)
百事達提供許多錄影帶可以出租(za 1)
台中文化中心 1 日晚間將有樂團來演奏(zu4)
隔壁的小貓因爲不常清洗而長了跳蚤(zo 3)

Condition 4: Consonantal Violated Condition

這隻貓身上有打架留下的傷疤 (da 1)
美術館假日不關閉 (di 4)
長時間調查後歹徒終於被逮捕 (du 3)
這家湯圓的餡多皮薄(dou 2)

韓國人吃飯不習慣用刀叉 (sa 1)
小朋友喜歡馬戲團可愛的小丑 (sou 3)
衛生署呼籲避免和鳥類接觸 (su 4)
山上有稀奇的昆蟲（song 2)
書店有雜誌可以閱讀(gu 2）
三打飲料喝掉一打剩兩打(ga 3)
大陸是台商投資的神奇之地 (gi 4)
政府提供住所幫遊民度過寒冬(gong 2)
駕照從下個月開始換發 （ha 1）
老師今天教我們成語是放下屠刀立地成佛（ho2)
大家不團結就無法建立一個民主的政府(hu 3)
這個盤子有道明顯的裂縫(hong 4)
小明不遵守交通規則隨便穿越鐵軌(bue3)
做錯事的小華表情顯的十分尷尬(ba 4)
針對這個計畫大家做了仔細的評估(bu 1)
逗點後面要空一格(ber 2)
許多詐騙集團犯罪手法十分狡猾 (fua 2)
因為尚未完工，這棟大樓夜晚十分漆黑（fei 1)
拿毒蘋果給白雪公主的是壞皇后（fou 4)
地震過後許多人的家園被摧毀（fue3)
煮麻油雞前媽媽先到商店買米酒(qiou 3)
最舒服的地方是自己的家（qia 1）
學生上課不專心也不抄筆記(qi 4）
因為吃藥使他產生錯覺(que 2)
弟弟因為吃不到糖而嚎啕大哭（pu 1)
哥哥看病忘了帶健保卡(pa3)
這個公園種滿了向日葵(pue 2)
他對進度有很好的計畫和掌控(pong 4)
中國人在除夕夜時重要的習俗是圍爐(bu 2)
這家的宮保雞丁好吃而不會太辣(ra 4)
妹妹在法國旅行時遺失了行李(ri 3)
球掉到池塘裡小明忙著打撈（rau 1)
許多父母不因為工作而為小孩請保姆(nu 3)
他從不透露自己的祕密(ni4)
媽媽希望頭髮美麗常吃芝麻(na 2)
爺爺養了一隻流浪貓(nao 1)
戰鬥機駕駛必須接受長時間的飛行模擬(mi 3)
這個袋子的設計不當因此很難拿(ma 2)

抗議的員工對企業表達了他們的憤怒(mu 4)
爲了做出麵包的拉麵，麵團要反覆揉捏(mie 1)
小明最愛的水果是枇杷(ta 2)
在演奏前小華反覆的練習樂譜(tu 3)
很久以前人類的主要交通工具是馬匹(ti 1)
民眾可以上網訂購電影票(tiao 4)
某家銀行被人在門口亂潑油漆(xi 1)
人在睡眠不足時會變得容易生氣(xi 4)
小朋友玩具玩完要收整齊(xi 2)
奶奶每天養成習慣早睡早起(xi 3)
新開的百貨公司約有五萬人湧入（lu 4)
新竹衛生局整合人力資源照顧獨居的老人(len 2)
能得到全國冠軍，來自屏東的小選手感到很光榮（long 2)
名人討厭私生活受到騷擾(lao 3)
爲了水果長得好，農藥必須定期噴灑 (za 3)
這次的珠寶展覽中有許多珠寶造形特殊(zu 1)
爲了確保衛生，雞蛋應該煮熟(zou 2)
經過無數練習哥哥贏得這次游泳比賽 (zai 4)
水災使得許多房屋毀壞倒塌(ka 1)
小寶因爲感冒流鼻涕(ki 4)
要種植美味的蔬菜需要肥沃的泥土(ku 3)
公園裡有座新蓋好的涼亭(king 2)
妹妹最喜歡的海鮮是螃蟹和龍蝦(jia 1)
避免暴露在潮溼的空氣，金屬才不會生鏽(jiou 4)
蔬菜水果應該仔細清洗(ji 3)
爺爺假日最常做的休閒是看電視休息(ji 2)
許多家長反應小學生的作業題目太複雜(ca 2)
這家書店提供許多書刊雜誌出租(cu 1)
台南文化中心 10 日晚間將有樂團來演奏(cou4)
寵物不定時梳洗容易長跳蚤(cao 3)

Appendix 2 Materials in Experiment 2

Condition 1: Semantic Congruous Condition
青春痘亂擠會留下傷疤 (ba 1)
信義商圈在跨年夜不關閉 (bi 4)
警方把歹徒逮捕 (bu 3)
畢業後大家就要離別（bie2）
張教授是電子學研究領域的巨擘（bo4）
小明喜歡看書和閱讀（ du 2）
三打啤酒喝掉一打剩兩打(da 3)
雲南成爲投資的神奇之地 (di 4)
這次作業的份量很多(duo1)
昆蟲館裡面有各種蝴蝶(die2)
新式身分證今天起開始換發（fa 1）
勸人改過的成語爲：放下屠刀立地成佛（fo2)
總統希望建立一個民主的政府(fu 3)
行人要遵守交通規則不能隨便穿越鐵軌(gue3)
說錯話的發言人表情顯的十分尷尬(ga 4)
到了年終許多企業要提出績效評估(gu 1)
戶外郊遊時大家開心地嘻嘻哈哈(ha1)
傳說中的水怪出現在尼斯湖(hu2)
前人教導我們要付出努力才會有收穫(huo4)
這片好山好水因爲人們的破壞而摧毀(hue3)
弟弟因爲跌倒而嚎啕大哭（ku 1)
年輕人常常不當使用信用卡(ka3)
這座花園種滿了向日葵(kue 2)
爸爸常告訴我做人胸襟要開闊(kuo4)
中國人在除夕夜要圍爐(lu 2)
這家的麻辣鍋爽口而不會太辣(la 4)
搭乘飛機前要先到櫃台托運行李(li 3)
以前的農田裡有許多田螺（luo2）
博物館裡面有許多物品陳列(lie4)
許多父母因爲工作而爲小孩請保姆(mu 3)
每個人多少都有自己的祕密(mi4)
希望頭髮美麗可以多吃芝麻(ma 2)
小偷偷東西時總是偷偷摸摸（mo1）
經過一番努力消防員終於把火撲滅（mie4）

飛行員必須接受長時間的飛行模擬(ni 3)
這個水壺的設計不當因此很難拿(na 2)
示威抗議的人民對市政府表達了他們的憤怒(nu 4)
為了做出美味的拉麵，麵團要反覆揉捏(nie 1)
王先生說話算話十分重承諾（nuo4）
姊姊最愛的水果是枇杷(pa 2)
在演奏前歌手要反覆的練習樂譜(pu 3)
在汽車發明前人類的主要交通工具是馬匹(pi 1)
人海中他們兩個人只是匆匆地一瞥(pie1)
妹妹不小心把玻璃打破(po4)
為了防止蟲害，農藥必須定期噴灑 (sa 3)
這次的國際汽車展中有許多車子造形特殊(su 1)
王老闆受到歹徒的勒索(suo3)
大明星出門總是有許多保鏢跟隨(sue2)
因為地震使得許多房屋倒塌(ta 1)
天氣冷了容易感冒流鼻涕(ti 4)
要種植美味的水果需要肥沃的泥土(tu 3)
沙漠的代步工具是駱駝(tuo2)
工業發展需要大量的鋼鐵(tie3)

Condition 2: Tonal Violated Condition
開刀過後小白的腳留下傷疤(ba 2)
誠品書局在跨年夜不關閉 (bi 1)
歹徒終於被逮捕 (bu 4)
吃完飯後大家就要離別(bie3)
王教授是生物學研究領域的巨擘（bo1）
哥哥喜歡在圖書館裡閱讀（ du 3）
三打牛奶喝掉一打剩兩打(da 4)
上海成為投資的神奇之地 (di 1)
這家餐廳的食物份量很多(duo2)
博物館裡面有各種蝴蝶(die3)
里長提醒大家身分證今天起開始換發 （fa 2）
放下屠刀立地成佛（fo 3)
副總統希望建立一個進步的政府(fu 4)
老師教小朋友不能隨便穿越鐵軌(gue 4)
說錯話的老師表情顯的十分尷尬(ga 1)
到了年終許多單位要提出績效評估(gu 2)
畢業旅行時大家開心地嘻嘻哈哈(ha3)

聽說有水怪出現在尼斯湖(hu4)
老師教導我們要付出努力才會有收穫(huo2)
這座山因為人們的破壞而摧毀(hue1)
妹妹因為跌倒而嚎啕大哭（ku 3)
民眾要知道如何正確使用信用卡(ka 1)
學校花園種滿了向日葵(kue 4)
老師常告訴我做人胸襟要開闊(kuo2)
中國人在除夕夜時全家一起吃飯稱作圍爐(lu 4)
這家餐廳的麻辣鍋非常辣(la 2)
旅行的時候要注意手邊的行李(li 1)
現在的農田裡很少看到田螺（luo4）
美術館裡面有許多畫陳列(lie2)
許多大人因為工作而為小孩請保姆(mu 1)
小明對小寶說了自己的祕密(mi 2)
希望身體健康可以多吃芝麻(ma 4)
學生作弊時總是偷偷摸摸（mo3）
大家終於把火撲滅（mie2)
空軍必須接受長時間的飛行模擬(ni 1)
這個杯子的設計不當因此很難拿(na 4)
示威抗議的人民對警察表達了他們的憤怒(nu 2)
為了做出美味的包子，麵團要反覆揉捏(nie 3)
老師說話算話十分重承諾（nuo2）
爸爸最愛的水果是枇杷(pa 4)
在演奏前歌手要記熟樂譜(pu 1)
在中古世紀人類的主要交通工具是馬匹(pi 3)
在車站裡他們兩人只是匆匆地一瞥(pie3)
弟弟不小心把花瓶打破(po2)
為了種出美麗的花，農藥必須定期噴灑 (sa 2)
這次的國際玩具展中有許多玩具造形特殊(su 4)
小明受到不良少年的勒索(suo2)
官員出門總是有許多保鏢跟隨(sue1)
因為偷工減料使得許多房屋倒塌(ta 2)
寒流來了容易感冒流鼻涕(ti 3)
要種植美麗的花朵需要肥沃的泥土(tu 2)
沙漠常常可以看到駱駝(tuo1)
製造武器需要大量的鋼鐵(tie2)

Condition 3: Tonal Violated Condition

小華因為跌倒頭上留下傷疤 (bu 1)

許多商店在春節不關閉 (ba 4)
搶匪終於被警察逮捕 (bi 3)
六月份以後大家就要離別(bo2)
李教授是生物學研究領域的巨擘（be4）
圖書館有報紙可以閱讀（ di 2）
兩打雞蛋打破一打剩一打(du 3)
香港成爲國際貿易的神奇之地 (da 4)
這次作業的份量很多(de1)
昆蟲館裡面有各種蝴蝶(do2)
新健保卡從今天起換發 （fu 1）
俗語說：放下屠刀立地成佛（fr2)
大家要團結建立良好的政府(fi 3)
父母告訴小朋友不能隨便穿越鐵軌(go3)
說錯話的小明表情顯的十分尷尬(gu 4)
到了期末老師們要做教學評估(gi 1)
戶外郊遊時大家開心地嘻嘻哈哈(hi1)
今天上演的戲劇是天鵝湖(ha2)
父母教導我們沒有努力就沒有收穫(hue4)
這條河川因爲污染而摧毀(huo3)
小華因爲跌倒而嚎啕大哭（ki 1)
看病要記得帶健保卡(ku3)
他家後院種滿了向日葵(kuo 2)
媽媽常告訴我做人胸襟要開闊(kue4)
中國人在除夕夜會一起吃飯圍爐(li 2)
這家的麻婆豆腐爽口而不會太辣(lu 4)
小明在旅行時遺失了行李(la 3)
小明養了多田螺（lue2）
昆蟲館裡面有許多標本陳列(lo4)
許多父母因爲忙碌而爲小孩請保姆(ma 3)
小孩也會有自己的祕密(mu 4)
姊姊希望頭髮美麗常吃芝麻(mi 2)
小偷偷東西時總是偷偷摸摸（me1）
由於火勢太大不容易把火撲滅（mo4）
新的模擬機器讓飛行員練習飛行模擬(nu 3)
這個水壺的設計不當因此很難拿(ni 2)
抗議的學生對學校表達了他們的憤怒(na 4)
爲了做出美味的饅頭，麵團要反覆揉捏(nuo1)
父母教導小孩要重承諾（nue4）

哥哥最不喜歡的水果是枇杷(pi 2)
爲了有好的演出，他反覆的練習樂譜(pa 3)
工業革命前人類的主要交通工具是馬匹(pu 1)
人海中他們兩個人只是匆匆地一瞥(po1)
哥哥不小心把杯子打破(pie4)
爲了蔬菜長得好，農藥必須定期噴灑 (si 3)
這次的服裝展出中有許多服裝設計特殊(sa 1)
張老師受到歹徒的勒索(sue3)
總統出門總是有許多保鏢跟隨(suo2)
因爲設計不良使得這棟房屋倒塌(ti 1)
小明因爲感冒流鼻涕(tu 4)
要種植良好的茶葉需要適當的泥土(ta 3)
動物園裡可以看到駱駝(tue2)
造船業需要用到大量的鋼鐵(tuo4)

Condition 4: Consonantal Violated Condition
這隻貓身上有打架留下的傷疤 (da 1)
美術館假日不關閉 (di 4)
長時間調查後歹徒終於被逮捕 (du 3)
討論完後後大家就要離別(die2)
吳教授是地理學研究領域的巨擘（do4）
書店有雜誌可以閱讀(gu 2)
三打飲料喝掉一打剩兩打(ga 3)
大陸是台商投資的神奇之地 (gi 4)
這次作業的份量很多(guo1)
昆蟲館裡面有各種蝴蝶(gie2)
駕照從下個月開始換發 （ha 1）
老師今天教我們成語是放下屠刀立地成佛（ho2)
大家不團結就無法建立一個民主的政府(hu 3)
小明不遵守交通規則隨便穿越鐵軌(bue3)
做錯事的小華表情顯的十分尷尬(ba 4)
針對這個計畫大家做了仔細的評估(bu 1)
戶外郊遊時大家開心地嘻嘻哈哈(fa1)
傳說中的水怪出現在尼斯湖(fu2)
前人教導我們沒有努力就不會有收穫(fuo4)
大自然因爲人們的破壞而摧毀(fue3)
弟弟因爲吃不到糖而嚎啕大哭（pu 1)
哥哥看病忘了帶健保卡(pa3)

這個公園種滿了向日葵(pue 2)
爸爸常告訴我做人胸襟要開闊(puo4)
中國人在除夕夜時重要的習俗是圍爐(ru 2)
這家的宮保雞丁好吃而不會太辣(ra 4)
妹妹在法國旅行時遺失了行李(ri 3)
以前的農田裡有許多田螺（ruo2）
圖書館裡面有許多書陳列(rie4)
許多父母不因爲工作而爲小孩請保姆(nu 3)
他從不透露自己的祕密(ni4)
媽媽希望頭髮美麗常吃芝麻(na 2)
小偷偷東西時總是偷偷摸摸（no1）
失火時村民一起把火撲滅（nie4）
戰鬥機駕駛必須接受長時間的飛行模擬(mi 3)
這個袋子的設計不當因此很難拿(ma 2)
抗議的員工對企業表達了他們的憤怒(mu 4)
爲了做出美味的拉麵，麵團要反覆揉捏(mie 1)
妹妹說話算話十分重承諾（muo4）
小明最愛的水果是枇杷(ta 2)
在演奏前小華反覆的練習樂譜(tu 3)
很久以前人類的主要交通工具是馬匹(ti 1)
人海中他們兩個人只是匆匆地一瞥(tie1)
妹妹不小心把玻璃打破(to4)
爲了水果長得好，農藥必須定期噴灑 (za 3)
這次的珠寶展覽中有許多珠寶造形特殊(zu 1)
小明受到同學的勒索(zuo3)
明星出門總是有許多記者跟隨(zue2)
水災使得許多房屋毀壞倒塌(ka 1)
小寶因爲感冒流鼻涕(ki 4)
要種植美味的蔬菜需要肥沃的泥土(ku 3)
在沙漠方便的交通工具是駱駝(kuo2)
發展工業需要大量的鋼鐵(kie3)