Unraveling the roles of word prosody and sublexical phonology in spoken English and spoken and written Chinese using event-related brain potentials

by

Jing Tian Wang

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie Univeristy
Halifax, Nova Scotia
March 2003

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

0-612-79398-2

DALHOUSIE UNIVERSITY

FACULTY OF GRADUATE STUDIES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled “Unraveling the roles of word prosody and sublexical phonology in spoken English and spoken and written Chinese using event-related brain potentials” by Jingtian Wang in partial fulfillment for the degree of Doctor of Philosophy.

Dated: March 26, 2003

External Examiner:

Research Supervisor:

Examining Committee:

Departmental Representative:

[redacted]

DALHOUSIE UNIVERSITY

DATE: March, 2003

AUTHOR: JING TIAN WANG

TITLE: UNRAVELING THE ROLES OF WORD PROSODY AND SUBLEXICAL PHONOLOGY IN SPOKEN ENGLISH AND SPOKEN AND WRITTEN CHINESE USING EVENT-RELATED BRAIN POTENTIALS

DEPARTMENT OR SCHOOL: PSYCHOLOGY/ NEUROSCIENCE

DEGREE: Ph.D. CONVOCATION: Spring YEAR: 2003

[redacted]

Signature of Author

DEDICATION

To MY BELOVED PARENTS,
Shi Fang Wang AND
Ya Wen Zhang

Table of Contents

List of Tables

List of Figures

Abstract

The main intentions of the thesis were to determine the roles that word prosody may play in speech and reading and to examine the effects of sub-lexical phonology on lexical activation. The thesis also aimed to elucidate the nature of phonological processing in silent reading. Given that event-related brain potentials (ERPs) can provide reliable information concerning the timing and sequencing of linguistic processes, they were used as the primary research tool. The Phonological Mismatch Negativity (PMN) associated with phonological encoding in speech and the N400 related to semantic analysis were studied in the first two speech experiments. The N270 sensitive to form-based (orthography and phonology) processing in reading and the N400 were further examined in the third reading experiment. In the first experiment, ERPs were recorded as native English speakers listened to a series of English sentences whose terminal words were varied according to lexical stress, syllable-sized sub-lexical phonology, and semantic appropriateness. The results of this experiment showed that both the PMN and the N400 were sensitive to lexical stress and contextually primed syllables. This suggests that lexical processing primacy for non-initial-stress words is given to primarily stressed syllables instead of word onsets. In the second and third experiments, ERPs were recorded as native Chinese speakers listened to (experiment 2) or silently read (experiment 3) a series of Chinese four character proverbs whose ending characters were manipulated by lexical tone, onset, rime, and semantic variables. It was shown that both the PMN and the N400 were modulated by lexical tone and segmental manipulations. Furthermore, the N270 findings suggest the presence of orthographic-to-phonological transformation in reading. It is argued that lexical tone is involved in both Chinese speech and reading comprehension. Segments play an important role in semantic constraint not only in speech but also in reading. In all, it is concluded that word prosody is represented in the mental lexicon, which reflects aspects of a language-universal mechanism. Lexical activation progresses with directionality but not in a strictly sequential fashion. Lexical processing does not proceed in an all-or-none commitment. Instead, sub-lexical phonology exerts crucial facilitating effects on whole word/character processing. The present research provides explicit evidence that phonology plays a primary role in both speech and reading.

Abbreviations and Symbols

ANOVA	analysis of variance
BAEP	brainstem auditory evoked response
EEG	electroencephalogram
EOG	electro-oculogram
ERP	event-related potential
Hz	Hertz
kΩ	kilo-ohm
LPC	late positive component
<u>M</u>	mean
ms	millisecond
MSS	metrical stress segmentation
PMN	phonological mismatch negativity
REA	right ear advantage
RT	reaction time
SD	standard deviation
SE	standard error
SEP	somatosensory evoked potentials
SOA	stimulus onset asynchrony
μV	micro-volt

ACKNOWLEGEMENTS

First and foremost, I would like with my great appreciation and sincere gratitude to thank my supervisor, Dr. John F. Connolly, for his mentorship, continuous support and encouragement during the whole course of my Ph.D. training. He has been always inspiring me with his open-minded thinking and wisdom in science. His word of support and encouragement has always bolstered my confidence and reinforced my courage on the journey to academic success.

I would also like to express my thanks to my external examiner, Dr. Shlomo Bentin, my committee members, Dr Kazue Semba and Dr Sherry Stewart, and my defense Chair, Dr. Raymond Klein, for their invaluable feedback and challenging questions.

I am also grateful for many friends, professors and colleagues throughout my graduate study:

-- to the Connolly lab, thank you all for your support. In particular, Dr. Elisabet Service and Alma Major have been extremely helpful in my research. Many thanks also to Dr. Yannick Marchant, Dr. Ryan D'Arcy, Randy Newman, and Chris Striemer for their kind encouragement and good sense of humor.

-- to my teacher and good friend, Dr. Ian Meinertzhagen, for his great encouragement and helpful advice.

-- to my MRI supervisor, Dr. Matthias Schmidt, for his rich knowledge and generous guidance.

-- to Dr. Shelly Adamo, Dr. Bryan Young, and Dr. Tracy Taylor for their great support in the completion of my comprehensive projects.

-- to my classmates and friends. Jennifer Klages, thank you so much for your precious friendship. I have so many fond memories of the time I have spent with you. Peggy Ruyak, thank you for your great friendship, kind help and professional advice throughout our graduate course. Scott Fairless and Mike Ellery, thank you for your support and encouragement that helped me to keep smiling during my stressful first year adjustment. I will miss you all!

-- to my dear friends in the Chinese community in Halifax for their generous support in my research and continuous encouragement that has made me strong and confident.

-- to my fiance, Zeiad Muntasser, I feel blessed to have you in my life.

-- to my parents and to each of my sisters and their spouses for their constant demonstration of love and integrity.

Unraveling the roles of word prosody and sublexical phonology in spoken English and spoken and written Chinese using event-related brain potentials

Chapter One: Introduction

1.1 Lexical prosody and sub-lexical phonology in spoken English word processing

Speech comprehension can be crudely seen as the process through which meaning is derived from continuous acoustic signal analysis. Meanwhile, given the fact that the brain is confronted with an infinite number of potential utterances, the brain cannot hold a representation of every whole utterance in memory. Rather, what are held in memory must be representations of discrete meaningful units, or words, which make up such utterances (Cutler, 1986; Cutler & Carter, 1987; Cutler & Norris, 1988). In this regard, the recognition of discrete words becomes of crucial relevance in fluent speech comprehension. It is noteworthy, however, that a problem that speech researchers have been facing is that words in fluent speech do not exist as separate acoustic events. There are no reliable equivalents of white spaces that conveniently present the onset and offset of a word in continuous text (Cutler, 1992). It has been argued that "words are the product of speech perception" and "they exist in the mind of a perceiver, and not in the physical stimulus" (Cole & Jakimik, 1980, p.133). On this view, much research effort has thereby been devoted to understanding the means and procedures through which the semantic representations of single spoken words in fluent speech can be accessed in the absence of clear cues to the locations of word boundaries.

Under such circumstances, a tentative solution in an attempt to locate spoken word boundaries on the basis of syllables to facilitate lexical access has been proposed. It is argued that mental representations of a spoken word may depend on syllabic units so that listeners can efficiently divide speech input into syllables in order to facilitate single word processing (Mehler, 1981; Mehler, Dommergues, Frauenfelder, & Segui, 1981). Based on this argument, a continuous speech stream can be constructed as a series of syllables that can be practiced as pre-lexical representations of the potential word. Therefore, lexical access can be attempted beginning on each syllable. The convincing evidence for this argument comes from an influential French study (Mehler, Dommergues, Frauenfelder, & Segui, 1981), the results of which revealed that syllable-sized targets were detected much faster when the target was compatible with the actual syllabification of the speech input. For example, the target "ba" is detected faster when it is in 'ba-llance' than in "bal-con" while the target "bal" is detected faster when it is in "bal-con" than in "ba-llance". Furthermore, if the target corresponds exactly to the first syllable of a word, the speed of its detection is significantly quicker than if it consists of more or less than the syllable (for a review, see Segui, Dupoux, & Mehler, 1990). It is thus reasonable to argue that this syllable-based strategy is efficient in speech segmentation at least in the French language. Unfortunately, this pattern of results did not hold up in English language studies (Cutler, Mehler, Norris, & Segui, 1983, 1986; Cutler & Norris, 1988). It is reasoned that perceptual procedures in speech processing are heavily influenced by the phonological features of the language (Cutler & Norris, 1988). The French language has regular syllable patterns and listeners generally rely on syllabic structure. In contrast, words in English have diverse syllable patterns. They may consist of one or multiple syllables; likewise, monosyllabic words in English can consist of a considerable range of syllable structures (the words "I" and "scrounged" are both monosyllables). This may make the syllable based strategy not appropriate to identify word boundaries for the English language.

Alternatively, several other computational models have been developed that have had a strong impact on most contemporary language literature. These models have been heavily influenced by written language theories that deemphasize the necessity of segmenting word boundary before lexical access and can fall under one or the other of the two major hypotheses in the literature to be introduced below. By this account, the two hypotheses are seen as general class hypotheses in this thesis. These models will be delineated and discussed by virtue of these hypotheses.

1.1.1 Two major hypotheses on the progression of spoken word recognition mechanisms

1.1.1.1 The "sequential" hypothesis

The "sequential" hypothesis aims to accommodate the sequential time-dependent nature of the speech stream. Proponents of this principle surmise that lexical activation of a spoken word proceeds with the deployment of speech in a strictly time-shadowing fashion. That is, speech information occurring first in time is always processed first. To illustrate, with the word *discount,* the initial sound *d* would initiate lexical search and be processed before the later sound *iscount* because *d* appears first in time. Indeed, from the anatomical standpoint, it is evident that acoustic signal is produced in a fairly sequential way with limited overlap between phonemes (Liberman, 1970; Massaro, 1975; Studdert-Kennedy, 1975). In this regard, it seems impossible that all the phonemes of a spoken word can be pronounced simultaneously, which renders the "sequential" hypothesis logical and likely to reflect the nature of all proactive language processing with early information contributing to the analysis of later information (for review see, Mattys, 1997).

The sequential hypothesis has been extensively developed in certain models of the process of spoken word recognition (e.g., Cole & Jakimik, 1978, 1980a; Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978; Marslen-Wilson & Zwitserlood, 1989). Amongst those models, most influential is the Cohort model. According to this model, a word's initial information activates a cohort of word candidates corresponding to the acoustic signal. With the sequential accruement of acoustic information, the cohort candidates, if matching the input, remain activated; otherwise, they immediately start to be deactivated. The consequence of this activation/deactivation process is the isolation of only one candidate that matches the sensory input while the others no longer do. The precise moment at which a word becomes isolated and recognized following this unfolding sequential process is called the uniqueness point (UP). Marslen-Wilson (1987) argues that such a strategy appears to be computationally inexpensive because only the acoustic signal before the UP in theory needs processing to have recognition occur. By this reasoning, words can be recognized with half or less than half of their full acoustic input and even sooner when the words are heard in a context constraining environment (Grosjean, 1996). The economic feature of this model is particularly evident, in theory, for long words with an early UP. For instance, the first four phonemes of the word *rhinoceros* are adequate to achieve recognition because it is the only word in English that starts with the sequence */rɋina/*. Another contribution of this model to speech processing is that early recognition may allow the offset of a word to be anticipated, which thereby solves the complex problem of word segmentation in continuous speech (Taft, 1984).

Until recently, the Cohort model had been dominating spoken word recognition research, which has been fostered by a considerable body of English literature with various tasks, such as shadowing (Marslen-Wilson, 1973), phoneme monitoring (Frauenfelder, Segui, & Dijkstra, 1990; Marslen-Wilson, 1984), lexical decision (Goodman & Huttenlocher, 1988; Taft & Hambly, 1986), mispronunciation detection (Jakimik, 1979; Ottevanger, 1984), gating (Grosjean, 1980; Tyler & Wessels, 1983), and phonemic restoration (Samuel, 1987). Although compelling, the sequential hypothesis and its supporting Cohort model have not gone unchallenged.

First of all, an increasing number of studies question the viability of the processing primacy of word onsets. It has been observed that the acoustic signal in fluent speech is naturally produced in a wide variety of conditions that could be disruptive (e.g., coarticulation, misarticulation, environmental noise) so that it cannot guarantee the processing system reliable information about word onsets (Connine, Blasko, & Titone, 1993; Marslen-Wilson & Zwitserlood, 1989). In this view, the over emphasis on word onset information puts a cohort-based recognition model into difficulty whenever the information at word beginnings is obscure. On the contrary, it appears that human listeners have the alleged ability to recognize such mispronounced items as "shigarette" or "dwibble" without difficulty (Grosjean, 1985; Norris, 1981). For example, in a cross-modal priming task, it was found that disruption of word onset information did not restrain recognition much more than violations of non-initial parts (Connine, Blasko, and Titone, 1993).

In response to such criticisms, Marslen-Wilson and Warren (1994) have argued that given the flexibility of the cognitive system, "shigarette" is processed as the word "cigarette" in fact by a correction strategy which may not be the same as the way the correctly and naturally spoken word "cigarette" is processed. Although this argument may seem logical in interpretation of the "shigarette" case, it fails to provide sound explanations for the challenge from another line of behavioral literature which argues that late segments of words generate important lexical searches (e.g., Radeau, Morais, & Segui, 1995; Shillock, 1990; Tabossi, 1993; Slowiaczek, Nusbaum, & Pisoni, 1987); this obviously runs counter to any predictions based on the sequential hypothesis. For instance, the identification of monosyllabic words in a noisy background has been shown to be facilitated by a preceding prime that shares one or more phonemes with the target, no matter where the shared phonemes are located in the target word (Slowiaczek, Nusbaum, & Pisoni, 1987). In addition, Connine, Blasko, and Hall (1991) provide converging evidence against the sequential hypothesis by demonstrating that contextual information after the offset of a word can also help disambiguate alternative lexical hypotheses of the acoustic input. Taken together, all these data demonstrate that the processing of spoken words is not strictly sequential. Instead, input occurring late in time has a certain impact on the processing of information occurring earlier.

Parallel to these studies, a substantial body of literature calls into question the viability of the UP effects in spoken word recognition. According to the sequential hypothesis, recognition operates in a strictly time-shadowing fashion. As each word is identified, the boundary of that word can be subsequently identified. In order to work efficiently, such an account demands that words can be reliably recognized before their offset. However, a seminal study done by Luce (1986) showed that when frequency is taken into account, over one third of words do not uniquely diverge from all other words until after their offset. It is noteworthy that 84% of polysyllabic words have at least one shorter word embedded in them (e.g., "carbon" contains *car*) and 63% contain more than one embedded word (e.g., "cartoon" contains *cart* and *car*) (McQueen, Cutler, Briscoe, & Norris, 1995). A majority of these embedded words appear at the onsets of the polysyllabic words (McQueen & Cutler, 1992). In this regard, even if the input information is intact with no variability, the sequential models cannot lead to accurate recognition for a considerable number of words. Words such as *cart, car*, and *cartoon* would not be deactivated until a pause is heard at the offset of the word *car*. In fluent speech without reliable word boundary cues, *car* may need a few more phonemes after its offset to be disambiguated. In line with this view, in theory, more than a third of words in daily-use English will be recognizable only after part of the following word is heard (McQueen, Norris, & Cutler, 1994). In fact, using the gating paradigm, experimental data further underscore the inability of the sequential hypothesis to deal with embedded words (Bard, Shillcock, & Altmann, 1988; Grosjean, 1985). For example, using a gating paradigm Bard and colleagues (1988) presented participants with words that were interrupted prior to their offsets (e.g., the presentation of a word such as “controversy” could be interrupted after /kontr∂/). The subject was required to guess the identity of the full word. By varying the locus of the interruption, it can be established how much of the word is required to be presented so that reliable recognition can be reached. The results in this study revealed that many words (short ones in particular) in fluent speech were frequently not identified until some time after their acoustic offset. Another study further demonstrated that even if the UP of a word occurs early, the decision on the recognition of that word might require the processing of post-UP segments too (e.g., more than the first four phonemes would be processed in *rhinoceros*) (Goodman & Huttenlocher, 1988). In line with this view, Taft and Hambly (1986) also provide compelling evidence demonstrating that the processing of a pseudo-word continues after the point at which there are no possible continuations that would make it a word.

The most salient problem of sequential models, as Norris, McQueen, and Cutler (1995) observe, is that the system will be "paralyzed" if recognition fails in the middle of an utterance given the principle that word processing can proceed only in a strictly left-to-right manner. In order to remedy this drastic weakness, Marslen-Wilson (1987) revised the Cohort model (Cohort II) to adjust the emphasis on signal onset and re-consider the role of the information that occurs after the UP. He emphasizes in the Cohort II model that word recognition is an event of relative activation rather than a cohort's steady reduction to one candidate. Although the Cohort II tends to be more flexible than its earlier version, its loss of temporal predictability as to the UP ultimately separates Cohort II from the sequential hypothesis (Mattys, 1997).

1.1.1.2 The "goodness of fit" hypothesis

In contrast to the sequential hypothesis in which word-initial phonology has primacy in initiating lexical access, the 'goodness of fit' hypothesis proposes that a word is recognized on the basis of its overall goodness of fit between the complete input and a given lexical representation and is based on a comparison of the input's goodness of fit to other potential candidates. The advantage of this hypothesis is that the identification of the word *pleasant* with an input error (e.g., *bleasant*) will not cost much because the extent of overlap with the lexical representation of *pleasant* is high and there is no other word candidate in the lexicon to which *bleasant* is a better fit. The influential representatives of this hypothesis in the auditory domain are the interactive TRACE model (McClelland & Elman, 1986) and the bottom-up Shortlist model (Norris, 1994), in which recognition is derived from a process of competition[1] between multiple lexical candidates starting at many different points in the acoustic stream.

According to the TRACE model, the cognitive system consists of nodes structured into feature, letter, and word levels. Excitatory activation spreads between levels. Within levels, however, there is lateral inhibitory activity so that nodes corresponding to overlapping candidates are interconnected with inhibitory links. Meanwhile, the more information shared by lexical candidates, the stronger the lateral inhibition between them. The most striking feature of this model is that "this process of competition is in principle open to any word in the vocabulary at any time, which is rendered possible by the alignment of a complete copy of the lexical network with each point in the input where a word might begin" (Norris, McQueen, & Cutler, 1995, p.1210). In line with this view, as well as given the fact that a very large number of candidates will be activated if the vocabulary is not trivially small, TRACE is regarded as a very computationally expensive model.

Alternatively, based on the Shortlist model, the cognitive system is made up of two clear-cut stages. At the first stage, any lexical candidates that can begin at any phoneme in the acoustic input are activated in a completely bottom-up fashion. Given that there is no interactive activation between potential candidates at this stage, the process in this stage does not take into account whether candidates share part of the

---

[1] Compared with TRACE and Shortlist, competition in Cohort II is passively completed through the analysis of the goodness of fit between the candidates and the acoustic-phonetic information.

acoustic input. It is not until the second stage that these candidates are linked into a small interactive activation network in which overlapping candidates compete in a way similar to how overlapping lexical nodes compete in the TRACE model. What is different from the TRACE model, however, is that the number of words permitted to enter the competition is relatively small because the competition occurs only between candidates (the "shortlist") for which there is some bottom-up evidence.

1.1.2 The common downside of the aforementioned speech hypotheses

It is worth noting that both the "sequential" and "goodness of fit" hypotheses lend the written language notion of words to spoken word recognition models. The assumption behind each is that the spoken word is an acoustic-phonetic analog of the written word, which is processed as a "unit" in the speech recognition process. It is further believed that the process that auditory word recognition undertakes is analogous to the domain in which printed word processing takes place (Grosjean & Gee, 1987). Although there is a large body of language literature arguing that there is a common semantic system shared by the written and spoken language comprehension (Kutas & Federmeier, 2000), there has been little empirical evidence in favor of the assumption that the written supposition of the word plays a role in the earlier stage of auditory lexical processing. Most critically, this excessive importance given to the written language notion of the word in the lexical access of spoken words leads to a major shortcoming of the two hypotheses. That is, the two hypotheses have trouble segmenting continuous and complex acoustic streams (speech) into discrete words, which is obviously not a problem in processing continuous text given that there are little white spaces between written words that serve as explicit word-boundary markers.

In point of fact, in an attempt to locate the speech analog of white spaces in text, several markers for word segments have been pinpointed and studied; these include pauses, fundamental frequency contours, bursts, aspiration, glottal stops, and allophonic variations (e.g., Garding, 1967; Lehiste, 1960; Nakatani & Dukes, 1977). Unfortunately, empirical data undermine any of them acting as a stable and reliable indicator of word boundaries (e.g., Klatt & Stevens, 1973; Lehiste, 1972; Reddy, 1976). In order to avoid the trouble in locating word boundaries, proponents of these two hypotheses claim that a word's boundaries are clarified only after the word is recognized (Cole & Jakimik, 1980). In line with this view, word segmentation in fluent speech is a by-product of the ongoing processes of word recognition. For instance, word boundary is demonstrated to be clarified by semi-exhaustive (in the Shortlist model) or fully exhaustive (in the revised Cohort II and TRACE models) activation of lexical candidates at any point in the acoustic input (Mattys, 1997).

However, this view is highly debatable, as one should keep in mind that lexical access operates with discrete units in continuous speech. Logically, word segmentation should be an essential and explicit process that occurs prior to lexical activation. Cutler emphasizes (1992) that "when a recognizer is presented with continuous speech… it cannot begin the process of lexical access until it has taken some decision about how the stream of continuous speech should be segmented into units that one might reasonably expect to be matched in the lexicon" (p.343). In fact, a rich and robust variety of evidence attests to this possibility. An increasing number of speech studies have demonstrated that prosodic information can provide reliable cues for speech segmentation. In particular, stressed syllables of an English word are suggested as reliable perceptual 'anchors' to parse the speech input into words (e.g., Nooteboom, Brokx, & de Rooij, 1978; Nakatani & Schaffer, 1978; Carlson, Grandstrom, Lindblom, & Rapp, 1973; Cutler & Butterfield, 1992).

1.1.3 The introduction of prosodic information into English spoken word processing theory.

Not a single language in the world is spoken without prosody. In general, speech prosody refers to the rhythmic patterns associated with speech's many phonetic elements (e.g., tone, stress, intonation, and rhythm) (Cutler, 1989; Gandour, Wong, & Hutchins, 1998). Variations in speech prosody offer valuable auditory information to phonological, lexical and grammatical codes (Gandour, Wong, Hsieh, et al., 2000; Gandour, Ponglorpisit, Potisuk, et al., 1997). A great deal of literature has underscored the fact that prosodic information plays an important role in speech comprehension. For example, it has been demonstrated that variations in the fundamental frequency contour have a clear impact on the intelligibility of sentences (Slowiaczek & Nusbaum, 1985; Sorin, 1983). Given the role prosody appears to play on ongoing language processing at the sentential level, it is reasonable to address the question of whether word prosodic cues have any effects on spoken word processing.

English language is a stress-based language. Word prosody in English refers to the stress pattern of an English word. Generally speaking, there are two means to define stress. A liberal definition of stress simply divides syllables into strong and weak syllables (Cutler & Norris, 1988; Norris, McQueen, & Cutler, 1995). Strong syllables are those containing a full vowel (e.g., /æ/) while weak syllables are those involving a reduced vowel (i.e., schwa /∂/). This dichotomous division of stress is called "metrical". The other definition of stress is relatively more lexically based and also more conservative (Mattys, 2000). According to this definition, strong syllables are further classified into primary stressed syllables, secondary stressed syllables, or unstressed vowel-unreduced syllables. For example, with respect to the multi-syllabic English word "generate", there are three stress levels for each syllable of the word: the first syllable "ge" has the primary stress, the syllable "rate" bears a kind of stress that is secondary to the primary stress, and the syllable "ne" is the unstressed. In English language, there are no routine regulations for the particular position for the primary stress in words. The primary stress can either occur on a word-initial syllable (e.g., "generate") or be placed on a word-non-initial syllable (e.g., "discount") (Cutler, 1992). What is noteworthy, however, is that, for each discrete English word, the primary stress occurs at a specific position either in isolated presentation or in context.

The intelligibility of stressed syllables in speech perception has long been taken as a well-established fact in phonetics as well as in cognitive psycholinguistics. Phonetic analyses have shown that the stressed syllable presents salience in terms of pitch, duration, and amplitude: it is longer in time, higher in pitch, and greater in amplitude (Chomsky & Halle, 1968; Fry, 1955, 1958; Lehiste, 1970; Umeda, 1977). In a phoneme detection task, for example, it was found that reaction times were faster when the phoneme was located in a stressed syllable than in an unstressed syllable (e.g., Cutler & Foss, 1977). Some other studies have also reported that when mispronunciations occur on stressed syllables they are detected more frequently than when they occur on unstressed syllables (Cole & Jakimik, 1978, 1980a, 1980b). Equally, the stressed syllables of words are misperceived or misinterpreted less often than the unstressed syllables (Bond & Garnes, 1980). Stressed syllables are identified more easily and consistently than unstressed syllables in noise backgrounds (Kozhevnikov & Chistovich, 1966) or in continuous speech (Liberman, 1965). They are also less fluently restored than unstressed syllables in shadowing tasks (Small & Bond, 1982).

The relevant effects stressed syllables appear to have on speech perception call great attention to the important role that stressed syllables may play during spoken word recognition in continuous speech. According to this account, a third "stress guide" hypothesis has been proposed (see review for Mattys, 1997). Contrary to the aforementioned "sequential" and "goodness of fit" hypotheses in which lexical segmentation emerges as a consequence of speech recognition, the "stress guide" hypothesis claims that the identification of word boundaries is a discrete process which can in fact facilitate spoken word recognition. Put differently, speech recognition involves a discrete procedure of segmenting continuing acoustic input so that lexical access can be guided by the knowledge of where word boundaries can most likely start. Given the salient status of stress in English language, the third hypothesis asserts that it is the stressed syllable that plays a crucial role in parsing fluent speech into words and initiating lexical processing.

The contribution of stress to speech processing has recently become a critical object of inquiry for a number of speech models, such as the "attentional bounce" model (Shieds, McHugh, & Martin, 1974; Pitt & Samuel, 1990). One of the most influential models based on this "stress based" hypothesis is the Metrical Stress Segmentation (MSS) model that was first proposed by Culter and Norris (1988).

On the basis of the metrical partition of stress, the MSS claims that the segmentation of speech stream is triggered by strong syllables at the onset of which the cognitive system starts a lexical access attempt. Using word-spotting methodology, Cutler and Norris (1988) presented English listeners with a list of bi-syllabic pseudo-words that ended with either a strong (e.g., thintayf or mintayf) or weak (e.g., thintef or mintef) syllable and asked them to judge whether or not a pseudo-word they had heard began with a real word (e.g., "thin" in "thintayf"). According to the MSS, it is expected that the second strong syllable in the strong-strong string "mintayf" triggers segmentation of the input into "min" and "tayf". This segmentation thereby disrupts the recognition of the embedded real word "mint". As for the second weak syllable "tef" in the strong-weak string "mintef", such a segmentation does not occur so that the real word "mint" can be easily detected. The findings obtained in this study fully support this expectation: the target "mint" was less likely to be detected in the string "mintayf" compared to the string "mintef", while the detection of the target "thin" presents no difference between the cases of "thintayf" and "thintef". Another piece of supporting evidence for the MSS comes from a "slips of the ears" study. Cutler and Butterfield (1992) examined misperceptions of the locations of word boundaries in fluent speech. They presented unpredictable and barely audible sentences that consist of strings of alternating strong and weak syllables. In line with the MSS, the results of this study demonstrated that listeners had a strong tendency to locate word boundaries at the beginning of strong syllables. On the basis of these data, it seems likely that English listeners practice a segmentation strategy that strong syllables are reliable markers for the onsets of words.

Further support in favor of the stress guide hypothesis comes from lexical distribution data. It is argued that the strategy of stress-guided segmentation, in theory, yields a high rate of correct word boundary detection. According to the lexical statistic data of the English vocabulary, most English words that contain content in everyday English are initially stressed (Cutler & Carter, 1987). Cutler (1992) reported that when their frequency of occurrence is taken into account, about 85 percent of content words start with stressed syllables. In this regard, for initial stressed words, the "stress guide" hypothesis would accurately guide the system to identify the word onset. With respect to the historical change in English language, it is also noted that words that gradually carry little informational saliency turn out to lose part of their phonological substance and eventually become function words or affixes to other words (Givon, 1975). Such historical changes in conjunction with the distribution data give rise to the viability of the stress guide hypothesis.

The viability of the stress guide hypothesis has not gone unchallenged, however. The major question about the feasibility of the stress segmentation strategy is that this hypothesis does not provide any clear account for how non-initial-stress words are processed. As proponents of the stress base hypothesis underscore, given the stress based nature of English language itself, stressed syllables are seen as "the milestones of the listener's attention" (Mattys, 1997, p.322). On the basis of this view, however, segmentation of word boundaries initiating at the onset of stressed syllables would result in faulty lexical processing on non-initial-stress words. This issue cannot be neglected given that there are about 10 percent of words starting with unstressed syllables in daily use English.

In an attempt to delineate the process during non-initial-stress word recognition, proponents of the hypothesis first maintain the credibility of the argument that lexical processing primacy is not given to temporally early-occurring acoustic information (in a strictly time-shadowing fashion) but to that which is perceptually salient (stressed syllables). At the same time, they further argue that word processing can proceed in a flexible time-independent fashion. It is reasoned that although stressed syllables in natural speech are capitalized from one to the other in a time-shadowing manner, the process initiated by stressed syllables can also realize and decode acoustic information immediately before and after stressed syllables in a proactive and retroactive fashion (Mattys & Samuel, 1997, 2000). To accommodate this hypothesis, a retroactive mechanism is proposed, which is argued to repair faulty lexical segmentation by backtracking to information presented earlier in time on the condition that lexical access through strong syllables is not successful. In this regard, the segmentation of speech using stressed syllables as onsets of words can be supplemented by a retroactive processing by which lexical search starting with initial unstressed syllables can be successfully accessed.

A number of experimental data have accumulated to substantiate the proposal that the initiation of lexical search on stressed syllables is accompanied with a retroactive processing during non-initial-stress spoken word recognition. For instance, in the identification of spondees[2] presented in white noise, Cluff and Luce (1990) found that accuracy of identifying final syllables was not influenced by the frequency and phonetic

---

[2] Spondees are bi-syllabic compound words like *chestnut* or *deadlock* that bear two stressed syllables.

neighborhood density of the initial syllable with which the final syllable was paired. The results showed that subjects' performance on the hard-easy spondee whose initial syllable was a low frequency word with high phonetic neighbors and whose final syllable was a high frequency word with few phonetic neighbors (e.g., "hacksaw") was nearly equal to the performance on the easy-easy spondee the two syllables of which are high frequency words with low phonetic neighbors (e.g. "catfish"). Similarly, accuracy of identifying final hard syllables was at nearly equal level of performance regardless of the type (easy or hard) of the initial syllable. What is interesting is that the pattern of results for final syllables was not obtained for initial syllables. It was shown, for example, that identification performance for initial easy syllables was decreased when the easy syllable was followed by a hard syllable (e.g., the easy-hard spondee "madcap") but increased when followed by an easy syllable (e.g., the easy-easy spondee "catbird"). The authors interpreted the data as suggesting that later occurring information (e.g., the final easy syllable of a spondee) can be used to resolve the ambiguity that has arisen in an initial hard syllable whereas earlier occurring information (e.g., the initial easy syllable of a spondee) would do little to aid in identifying the final syllable (the easy-hard case). The data have therefore been invoked extensively to support the idea of the retroactive strategy in spoken word recognition. It seems likely that under certain conditions (e.g., initial lexical uncertainty), retroactive processing occurs in continuous speech.

The viability of this retroactive position is further bolstered by a phoneme-detection study. In this study, Mattys and Samuel (2000) conducted three experiments. In the first experiment, subjects were presented with a list of seven words (e.g., list 1: sauna—gazelle—awkward—profane—depart—dissect—pervade; and, list 2: water—

basic—dictate—sage—zealous—furry—beyond) with no pauses between them. Subjects were instructed to detect the initial consonant (target) of a pre-specified syllable (e.g., /g/ of /g∂/ or /z/ of /zel/). The target-carrying syllable was either stressed (e.g., /zel/ in "gazelle") or unstressed (e.g., /g∂/ in "sage") and was either the first (e.g., /g∂/ in "gazelle") or second syllable (e.g., /g∂/ in "sage") of a bi-syllabic word. Meanwhile, initial stress and non-initial stress words were matched phonemically and acoustically on their target-carrying syllables. It was predicted that if the retroactive strategy is exploited during non-initial-stress word perception, a penalty – a delayed process, should be imposed on non-initial-stress words. The finding in this experiment was consistent with this prediction in demonstrating that the phoneme-detection reaction times were shorter for initial-stress words than for non-initial-stress words, no matter where the target was in the word and no matter whether or not the syllable that the target was embedded in was stressed. In their second experiment, they attempted to examine the cost of retroactive processing by monitoring the detection times to word-initial phonemes in a strong-weak and weak-strong word in which the syllable that the target phoneme was embedded in was identical through splicing. For instance, /p/ in PERmit contrasts with perMIT or /k/ in CAMpus with camPAIGN). The results of this experiment demonstrated that the detection times were prolonged for a word-initial consonant in a word with late stress (e.g., perMIT or camPAIGN) than in a word with initial stress (e.g., PERmit or CAMpus). These performance data are believed to provide compelling evidence for the retroactive strategy utilized in non-initial-stress word processing. It is logical to argue that there should be a substantial demand on memory if the retroactive processing is required given that the information prior to the stressed syllable of a non-initial-stress word needs to be kept available in the memory store. In line with this view, the authors carried out a third experiment to limit the phonetic memory sources during spoken word processing by asking subjects to perform two simultaneous cognitively demanding tasks. In support of the retroactive processing view, the results demonstrate that non-initial-stress words demand greater mnemonic needs than initial-stress words. Overall, the findings in this study indicate that words with different stress patterns are processed differently: non-initial-stress words require extra processing so that early information can be incorporated into the encoding of later stressed sounds to assist inappropriate lexical access begun on a later strong syllable.

It is worth keeping in mind that the behavioral data particularly represent the complex combination of several processes that involves extensive post-lexical bias effects. Although the data collected in these behavioral studies support the argument that non-initial-stress words require an additional, time-consuming backtracking processing, the data fail to provide a direct account of the mechanisms underlying the non-initial-stress word recognition. Merely based on the performance data, we cannot screen out the possibility that the delayed responses to non-initial-stress target words in the aforementioned studies reflect delayed post-lexical processes rather than activity at the prelexical or lexical stage. In this regard, the discrete processes of lexical access remain to be explored.

1.1.4. Sub-lexical phonological priming in spoken English word recognition

Most theories of word recognition are in general agreement that contextual knowledge has an impact on lexical processing (Henderson, 1982; Henderson, Wallis, &

Knight, 1984; Kutas & Van Petten, 1988). First and foremost, there exists a considerable body of empirical findings that point to contextual semantic priming effects on lexical processing. The basic and general experimental design is to examine the contextual facilitation effects with various reaction time measures. In a typical word-pair paradigm, for instance, subjects are presented with a stream of words upon which to perform some task such as naming or lexical decision (i.e., judge whether or not a string of letters is a real word). Using this paradigm, the classical semantic priming effect refers to the phenomenon that both speed and accuracy are improved if the target word (e.g., "dog") is preceded by a semantically related word (e.g. "cat"—prime) as opposed to a semantically unrelated word (e.g., "table"), a non-word, a "pseudo-word" (pronounceable non-word), or some non-linguistic stimulus such as a line of *'s (Bolinger, 1965; Collins & Quillian, 1969; Johnson-Laird, Herrman, & Chaffin, 1984; Katz & Fodor, 1963; Rips, Shoben, & Smith, 1973; Smith, Shoben, & Rips, 1974). This response facilitation for related words has been taken as empirical evidence in favor of the argument that the nodes corresponding to discrete words in the mental lexicon are organized and linked in some meaningful fashion. Put differently, the above facilitation reflects some abstract knowledge-based association between the target and its preceding contextual information (either words or sentences) (for a review, see Neely, 1991).

Faced with the semantic contextual effects on word processing in both speech and written text, it is logical to ask whether contextual information exerts any effect on form based (phonology and/or orthography) representations and whether form-based contextual cues play any role in semantic activation of the word. Form-based processing in the auditory modality, according to Marslen-Wilson (1989), is described as "the basis for a mapping of the speech signal onto the representations of word forms in the mental lexicon" (phonological processing) (p.3). Although it is evident that reading or hearing one word has influence on the speed and quality of the processing of associated words not only in the semantic domain but also in the form based domain, the effects of form based contextual information on ongoing lexical processes remains elusive.

At present, the interaction of contextual factors with form-based information in word recognition has been documented in numerous studies. One line of performance evidence in favor of the argument that the form based contextual cues come into play in word recognition is derived from a group of repetition priming studies. It has been found that the presentation of words or other items (e.g., drawings) influences later performance on those items. This repetition effect, in fact, can even be reliably observed in a task that does not make explicit reference to the past study episode, such as the word perceptual-identification task (Jacoby & Dallas, 1981) and the word-stem completion task (Graf, Mandler, & Haden, 1982). This supporting evidence has been questioned by the possibility that the priming effects observed do not reflect phonological (if in speech) or orthographic (if in reading) contextual information on word processing because many other variables (e.g., meaning) are also involved in this process and thereby bias the results.

Notwithstanding these criticisms, a variety of other performance data underscore that the influence of contextual information on lexical processing includes more than the associations in the semantic domain. Indeed, an increasing amount of literature has suggested that the phonotactic (Luce, 1986; Pisoni, Nusbaum, Luce, et al., 1985), phonological (Pisoni, Nusbaum, Luce, et al., 1985; Slowizczek, Nusbaum, & Pisoni, 1987), and orthographic (Hillinger, 1980; Jakimik, Cole, & Rudnicky, 1985) information about a word reflects attributes of a corresponding node in the mental lexicon and its relevant contextual information plays some role in word recognition during listening and reading.

In view of the common agreement that form based contextual information contributes to ongoing linguistic processing, the question at stake is whether form based information has a facilitation or inhibitory effect on word recognition if only partial cues about the word (e.g., the phonological cue /d/ for the word "desk") are provided. In an attempt to clarify this question, a number of studies were conducted and computational theories were developed.

Some compelling performance data demonstrate that there is a facilitation effect of sub-lexical form based information on word recognition. In a seminal lexical decision study on phonologically and orthographically related stimuli, Hillinger (1980) presented subjects with pairs of rhyme words (e.g., BRIBE-TRIBE) and asked them to decide whether or not the target item is a real word (lexical decision task). The finding demonstrated a reliable response facilitation effect: subjects responded much faster to the visually presented words in a pair that were similar orthographically and phonologically (BRIBE-TRIBE) than to control pairs (BREAK-DITCH). The phonological facilitation effect was further replicated under the condition that the rhymed prime-target pairs were orthographically dissimilar (e.g., "eight—mate") and when the prime was presented aurally and the target remained visually presented. Based on these findings, it seems tenable to argue that sub-lexical form-based cues provide facilitation functions on whole word recognition. However, this argument was questioned by some results obtained in the same visual lexical decision task: contrary to the aforementioned facilitation findings, an inhibition effect was instead observed as the response times were found prolonged when the prime-target pairs (TOUCH-COUCH) shared only orthographic attributes.

In a similar vein, inconclusive results were observed in some other behavioral studies. For instance, a sub-lexical contextual facilitation effect was seen when the auditory target (e.g., car) was preceded by a prime that started with the same one or two phonemes (e.g., carbon) in identification-in-noise, lexical decision-in-noise (Goldinger, Luce, & Pisoni, 1992), and shadowing tasks (Slowiaczek & Hamburger, 1992). However, the findings that run counter to the facilitation data were also observed when the auditory target was preceded by a prime that shared more than three phonemes (Hamburger & Slowiaczek, 1996; Slowiaczek & Hamburger, 1992).

In view of these inconclusive results from performance studies, one of the intentions in this thesis is to explore the neural correlates of sub-lexical form-based effects on semantic representations of words during listening and reading, the consequence of which can provide a better insight into the architecture of the organization of word processing. This thesis seeks to define the role of syllable-sized phonological information in multi-syllabic word processing in spoken English sentence comprehension (Experiment 1) and to explore the function of discrete onset and rime phonological information in Chinese character recognition during Chinese proverb listening (Experiment 2) and reading (Experiment 3) comprehension.

1.2 Lexical prosody and segmental phonology in spoken Chinese character processing

In comparison to the salient stress and phonological complexity of English words, Chinese characters present different physical attributes. Chinese is a tonal based language. Chinese characters consist of monosyllables affiliated with word prosody by virtue of the tonal pattern. In this regard, syllable structure of a Chinese character in the phonological domain is relatively regular. Meanwhile, contrary to the stress pattern in English, tonal information in Chinese is relevant for lexical purposes (Lee & Nusbaum, 1993; Gandour, Wong, & Hutchins, 1998; Gandour, 2000). In Chinese, segmental (onset and rime) distinctions are not the only criteria used to distinguish words but contrasting tonal variations as well. For example, in Chinese Mandarin, although two characters may be homonyms insofar as they share the same onset-rime pattern, they can have totally different meanings due to tonal variations (麻'ma2'-linen; 马'ma3'-horse) (Chao, 1948, 1968).

It is interesting to note that while the majority of the world's languages are tonal in nature (Fromkin & Rodman, 1993; Gandour, 2000), most of theories regarding speech processing have been derived from stress based language studies (e.g., English) and not many studies have focused on the specific features of tonal language speech. It is thus of great importance to incorporate an understanding of the nature of tonal speech processing if we hope to shed light on language-universal mechanisms.

It is clear that from an acoustic vantage point, voice fundamental frequency states and movements are the primary carriers of tonal information in tonal languages (Cutler & Chen, 1997; Gandour, Wong, Hsieh, et al., 2000; Ye & Connine, 1999). However, little explicit evidence exits regarding the relationship between segmental and tonal perception as well as the contribution of tonal information to ongoing linguistic processing.

Most of the related literature argues that, although tones are lexically meaningful in tonal languages, tonal information is still a supra-segmental element of language (Miller, 1978; Nusbaum & Lee, 1993; Repp & Lin, 1990; Ye & Connine, 1999). Like other prosodic information such as intonation and syntactic stress, word tones have been seen as being carried upon segmental phonology. Contrary to this view, however, Goldsmith (1976) argues that tones in tonal languages are "segmental in their own right", and has proposed the term 'autosegmental' for tones. He underscores that segments and tones do not depend or ride upon each other but can be seen as independent of one another. Viewed in this light, it is logical to ask "are tones and segments processed in an integrated or separate manner during ongoing linguistic processing?" Likewise, one must ask whether they are processed in the same neural network as a single, integral structure.

The majority of behavioral studies (Gottfried & Suiter, 1997; Lee & Nusbaum, 1993; Miller, 1978; Repp & Lin, 1990) addressing these two concerns were conducted using a speeded classification paradigm that was first introduced by Garner (1970). This paradigm is argued to reveal the nature of the interactions between different dimensions of speech information (Garner & Felfoldy, 1970). In this paradigm, subjects are required to attend to variations along one dimension while trying to ignore changes along a second dimension. For example, in Wood's (1974) study, there were four synthesized stimuli: /ba/ and /da/, affiliated with either a high (140Hz) or a low (104Hz) constant F_0. When pitch (high or low) was chosen as the target dimension, the participant was asked to merely focus on the change in pitch dimension while ignoring the variations in the consonant dimension (/b/ or /d/). The theory behind this paradigm is that listeners will have difficulty in selectively attending to only one dimension if two dimensions (i.e., consonant and pitch) are processed in an integrated fashion. In this case, the extent to which two dimensions are integrated can be indicated by reaction times (RTs).

Using the speeded classification paradigm and choosing native English listeners as participants, it was found that there was an asymmetric interference in the processing of consonant and tonal information. That is, tonal variation slowed down consonant identification decisions, whereas consonantal variation had no influence on tonal pitch decisions (e.g., Repp & Lin, 1990; Wood, 1974). However, symmetric interference was observed in the discrimination of vowel quality and tonal perception. That is, tone change interferes with the processing of vowel as much as vowel change with the processing of tone (Carrell, Smith, & Pisoni, 1981; Miller, 1978; Repp & Lin, 1990). Interestingly, using the same paradigm but applying Chinese language-related tones to Chinese listeners, Repp and Lin (1990) found that, unlike English listeners, Chinese listeners showed an underlying processing asymmetry between vowels and tones in consonant-vowel (CV) syllables. This finding was interpreted as supporting evidence for the hypothesis that vowels and tones in Chinese may not be processed in an integrated fashion. However, given a paucity of empirical evidence on the nature of segmental and tonal processes in Chinese, the relation between segments and tones remains to be further elucidated.

At the same time, the aforementioned differences in performance results based on the testing of different language populations have encouraged researchers to define the role that tonal information may play in tonal language processing. It has been argued that tonal information serves to distinguish lexical identity in tonal languages. A line of compelling evidence bearing on this issue is derived from neuropsychological investigations of hemisphere functions. For example, in a dichotic-listening task (Van Lancker & Fromkin, 1973), native speakers of Thai demonstrated a right ear advantage for Thai tonal variations. Interestingly, this right ear advantage disappeared when Thai tones were replaced by hummed stimuli (i.e., sounds that carried tonal features but were unrecognizable as words and thus carried no linguistic meaning). In a similar vein, some clinical studies reported that left brain-damaged aphasics seemed to meet with much greater difficulty in distinguishing words that minimally differed in tones than normal subjects, while this difficulty could not be observed in the right brain damaged patients (Gandour, 1988, 2000). Given the view that the left hemisphere plays a predominant role in processing lexical information, this left lateralization for tonal features was interpreted as compelling evidence for the proposal that tones in tonal languages are represented in the mental lexicon.

Another line of convincing evidence for the engagement of tonal information in lexical processes comes from some psycholinguistic performance studies. For instance, in a tonal discrimination task, Lee, Vakoch, and Wurm (1996) found a lexical effect when lexical and non-lexical tones in Chinese Cantonese were presented to native Cantonese listeners. However, such an effect was not observed when the same materials were presented to native English and to native Chinese Mandarin listeners.

The lexical view of tones, however, is not without controversy. For example, Chao (1980) states that in tonal languages, the importance of tones to a spoken word cannot be overemphasized so that a "syllable of the same consonant and vocalic composition but pronounced using a different tone is perceived as a thoroughly different word" (p.42). In line with this statement, several lines of empirical data appear to run counter to the lexical view. For example, using the same dichotic listening technique as Van Lancker and Fromkin (1973), Yang (1991) found no right ear advantage in language related tonal identification. These conflicting results were taken as a counter example to challenge the validity of the left hemisphere dominance theory in tonal processing. Evidence questioning the effects that tonal information may exert in lexical processing also comes from a visually presented homophone judgment experiment (Taft & Chen, 1992). The results in this experiment demonstrate that native Chinese speakers gave less rapid and less accurate responses when the pronunciation of the two characters differed only in tonal dimension, as opposed to in segments.

In a similar vein, Cutler and Chen (1997) carried out an auditory identical judgment experiment and found that listeners without any knowledge of Cantonese showed an identical same-different judgment pattern for Cantonese syllables as the native Cantonese listeners. Nevertheless, there were more errors in lexical decision when non-words mismatched real words in the tonal dimension compared with a mismatch in segmental attributes. Based on these findings, it was concluded that tonal processing is only achieved at a simple perceptual level without any lexical specifications.

Some caution, however, has to be exercised in drawing firm conclusions about the effects of tones on on-going language processing. It has been argued that behavioral performance approaches (e.g., the response time and error rates) usually measure the 'end-state characteristics of a process' (Swinney, 1981, p 307), which unavoidably lead to extensive post-lexical bias effects if earlier stages of language processes are of interest. In this regard, these measures may obscure the nature and form of processes that occur earlier in the language sequence. For instance, the results of the study by Taft and Chen

(1992) showing less accuracy in tonal change than in segmental change may not necessarily imply that tonal information does not come into play at the lexical stage. Indeed, based on these performance data, there is no evidence to undermine the possibility that slow responses to tonal variation merely reflect the nature of judgment-related post-lexical processes.

The drawback of behavioral measures renders the structure and process of tonal and segmental information difficult to establish. In an attempt to delineate the functional locus and time course of tonal and segmental processes during Chinese Mandarin listening and elucidate the relationship between tones and segments, experimental methods and designs sensitive to both temporal and spatial domains are required.

1.3 The roles of segmental phonology and prosodic phonology in silent Chinese reading

The role of phonology in visual word recognition has been a key question in reading research for the past 20 years. Although there is an abundance of accumulated data, the major controversies surrounding phonological functioning in reading are still far from resolved. Central to the inquiry remains the question of whether “to read, in effect, is to translate the writing into speech” (Edfelt, 1960, p.14).

It has been widely accepted that the foremost association developed in the process of language acquisition is the contact between spoken words and meaning. Levelt, Roelofs, and Meryer (1999) observe that infants (literally derived from Latin *infans*, meaning speechless) are human beings who cannot speak. It takes most of us the whole first year of our lives to learn a simple mapping from sound phonology to semantic meaning in order to understand and produce meaningful spoken words. In a similar vein, it is believed that learning to read may start with the development of orthography-phonology mapping so that words can be overtly articulated and understood via the original sound phonology-semantic route for spoken language (Taft & Van Graan, 1998). By this account, a number of researchers (Blank, 1978; Siegel, 1985; Stanovich, 1991) have postulated that phonological functioning plays a critical role in the development of reading skills. In fact, a large number of studies on the reading skills of reading disabled children provide convincing evidence for the argument that the deficits in basic phonological skills are the basis of a reading difficulty in alphabetic (e.g., English) as well as logographic (e.g., Chinese) languages (Rack, Snowling, & Olson, 1992; So & Siegel, 1997).

However, of crucial interest is the question as to whether phonological decoding is still central to reading comprehension when skilled readers perform reading. Can the printed word be processed directly from orthography to meaning without any phonological involvement when reading proficiency is reached? Although hundreds of studies conducted over the past several decades have attempted to answer these questions, such issues have consistently remained on the agenda.

Regarding the different views on phonological functioning in reading, there are three major theories concerning the nature of visual word recognition. The first is the 'phonological recoding' theory. The major feature of this theory is its emphasis on the primary role of phonological processing in word recognition (Frost, 1998; Van Orden, 1987, 1990). Proponents of this theory argue that phonological processing is indeed the default procedure of the cognitive linguistic system. According to this account, the process of reading is believed to proceed with a visually presented word being converted into its phonological codes in order to activate its semantic representation (Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994; Peter & Turvey, 1994).

The second theory "phonological mediation" claims that phonological information exists and is automatically activated in the course of silent reading. In contrast to the first theory in which phonological processing in written word recognition is central, advocates of the second theory argue that the existence of phonological activation does not mean that it is indispensable to transform orthography to phonology in order to activate semantic meaning. The existence of an orthographic-to-semantic route is considered to be a viable proposition. At present, there are two influential models in favor of this theory. The first is the dual-route model that further proposes that there are two independent routes to lexical meaning: a direct visual route and a back-up phonological route (Coltheart, Davelaar, Jonasson, et al., 1977; Coltheart, Curtis, Haller, et al., 1993). The direct visual route is responsible for the processing of the physical features of the word form and is generally the primary approach to lexical access for skilled readers. Similar to the dual-route theory, the second named "parallel-distributed time-course processing (PDP)" model also provides two routes for reading (Seidenberg & McClelland, 1989). What is different from the dual-route model, however, is that, according to the PDP model, the two routes are not independent of each other. Rather, both the visual and the phonological pathways interact within a distributed network to activate the mental representation of a written word.

Using different experimental paradigms in which semantic meaning is either required by the task or appears to be involved in the task, a large body of behavioral research has been carried out to determine the impact of phonology in reading. The issue of phonological recoding has been extensively studied and been in particular inspired in tasks showing (pseudo)homophone confusion effects. One such task is lexical decision in which subjects are required to decide whether or not a visual stimulus (e.g., letter strings) is a word. Homophones are described as those words that have the same pronunciations but different derivations, spellings and meanings (e.g., *you* and *ewe*). Pseudo-homophones are non-words (e.g., *brane*) that are homophonic with a real word (*brain*). A typical finding in the lexical decision task is that subjects take more time to reject pseudo-homophone foils (e.g., *brane*) than control foils (e.g., *brene*) (e.g., Coltheart, Davelaar, Jonasson, & Besner, 1977; McCusker, Hillinger, & Bias, 1981; Rubenstein, Lewis, & Rubenstein, 1971). One influential explanation for this effect is that the pseudo-homophone *brane* encodes the phonological representation /brein/. The encoded phonological representation further activates the lexical code corresponding to the word *brain.* As a consequence, this lexical activation makes the speech cognition system take longer time to classify *brane* as a non-word. This pseudo-homophone effect is thereby taken as evidence for phonology playing a predominant role in written word recognition.

Another task showing homophone confusion effects is semantic categorization (e.g., Coltheart, Patterson, & Leahy, 1994; Peter & Turvey, 1994; Van Orden, 1987). In this task, subjects are required to judge, for example, whether LOOT that is homophonic with LUTE is a musical instrument. The common finding is that homophones and/or pseudo-homophones of category members produce a higher false alarm rate than orthographically similar controls.

Apart from these (pseudo)homophone confusion data, the involvement of phonology in reading has also been demonstrated in some other studies showing homophone facilitation effects. For example, the eye monitoring technique has been used in many studies to examine the effect that homophony may exert on eye movements and fixation during reading comprehension (Daneman, Reingold, & Davidson, 1995; Rayner, Pollatsek, & Binder, 1998; Rayner, Sereno, Lesch, et al., 1995). The underlying assumption is that the duration of fixation would be shorter if the processing of the target word is facilitated. Consistent with the phonological mediated theories, the results obtained in these studies demonstrated that the duration of eye fixation becomes shorter when homophones (Beech vs. Beach) were involved in contrast to visual controls (e.g., Bench vs. Beach).

Another line of literature employing a phonologically mediated semantic priming approach provides further compelling evidence for the role of phonological processing in reading (e.g., Lesch & Pollatsek, 1993; Lukatela, Lukatela, & Turvey, 1993; Lukatela & Turvey, 1994). It was found that naming times as well as lexical decision times for a target word could be facilitated not only by the semantically related primes (e.g., the target word 'war' in the sentence 'the dove is a sign of war (peace)') but also by homophones of semantic associates of the target word (e.g., the target word 'piece' in the sentence 'the dove is a sign of piece (peace)').

Taken together, these performance data lend strong support to the hypothesis that phonological activation is an intrinsic constituent of the reading process. In particular, the results from some performance studies further indicate that the sound of a word is indeed activated automatically even in tasks in which phonological information is not explicitly required (e.g., Ferrand & Grainger, 1992, 1994; Grainger & Ferrand, 1994; Zeigler, Van Orden, & Jacobs, 1997). One such task is letter search that requires only a shallow level of processing. It was hypothesized that phonological recoding should be seen as an automatic process if homophony effects can be observed in a letter search task that is in fact carried out on a purely visual basis. The results in such letter search studies support this hypothesis. It was demonstrated, for instance, that the identification of the letter "I" in the word "hail" was interfered with by the homophone of "hail" (hale) (Zeigler, Van Orden, & Jacobs, 1997).

The above arguments, however, are not without controversy. With regard to the findings in the eye monitoring tasks, there is controversy as to the locus of the phonological activation. Some argue that this phonological activation occurs earlier than semantic access (Pollatsek, Lesch, Morris, et al., 1992; Rayner, Sereno, Lesch, et al., 1995), while others believe that it comes into play only after meaning has been activated (Daneman & Reingold, 1993; Daneman, Reingold, & Davidson, 1995). More importantly, the homophone effect was not replicated in other eye monitoring studies (Rayner, Pollatsek, & Binder, 1998; Pollatsek, Lesch, Morris, et al., 1992). In semantic categorization studies, conflicting results were likewise observed (Jared & Seidenberg, 1991). Most interestingly, Taft and Graan (1998) did not find any regularity effect when asking participants to decide whether the target word belongs to the category "words with definable meanings" or the category "given names". Given the assumption that the regularity effect derived from the transformation of orthography to phonology should be revealed if semantics is to be accessed via phonological assistance, the results of this study were interpreted as strong evidence against the argument that phonological processing is a prerequisite for semantic activation. In an attempt to accommodate these controversial findings with reading models, the theory of "direct access" has been proposed, which stands in contrast with the first two reading theories in that it questions the influence of phonological activation for skilled readers. The "direct access" theory claims that the semantic representation of a written word can be accessed directly via the visual representation of the word without reference to the word's phonology.

The past several decades have witnessed a pendulum-like swing between the strong phonological view (phonological recoding) and the strong visual view (direct access) of access to semantics. The conflicting findings in the literature regarding the role of phonology in reading English make one wonder why there has been such difficulty in establishing the facts. Apart from the limitations of behavioral research in reading that will be discussed later on, one main reason is possibly due to the confounding nature of the English language itself. Most of the research in this area has been conducted in English and the stimuli used in many of these English studies have been homophones. Skeptics of the viability of the reading models based on those studies may argue that even though orthography itself has no direct impact on access to semantics it may come into play in reading tasks where it interacts with appropriate phonological information (Zhou & Marslen-Wilson, 1999). For instance, with the exception of very few examples such as *ate* vs. *eight,* English homophones (e.g., *rose* and *rows)* not only share identical phonological attributes but substantial orthographic characteristics. By this account, it seems logical to claim that the confounding between word form and pronunciation in alphabetic languages may hamper the examination of phonological roles in reading alphabetic materials.

The frustration inherent in attempting to control the orthographic confounding in English has gradually directed researchers' attention to languages with non-alphabetic orthographies. The Chinese language, for instance, is a logographic language where each character has its own symbol. In stark contrast with the transparency between orthography and phonology in alphabetic languages, there is an arbitrary relationship between orthography and phonology in Chinese characters. To illustrate, although two Chinese characters may be homophones insofar as they share the identical segmental and tonal attributes, they can have distinct orthographic forms (填 /tian1/-“add”; 天 /tian1/-“sky”). This arbitrary relationship in Chinese provides an excellent opportunity to examine the ‘pure’ contribution of phonological information to reading comprehension (Tan & Perfetti, 1998).

In point of fact, several lines of behavioral research concerning the nature of Chinese reading parallel studies using English language which implicate the involvement of phonology in reading. For example, employing forward priming and backward masking techniques, a good number of Chinese studies have shown that there is phonological activation during Chinese reading (Perfetti & Tan, 1998; Perfetti & Zhang, 1991, 1995; Tan, Hoosain, & Peng, 1995; Tan, Hoosain, & Siok, 1996). Critically, these experiments further provide evidence that phonological processing appears to occur in the early stage of Chinese character recognition. Most impressive perhaps is the study by Tan, Hoosan, and Peng (1995). Using the backward masking paradigm, the authors exposed the character target for a brief time, which was immediately followed by a character mask that was replaced by a pattern mask of 1.0 s. The character mask was visually, phonologically, or semantically either similar or dissimilar to the character target. The duration time of the character target and mask presentation varied across two experiments in this study. In Experiment 1, the target and mask were exposed for 50ms and 30ms, respectively. In Experiment 2, the exposure durations for the target and mask were, respectively, extended to 60ms and 40ms. The results in this study demonstrate that the exposure durations of the target and mask were shorter when the phonological effect was first found than when the semantic effect was first observed. Given the assumption that an early-occurring event may mediate a later-occurring event during the same reading task, the earlier activation of phonology in contrast to semantic activation was taken as convincing evidence for the proposition that phonology constrains access to semantic meaning.

In a similar vein, another line of research using lexical decision and meaning judgment tasks presents further evidence that phonological processing takes place in an automatic fashion during written Chinese character recognition (Chen, Flores d'Arcais, & Cheung, 1995; Leck, Weekes, & Chen, 1995; Perfetti & Zhang, 1995). For example, Perfetti and Zhang (1995) carried out two judgment tasks in which Chinese characters were visually presented. In the synonym judgment task, they asked subjects to decide whether two successive characters had the same meaning. In the second homophone judgment task, they required subjects to decide whether two successively presented characters had the same pronunciation. In their study, some succeeding characters had the same pronunciation as, but had a different meaning from, the preceding characters. It was found that phonological activation was present in both tasks. This evidence was further construed as the evidence for the "automaticity" of phonological processing.

However, this conclusion about the influence of phonology in Chinese reading has not gone unchallenged. In fact, some of the most striking phonological effects documented in the literature (Perfetti & Tan, 1998; Tan & Perfetti, 1997) were not replicated in some other behavioral studies even though the same stimuli, procedures and approaches were employed (Chen, 1998; Zhou & Marslen-Wilson, 1999). Setting this problem to one side, it is noteworthy that behavioral approaches are critically sensitive to the 'end-state characteristics of a process' and mainly associated with particular response requirements. In this regard, behavioral approaches may not be able to reveal the locus of early components in reading and distinguish distinct cognitive stages of a process. More importantly, language comprehension consists of very rapid and complex processes that may overlap in both spatial and time domains. In line with this view, while phonology is, no doubt, involved in reading, its potential role in constraining the semantic activation of a written word remains unclear. For the investigation of phonology's role in accessing semantic meaning during reading that may take place at an earlier stage of lexical processing, an online measure that can capture detailed language processing as it unfolds across time becomes of great value.

1.4 Event-related brain potentials in language research

Given the inherent limitations in behavioral approaches, capturing continuous online traces of language processing has been one of the most daunting problems in language research. However, the emergence of event-related brain potentials (ERPs) has made the future prospects in this respect appear very promising. The central position of this thesis is that event-related brain potentials can be taken as "an on-line and continuous measure, one that taps into the state of affairs in the lexical processing system while it operates in real time" (Zwitserlood, 1989, p. 29).

ERPs, as a noninvasive, objective, online electrophysiological measure of human brain functions, are time locked to an external event (e.g., a word or a picture) of interest and have sufficient temporal resolution to reveal the temporal dynamics of the neural activity underlying linguistic processes. Of crucial relevance is that ERPs provide direct real time insights into covert neural correlates of specific stages of information processing which are not contingent upon overt behaviors (Brandeis & Lehmann, 1986). In addition, although ERPs do not indicate the source location of neural activity since they are the summed brain activity of synchronized post-synaptic potentials, the scalp distribution of ERPs reflects the locations and orientation of the neurons within the ensemble (Kutas, 1997; Kutas & Federmeier, 2000). According to these features, analysis of language event-related brain potentials is regarded as "an attractive means of complementing the behavioral analysis of information processing" (Rugg, 1987, p.126) and poses effective constraints on the viability of language processing models derived from behavioral data.

Simply put, when a pair of electrodes are placed on the scalp surface and attached to an amplifier, a pattern of variation in voltage over time can be obtained from the output of the amplifier. This variation pattern is referred to as the 'electroencephalogram' (EEG) (Coles & Rugg, 1995). Since Berger (1929) first successfully recorded the scalp EEG, considerable effort has been devoted to the understanding of the neural origin of this scalp-recorded phenomenon. According to the present knowledge, two types of neuroelectric events in nerve cells contribute to the scalp EEG: inhibitory and excitatory synaptic potentials and action potentials. At the microscopic level, currents in relation to individual action potentials are generally detected more strongly than synaptic currents.

However, the magnitude of an action potential in the spatial domain is quite small because the action potential is an intracellular current that cannot easily be observed at a distance (closed field). In addition, in most cases there is little synchronous activation across neighboring axons that conduct action potentials. In this regard, at the macroscopic level, the net current change produced by multiple action potentials even across a large population of neurons is fairly small (Lewine & Orrison, 1995).

In contrast, synaptic potentials are primarily recorded as extracellular currents. The pattern of current flow of synaptic potentials is very sophisticated at the microscopic level but can be simply seen as if it were produced by current flow between two poles at the macroscopic level. This dipole approximation results in an electric potential distribution that can be recorded at a distance. What needs to be kept in mind, however, is that even with this dipole configuration synaptic activity generated by some neuronal populations may not be recorded at a distance. Only does the neuron population in which the individual neurons are synchronously active and which have a certain geometric configuration (e.g., the alignment of neurons is in a parallel orientation) yield potentials of sufficient magnitude that can be recorded at the scalp. At present, it has been strongly suggested that the EEG recorded at the scalp is principally a summation of graded post-synaptic potentials produced by the inhibitory and excitatory pyramidal cells within the cortex. The name of those cells comes from the peculiar pyramidal shape of their cell body. Pyramidal cells in the cortical layers are large and account for about 70 percent of the cortical neurons (a sizeable population). Another important feature is that all pyramidal cells in the laminar structure of the cortex and even their dendrites are organized in a parallel fashion that creates a large open bipolar field (Coles & Rugg, 1995; Knight, 1990; Kutas & Van Petten, 1994; Lewine & Orrison, 1995).

The ERP is derived from the EEG. It can be seen as a series of transient voltage perturbations of the spontaneous EEG, which are time locked to a definable event of interest such as the onset of a word and are specifically associated with the brain's response to the event (Kutas & Van Petten, 1988). Simply put, when the subject is presented with certain events of interest at the same time when we record the EEG, an epoch of the EEG that is time-locked to the language target can be defined. To illustrate, the epoch of the EEG analyzed in the three experiments of this thesis begins 100 ms before the onset of the target word and ends 1000 ms after the target onset. The raw EEG data normally contain both ERP responses (signal) to the language stimulus that are associated with language related information processes and spontaneous electrical brain activity (noise) that is not associated with word processing mechanisms. It is noteworthy that ERP signals are small in amplitude (e.g., 5-10 μV) compared to the raw EEG data that varies from 10-200 μV. In order to achieve reliable ERP signals, they must be extracted from the background raw EEG data by the means of signal averaging. This technique includes the recording of ERPs to repeated presentations of the 'same kind' of stimuli. The assumption behind this is that EEG activity that is not involved in stimulus-evoked processes is not synchronous with the timing of stimulus onset and thus cancels out by averaging. Conversely, the event-related ERP signal will summate and emerge from the background EEG (Kutas & Van Petten, 1994)

The relationship between neural processes in the brain and scalp recorded ERP waveforms is not yet completely understood. Nevertheless, it is evident that neural activity can be reliably recorded on the scalp in relation to the stimulus (Allison, Wood & McCarthy, 1986). ERPs have been widely used in identifying the neural correlates of specific stages of information processes in cognition, particularly with respect to word recognition and associated processes (Rugg, 1987).

Traditionally, a peak or trough of the ERP waveform that is linked or associated with a specific process such as semantic comprehension is defined as a "component" (Donchin, Ritter, & McCallum, 1978). It has been found that the particular component that is evoked relies predominantly on the nature of the stimulus of interest and the nature of the task demand that is applied in the study.

The earlier components of the ERP are usually considered to be those with latencies of less than 100ms after the stimulus onset. They can be elicited by a given stimulus and their amplitude and latency are determined primarily by stimulus parameters (e.g., intensity, modality and rate of stimulus presentation) (Knight, 1990). As such, the earlier evoked potentials are referred to as "exogenous", sensory or stimulus-bound components (Kutas & Van Petten, 1994). Examples of exogenous components are the brainstem auditory evoked response (BAEP) and primary somatosensory evoked potentials (SEP). Given the fact that those components are insensitive to a subject's state of alertness or attentiveness, they have been widely used for clinical diagnostic (e.g., multiple sclerosis disorder) and prognostic (e.g., coma) needs to measure neural activity in sensory pathways in neurological disorders (for a review see Wang, Young, & Connolly, 2003).

Within the cognitive domain, however, the more informative ERP components are those referred to as endogenous components. In contrast with exogenous components, endogenous components generally occur relatively late - hundreds of milliseconds after or before an evoking event. The triggering event may be a stimulus, a response, a voluntary movement, or a cognitive operation (Kutas & Van Petten, 1994). In contrast with the sensitivity of exogenous components to the physical stimulus parameters, endogenous components are primarily sensitive to nature of task demands, attention, decision-making, expectancies, or performance strategies. Put differently, endogenous ERP components are not strictly “evoked” by the presentation of the stimulus but are related to a range of processes constrained by the psychological demands of the event (Donchin, Ritter, & McCallum, 1978; Kutas & Van Petten, 1994). Given that endogenous potentials occur in relation to a certain cognitive event, they are also, in practice, referred to as event-related brain potentials (ERPs) (Knight, 1990). Examples of endogenous ERP components include the P300 which is well known for reflecting attention and memory processes (Polich & KoK, 1995), the N400 which is primarily seen as an index of semantic analysis (Kutas & Hillyard, 1980a,b), the Phonological Mismatch Negativity (PMN) which is related to phonological processing in the auditory modality (Connolly & Phillips, 1994) and the N270 which is associated with form-based processes in reading (Bentin, Mouchetant-Rostaing, Giard, et al., 1999; Connolly, Phillips, & Forbes, 1995).

To date, a large body of literature has demonstrated the usefulness of employing certain endogenous ERP components in language research and has laid the groundwork for the hypothesis that discrete lexical processes can be dissociated and indexed via ERPs (for a review see, Kutas & Van Petten, 1994). For the purposes of this thesis, given the crucial interest in the cognitive mechanisms engaged during language processing, ERP components associated with linguistic function will be introduced and the underlying mechanisms will be discussed.

1.4.1 The N400 component and lexical processing

A considerable amount of ERP literature has supported the practical and unique value of an ERP component called N400 in language research. The N400, a negative ERP component peaking around 400ms after the stimulus onset, was first described by Kutas and Hillyard (1980a) in response to sentence terminal words that did not match the semantic context of the sentence.

In an attempt to use ERP measures to index the function of sentential contextual constraints in subsequent word recognition, Kutas and Hillyard (1980a) conducted three experiments with different types of deviant sentence completions: moderate and strong semantic deviations, and physical deviations. Examples of moderate and strong semantic deviations are, respectively, a sentence ending with a relatively unexpected but semantically appropriate word like "He shaved off his mustache and *eyebrows*" and the sentence terminated with a semantically anomalous word like "He shaved off his mustache and *city*". The physical deviations made use of expected sentence ending words, which were, however, presented in upper case, an example of which is "She put on her high heel *SHOES*". In each experiment, 160 different seven-word sentences were visually presented on the screen with each word presented for 700 ms. Subjects were asked to silently read the sentences and told that they would be required to complete a questionnaire after the recording session. Three fourths of the sentences ended with expected and meaningful words (e.g., "The pizza is too hot to eat") and the remaining one-quarter of the sentences in each experiment were terminated by a deviant word according to the three experimental manipulations described above. The results demonstrated that two types of semantically deviant sentence-ending words were characterized by a clear negative brain wave with a centro-parietal distribution (N400). Different from semantic mismatch conditions, physically deviant targets were not followed by a N400 but a robust positive component (P300). What is more interesting is that the strong semantically deviant target word (e.g., "city" in the example above) was characterized by a much larger N400 component than the moderate semantically deviant target word (e.g., "eyebrow"). According to these results, Kutas and Hillyard argued that the N400 is not a manifestation of the analysis of physical attributes, but instead reflects aspects of cognition specific to linguistic processes.

A substantial body of later ERP literature reinforces the viability of this view. With the issue of whether anomalies within non-linguistic contexts can also be characterized by the N400 component, Besson and Marcar (1987) carried out an experiment in which four experimental conditions were manipulated: semantically congruity and incongruity in visually presented sentences, appropriate and inappropriate notes in well known French melodies, increasing (congruous) and decreasing size (incongruous) of geometric figures, and the ascending pitch (congruous) and descending (incongruous) pitch in music notes of the musical scale. Consistent with the first N400 study by Kutas and Hillyard (1980a), one quarter of the stimulus sequences ended incongruously (Besson & Marcar, 1987). The findings demonstrated that only semantically anomalous sentence-ending words were followed by a clear N400 component. The other three types of incongruous stimuli were characterized by a positive-going deflection (P300). Compatible with these findings, other literature reported that speech sounds played backward and orthographically illegal non-words in reading also failed to elicit a reliable N400 component (Holcomb & Neville, 1990; Rugg & Barrett, 1987; Smith & Halgren, 1987). Overall, these findings provide convincing evidence for the argument that the N400 is not sensitive to physical attributes of a language stimulus. Rather, it reflects some aspects of linguistic processes.

The N400 is indeed an extremely robust component in language research. Many ERP studies in reading have demonstrated that sentence-based semantic N400 effects can be obtained in a variety of languages. Semantically anomalous sentence-ending words in Chinese (e.g., Wang, 1997), French (e.g., Besson & Macar, 1987), Spanish (e.g., Kutas, 1985), Finnish (Helenius, Salmelin, Service, & Connolly, 1999) and American Sign language (e.g., Kutas, Neville, & Holcomb, 1987) are all reliably followed by the N400 component. Also, the N400 has been reliably elicited by semantic incongruity in both visual and auditory modalities in language research (Kutas & Van Petten, 1988).

The N400 is characterized as a ubiquitous marker of lexical processing. At issue is which linguistic processes are reflected by the N400. There has been general agreement that the default amplitude of the N400 is high and experimental manipulations, such as highly constrained semantic cues, decrease it (Van Petten, Kutas, Kluender, et al., 1991). Moreover, it is also argued that the amplitude of the N400 reflects the ease of which the target word is integrated into its linguistic (e.g., sentence) context. Integration generally refers to the process of introducing a target word into a higher-order meaning representation of the entire sentence or discourse (Brown & Hagoort, 1993). It can be construed as the process of matching the semantic specifications of an activated lexical item with a representation built up with respect to the semantic specification of the preceding context. The better the match between these two domains, the easier the target word is integrated into the entire context. The evidence supporting this integration view on the N400 effect comes from several lines of ERP research.

The first line of convincing evidence arises from a substantial body of ERP literature that examines the relationship between word integration and the extent of contextual constraints. For example, some early ERP studies have demonstrated that the amplitude of the N400 varies according to the target word position in the sentence. Kutas and Hillyard (1983) used prose passages with semantic anomalies in reading comprehension. The semantically incongruous words were placed either at intermediate positions (e.g., Other well-known reptiles are snakes, lizards, *eyeballs* and alligators.) or at the ends of the sentences (e.g., Turtles are smarter than most reptiles but not as smart as mammals such as dogs or *socks*). The findings of this study demonstrate that the N400 is responsive to both intermediate and terminal semantic anomalies, which replicates and extends the findings in Kutas and Hillyard's study (1980a). They undoubtedly provide convincing evidence in favor of the argument that words are integrated with the preceding context in a step-by-step fashion instead of being loaded until the end of a sentence (Carpenter & Daneman, 1981; Kutas & Van Petten, 1988). Furthermore, the striking finding in this study is that the amplitude of the N400 in response to a semantically coherent content words decreases as a function of increasing word position in sentence (i.e., N400 amplitude is negatively correlated with serial word position within a sentence). This systematic decline in the N400 amplitude is construed as compelling evidence for the integration processes manifested by the N400. It is reasoned that semantic context constraints accumulate with the increasing number of words in the sentence. In this regard, words occurring late in the sentence benefit from the more constrained contextual semantic cues established by the preceding words than those in the earlier positions of the sentence. The electrophysiological manifestation of this sentence position effect is that sentence-ending words evoked a larger N400 than sentence onset words.

It is interesting to note that the above N400 effect fails to emerge when sentences are embedded within coherent discourse (Van Petten, 1993). In an attempt to accommodate these two contradictory findings with one explanation, Van Petten (1993) reasoned that due to the coherence amongst embedded sentences in a discourse, the contextual cues even in the beginning of a sentence have sufficient contextual constraint to attenuate the N400 amplitude even to expected words that occur early in the sentence.

In fact, the amplitude of the N400 has been intentionally studied as an inverse function of semantic integration at both local (sentence) and global (discourse) levels. For example, St. George, and his colleagues (1994) presented subjects with paragraphs that made sense only if a title that summarized the global subject was given. All the paragraphs presented without titles looked like a series of relatively incoherent sentences each of which was nonetheless semantically appropriate and acceptable. Given that the semantic cues provided by the title would dramatically facilitate the semantic integration of such paragraphs (Bransford & Johnson, 1972), the authors predicted that inferences drawn with respect to the title presentation would attenuate the amplitude of the N400 if the amplitude of the N400 were sensitive to semantic integration at the global level. As predicted, the results demonstrated that N400 amplitudes were larger to the paragraphs without titles than to those with titles. In line with this argument, Van Berkum, Hagoort, and Brown (1999) conducted two experiments to examine the relationship between an incoming word to the preceding local semantic context and the wider global discourse. What they found was that the target word that was semantically anomalous to the wider discourse (e.g., "Jane told the brother that he was exceptionally *slow*" in a discourse context in which he was in fact very fast) was characterized by a large N400 compared with the discourse-congruous word (e.g., "fast"). This finding, in conjunction with the aforementioned studies, provides convincing evidence that the amplitude of the N400 reflects the extent to which the incoming word is integrated into its preceding contextual cues at both local and global levels.

Another line of research complements the view that the amplitude of the N400 is not restricted to the violations of semantic knowledge. Kutas and Hillyard (1984) manipulated experimental conditions in terms of sentence constraints and "cloze probability" of the sentence ending words (targets). "Cloze probability" refers to the proportion of individuals using that particular word as the most likely chosen word to complete a sentence fragment. In this experiment, sentence-ending words were divided into three kinds according to the degree of their cloze probability (high, medium, and low). Meanwhile, sentences used were classified into three levels according to contextual constraints (high, medium, and low). Highly contextually constrained sentences included those that resulted in very predictable endings (e.g., "He mailed the letter without a *stamp*") while low constraints sentences were those whose contextual cues could not lead to such strong expectations (e.g., "There was nothing wrong with the *car*"). The data showed that the highly probable words in highly constrained sentences were characterized by a robust positive component (P300) while low probability words were followed by a significant negative deflection (N400). Although Cloze probability and sentential constraint to some degree interact with each other, the N400 was argued to be more sensitive to Cloze probability than to the extent of sentential constraint. The supporting evidence is that there was no difference in the N400 amplitude for words with similar Cloze probability (e.g., the highly probable word "car" in the low constraints sentence "there was nothing wrong with the *car*" and the medium probable word "jobs" in the medium constraints sentence "Too many men are out of *jobs*"). What's more, the amplitude of the N400 decreased with the increasing Cloze probability of words in medium constraints sentences.

Based on these findings, it was argued that the N400 can be seen as a sensitive indicator of the semantic relationship between a word and the context in which it appears. In particular, the amplitude of the N400 is modulated as an inverse function of semantic expectation (Cloze probability). The sensitivity of the N400 amplitude to the Cloze probability is in turn taken as compelling evidence for the prevalent view that the N400 reflects contextual integration. The higher the Cloze probability, the easier the integration of the eliciting item into context-based information, and the smaller the N400 amplitude.

However, the view that the amplitude of the N400 reflects the extent to which the eliciting item is integrated into its preceding contextual information does not always succeed in interpreting the findings obtained in N400 language studies. For example, Kutas, Lindamood, and Hillyard (1984) presented their subjects with three types of sentences with respect to the semantic violations of the sentence-ending words. The first type involved sentences ending with the highest Cloze probability word (e.g., "The pizza is too hot to eat"), the second type used sentences ending with semantically incongruous words (e.g., "The pizza is too hot to cry"), and the third consisted of sentences terminated with semantically incongruous words but whose meaning was related to the highest Cloze probability word (e.g., "The pizza is too hot to drink"). According to the integration hypothesis, the words in the latter two conditions should have been difficult to integrate into their respective contexts and thus, should have evoked the same-sized N400. In contrast to the prediction, however, the N400 in response to the word "drink" proved to be very small compared with that to the word "cry".

The viability of the integration view has also been called into question in some other ERP literature. Using the sentence verification paradigm, Fischler and collaborators (Fischler, Bloom, Childers, et al., 1983; Fischler, Childers, Achariyapaopan, et al., 1985; Fischler, Bloom, Childers, et al., 1984) conducted a series of studies to investigate the influence of the sentential meaning constraints on the amplitudes of the N400 by asking subjects to judge the accuracy of simple statement sentences such as "A robin is a bird". Surprisingly, the amplitudes of the N400 show no relationship with the truth or falsity of the propositions. Rather, they varied with the semantically associative relationship between the subject and object of the sentence. Simply put, a robust N400 was found in response to the target word (object of the sentence) that was semantically unassociated with the subject of the sentence even if the statement was true (e.g., "A robin is not a *vehicle*"). By this account, the N400 is not restricted by the propositional content of the statement sentence but by the strength of semantic associations between the codes corresponding to the subject and object of the sentence in the mental lexicon.

The N400 has also been used to investigate semantic association in lexical decision, naming, and letter search studies (Bentin, Kutas, & Hillyard, 1995; Brown & Hagoort, 1993; McCarthy & Nobre, 1993; Kutas & Hillyard, 1989). For example, it has been reported repeatedly that semantic associations between individual words in lists modulated the amplitude of the N400. A word (e.g. “eat”) that is semantically related to its prior word (“drink”) is readily characterized by a smaller N400 than when it is preceded by a semantically unrelated word (“car”). Based on those findings, some (Brown & Hagoort, 1993; Kutas & Hillyard, 1989; Kutas & Van Petten, 1994) argue that N400 amplitude measures are indeed constrained by semantic associations given the view that contextual constraint effects can be accounted for with respect to lexical associations between the constituent items (e.g., Bradley & Forster, 1987).

In line with this argument, the term ‘semantic priming effect’ was introduced in an attempt to delineate the mental processes manifested in the N400. In lexical decision research, priming generally refers to the facilitation phenomenon that it takes less time to decide that a string of letters or phonemes is a word if it is preceded by a related item than if the preceding item is unrelated (e.g., Meyer, Schvaneveldt & Ruddy, 1974). Taking a semantic priming effect as an example, the reaction time to decide whether the item “car” is a real word is shorter when it follows the word “vehicle” or “bus” compared to when it follows the word “desk”. A common explanation for this priming effect is that it is attributable to the automatic spread of activation among nodes in the mental lexicon. It is important to note that the relationship between nodes in the mental lexicon is not only restricted to associations with respect to semantic knowledge but also involves identity or associative information in orthographic and/or phonological forms (Kutas &

Hillyard, 1989). Connecting low-resistance links exist between nodes in the mental lexicon that share the same or associative attributes (e.g., semantic relatedness or similar phonological forms). When a word in the network is processed, the corresponding node in the mental lexicon is consequently activated. This activation is not limited merely to the node corresponding to the processed word. Rather, it spreads along the links in the network to other associative nodes so that all related nodes leave the resting state and become closer to the threshold of recognition. As a consequence, it takes less effort and time to activate an associative "restive" node that facilitates semantic activation of the corresponding word when the processing of the word is subsequently required. In addition, the activation level of nodes corresponding to unrelated words does not vary with this priming effect (Brown & Hagoort, 1993).

Indeed, the "spread of activation" semantic priming theory can easily account for the modulations of the N400 amplitude by sentence constraints, Cloze probability, related semantic anomalies (e.g., "the pizza is too hot to drink"), and semantic relatedness. Based on the 'spread of activation' priming theory, for instance, the sentence fragment "the pizza is too hot to _" provides clear and clean semantic information which provokes the cognitive system to pre-activate nodes in the mental lexicon bearing such semantic features, such as the node corresponding to the word "eat". Meanwhile, the pre-activity in the "eat" corresponding node automatically spreads along certain links to any other nodes that share associative features of the pre-activated node in either semantic or any other feature domains. The consequence of this spread of the pre-activation leads to a small N400 amplitude in response to the target word "drink" which is semantically associated with the highest Cloze probability word "eat" although it is a semantically anomalous word for the sentence. Meanwhile, according to the "spread of activation" priming theory, the different magnitudes of the N400 amplitudes in response to different Cloze probable sentence-ending words can also be also understood in terms of the various strengths of the pre-activation determined by its preceding sentence fragment.

A number of N400 studies further demonstrate that N400 amplitude can also be modulated by the number of times a target word is presented (word repetition effect) and daily use word frequency. These two N400 effects, together with the aforementioned studies, accord the spread of activation theory a prominent position in current models of the nature of the N400. As found in the large body of word repetition ERP studies (e.g., Van Petten, Kutas, Kluender, et al., 1991; Rugg, 1987), the N400 amplitude in response to a repeated word is reduced compared to the N400 following the word's first presentation. It is reasoned in terms of the spread of activation theory that the node in the mental lexicon corresponding to the repeated word has some "leftover" activity due to its previous activation that may remain for a period of time. As a result, the "leftover" activation facilitates the repeated node to pass over the threshold of recognition that in turn facilitates semantic activation.

Another line of evidence supporting this view comes from the word frequency N400 effect (Van Petten & Kutas, 1990). It has been reported that N400 amplitudes are reduced to target words with higher word frequency values. It has been a well-known behavioral finding that subjects usually require either less time or less information to process common words than they do to rare words (Rubinstein, Garfield & Millikan, 1970; Solomon & Howes, 1951). The word frequency N400 effect is compatible with this behavioral frequency phenomenon. According to the "spread of activation" priming theory, variations in N400 amplitudes to words with varying word frequencies can be interpreted as reflecting the various general "resting" levels of the nodes corresponding to words with different word frequencies. The "resting" level of the common node is closer to the activation threshold of recognition (due to frequent practice) than that of the rare node.

In order to accommodate all of these N400 effects within one underlying mechanism, a compelling argument has been put forward with respect to the processing nature of the N400. It states that the amplitude of the N400 can be seen as an inverse function of the ease of semantic analysis and is sensitive to the strength and precision of the spreading activation along the links of the network in the mental lexicon (e.g., Praamstra & Stegeman, 1993; Kutas & Federmeier, 2000; Kutas & Hillyard, 1988). If this argument is accurate, any associative features including semantic and other linguistic attributes (e.g., phonology and orthography) that can influence the pre-activation or threshold level of the corresponding nodes should facilitate semantic activation of the target word, the consequence of which would reduce N400 amplitudes.

Indeed, it has been argued that the N400 has also found in response to priming manipulations in terms of not only semantic but also orthographic or phonological characteristics of the stimuli. Kutas and Van Petten (1988) have suggested that Sanquist and colleagues (1980) first revealed the sensitivity of the N400 to phonological cues despite their primary concern being the relationship between P300 and language related cognitive processes. Their study consisted of a series of experiments in which pairs of words were presented and judged as "same" or "different" with respect to orthographic, phonemic, or semantic characteristics. Kutas and Van Petten (1988) have interpreted these data as suggesting that a large negative component (N400) followed the "different" words in the semantic judgment experiment. More interestingly, a discernible N400 was obtained in response to "different" words in the phonemic judgment task that did not put any explicit emphasis on semantic analysis of words.

In a similar vein, Kutas and Van Petten (1988) reinterpreted another data set (Polich, McCarthy, Wang, & Donchin, 1983) as demonstrating the sensitivity of the N400 to the engagement of both phonological and orthographic processes in a visual word pair. In this study, subjects were presented with pairs of printed words and required to judge whether the second word (target) was orthographically similar to, or rhymed with, the first member (prime) of the word pair. The word pair stimuli included the word pairs which not only rhymed but also looked alike, those which rhymed but did not look alike, those which looked alike but did not rhyme, and those which neither rhymed nor looked alike. Kutas and Van Petten (1988) inspected the ERP waveforms and suggested that there was a robust N400 component whose amplitude varied with the phonological and/or orthographical modulations. To illustrate, when rhyme matching was the task demand, the N400 following the rhymed word pair appeared much smaller than that following the non-rhyming word pair. Meanwhile, orthographically dissimilar words elicited larger N400s than orthographically similar words. When the task demand was focused on orthographic similarity matching, orthographically dissimilar words were likewise characterized by a large N400 compared with orthographically similar ones. What is interesting is that under this task demand the N400 became insensitive to phonological variations (rhyme vs. non-rhyme). Based on theses findings, it was argued that the elicitation of the N400 was not exclusively restricted to semantic processes.

Rather, the amplitude of the N400 can be also influenced by the engagement of phonological and orthographic information. From another perspective, the disparity of these ERP data that orthographic effects of the N400 occurred in both task manipulations while phonological effects of the N400 existed only in the explicit phonology-oriented task provides compelling evidence in favor of the view that phonological and orthographic representations can be activated independently. The additional implication of this "disparity" effect also extends the tentative argument that phonological processing may not be a prerequisite process during reading (Kutas & Van Petten, 1988).

In an attempt to explicitly replicate the above phonological effects of the N400 and provide more insights into the nature of this effect, some ERP studies intentionally recorded N400s in response to phonological and/or orthographic manipulations. For instance, in the first experiment of Praamstra and Stegeman's study (1993), aurally presented non-rhymed word pairs were contrasted with physically incongruous ones, i.e., being spoken by different voices (one male, the other female). The results indicate that physically dissimilarity (different voices) did not elicit an N400 but rather a late positive component (P300). However, ERPs to non-rhyming words were more negative than those evoked by rhyming words in the 250-500 ms time range where the classical N400 semantic effect is obtained. It was argued that this rhyme priming ERP effect is sensitive to phonological cues instead of acoustic variables (i.e., voice difference). It was further concluded that this priming effect reflects the sensitivity of the N400 to phonological manipulation.

As well, a large body of ERP literature using words and non-words in both visual and auditory related word pair tasks has obtained similar phonological effects (Kramer &

Donchin, 1987; Praamstra & Stegman, 1993; Rugg & Barrett, 1987). A typical finding is that phonological incongruity (non-rhyming) evoked a significant N400-like deflection regardless of whether the evoking stimuli were real words or phonotactically illegal non-words. Given the phonotactically illegal features of non-words used which would inhibit the activation of lexical processes, it is questionable whether the phonological effects of the observed N400-like component reflect the activation of the phonological representations that bridge between physical input and corresponding lexical nodes in the mental lexicon. Taking a closer examination of these studies, however, most of them place task demands explicitly on the phonological (e.g., rhyme matching) or orthographic (i.e., "look like") levels, which, as Praamstra and Stegeman (1993) argued, may gear processing strategy away from natural word processes. In addition, as Connolly and colleagues (1995) have suggested, this rhyme-sensitive negative deflection may involve a pre-N400 response (PMN) that could have been related to those phonotactical violations. In particular, the emphasis of those task demands on phonological analysis renders possible that the rhyme priming ERP effect described above was the manifestation of this earlier component rather than the N400.

Using the same word and non-word stimuli as used in the above studies, Praamstra and Stegeman (1993) asked subjects to carefully listen to word-word and word-nonword pairs, and to "respond to the second member of each pair, and decide fast and accurately whether it was a word or not"(p.82). The authors reasoned that phonological processing is crucial in spoken word processing which captures phonological information that is related to the physical-acoustic instantiation of a word and establishes contact between the acoustic signal and corresponding lexical nodes in the mental lexicon. In line with this argument, the N400 should be sensitive to phonological manipulations. In this case, a clear phonological effect of the N400 should be elicited by nonrhyming words even if phonological cues are not explicitly placed in the task demand. Indeed, the findings were compatible with the expectation. It was found that the amplitude of the N400 elicited by nonrhyming words was larger than that by rhyming words. Importantly, such an N400 effect was not observed in word-nonword pairs. This result seems not in accordance with what was found in the above explicitly rhyming judgment task but consistent with the results of some other investigations that used orthographically illegal nonwords in lexical decision tasks (Holcomb & Neville, 1990; Smith & Halgren, 1987; Van Petten, Kutas, Kluender, et al., 1991). At the same time, these results extend the scarce knowledge of the phonological effects of the N400 that are assumed to occur only when phonological codes are task relevant. Taking all the findings mentioned above into consideration, the phonological effects of the N400 more likely reflect the consequence of phonological processes instead of phonological processing per se (Praamstra & Stegeman, 1993).

Overall, the N400 component has been used as a marker of cognitive processes associated with semantic analysis. Given the sensitivity of the N400 to both semantic analysis and the consequence of form-based (phonology and orthography) encoding, the N400 can be seen as of great value for the investigation of the influence of semantic information and phonological cues on word recognition during reading and listening.

1.4.2 Phonological Mismatch Negativity (PMN) and spoken language research

The Phonological Mismatch Negativity (PMN) as an auditory-specific ERP component was first identified by Connolly, Stewart, and Phillips (1990) and labeled as the PMN later (Connolly & Phillips, 1994). The discovery of this component was a consequence of research efforts to elucidate the modality-specific cognitive processes supposedly characterized by the time course of the N400 component. Before discussing the PMN it is important to review the literature that preceded and in some ways contributed to its discovery.

An increasing body of research has demonstrated that the semantic priming effect on the N400 that has been recorded primarily in reading research can be reliably observed in speech comprehension tasks. The common finding is that a spoken word that is semantically related to its preceding contextual information (either sentence or the prime word in a pair) evokes a smaller N400 than a semantically unrelated spoken word (e.g., Bentin, Kutas, & Hillyard, 1993; Connolly, Stewart, & Phillips, 1990; Osterhout & Holcomb, 1993).

The semantic priming effect on the auditory N400 was first recorded in an auditory replication of Kutas and Hillyard's classic visual N400 study (1980). In this study, McCallum, Farmer, and Pocock (1984) presented sentences spoken by a male speaker with highly expected sentence-ending words, semantically anomalous ending words, and highly expected ending words but unexpectedly spoken by a female voice. Consistent with the classical N400 findings in the visual study, the anomalous sentence-endings were characterized by a large negative component (N400) and congruous endings were followed by a positive component (P300) in the 400 ms time range. What is different, however, is that the auditory N400 appeared to have a more anterior distribution than the visual N400. In addition, this study and several subsequent investigations found that the semantic priming effect on the auditory N400 appeared to have a different time course from that of the visual N400 (Anderson & Holcomb, 1995; McCallum, Farmer, & Pocock, 1984; Holcomb & Anderson, 1993; Holcomb & Neville, 1990, 1991).

For example, in an attempt to examine the timing effects on the N400 component of a single word semantic context (prime), Holcomb and Neville (1990) employed a lexical decision paradigm and systematically studied contextual priming effects to auditory and visual stimuli. In this study, there were two blocks of trials: stimulus pairs in one block were spoken by a male voice and those in the other were written. The paired stimuli used in both blocks included word/word pairs that were semantically related, word/word pairs that were semantically unrelated, and word/nonword pairs. In addition, the target was presented 1150ms after the prime. Compatible with the findings of most of the N400 literature, the results of this study demonstrated a robust semantic priming effect in both modalities: the amplitude of the N400 was smaller when a target word was semantically related to its prime than when the target word was preceded by a semantically unrelated word. What is noteworthy, however, is that the time-course and scalp distribution of the N400 effect showed modality specificity. First, the auditory N400 effect looked more bilaterally symmetrical while the visual N400 effect had a right hemisphere distribution. Given that the distribution of the N400 can be symmetrical even in some reading tasks, this finding was not taken as convincing evidence for the modality-specific effect. What was particularly interesting was the finding that the divergent point of the negative ERP component to semantically incongruous target words (in comparison to the ERP to congruous target words) in the auditory modality appeared to start earlier (between 200ms and 290ms after the target word onset) than that in the visual modality (between 300ms and 360ms). The authors' explanations of these effects appear counter initiative in some respects. The authors proposed that, in contrast with the written word that can be fully available at word onset, sounds of the spoken word unfold across time. They noted that the durations of all spoken words in this study were not shorter than 270ms. By this account, an earlier divergence effect would have been expected in the visual instead of the auditory modality. However, the authors suggest that contextual priming started its impact on the spoken word processing prior to the occurrence of the final sounds of a spoken word. Several theories of spoken word recognition have claimed that the locus of preceding semantic contextual priming influences on word recognition occurs only at the level of higher-order semantic processes instead of at the form-based (e.g., phonological) levels (Holcomb & Neville, 1990; McClelland & Elman, 1986). For instance, according to the Trace model (e.g., McClelland & Elman, 1986), the analysis of ongoing acoustic input is purely form-driven and bottom up without any contextual influence. Top-down contextual information starts to exert its influence only after the activation of form-based representations. Taking all the above arguments into account, semantic processes engaged during listening comprehension seem to begin prior to the arrival of all the acoustic information of a spoken word.

In line with this argument as well as the argument that the N400 is a sensitive index of semantic analysis, Holcomb and Neville (1990) take the above divergent moment between incongruous and congruous ERPs as the onset of the N400 component.

The earlier influence of contextual priming effect for auditory than visual words was further taken as evidence for the argument that different semantic mechanisms which are characterized by the N400 component serve for spoken and written word recognition.

The view that the early onset of the auditory N400 effect reflects the early engagement of semantic analysis has not gone unquestioned, however. In an attempt to extend the findings of the above study, Holcomb and Neville (1991) carried out two sentence comprehension experiments. Their study was aimed to further examine the time course of the contextual priming effect on the N400 by manipulating semantic congruity of sentence-ending words and the temporal interval between words in the sentence. The stimuli used in both experiments were composed of three types of sentences: those that ended with best completion words (contextually meaningful), those with related but anomalous words (contextually meaningless but related to the best completion), and those with unrelated anomalous words (contextually unrelated). The interval between words varied across two experiments: in the first experiment, the sentence stimuli were presented at the natural speech rate. In contrast, the second experiment introduced a constant 750 ms word interval between words within spoken sentence stimuli. Consistent with the aforementioned N400 findings, the results in both experiments revealed a significant N400 semantic priming effect. However, in contrast with the time course pattern of the N400 effect in the Holcomb and Neville's (1990) study, the divergent moment of the negative-going ERPs elicited by unrelated anomalous sentence-ending words from the ERPs in response to the best completion words in connected natural speech occurred even earlier (about 50 ms after the word onset). Given the fact that natural speech bears more non-semantic cues such as prosodic and co-articulation than word pairs, Holcomb and Neville attributed this early priming event to the facilitation effect of prosodic and word co-articulation cues on the best completion word processing. More interestingly, the results in this study further demonstrated that the onset of the N400 effect varied with the word interval duration. With 750 ms word interval, the onset of the N400 effect appeared late in the time range of 220-300ms after the terminal word onset. It has been emphasized that although the word interval of 750ms may make sentences sound less than natural the relatively long word interval did not inhibit sentence processing. Given that the manipulations in both experiments of this study are semantic in nature and that the same stimuli were used, the word interval rate effect suggests that the varied onsets of the contextual priming effect are not related to semantic-sensitive processes.

As well, the data obtained in some other studies cast further doubt on the argument that the different onsets of the auditory and visual N400 effects manifested modality-specific semantic processes. Based on the manipulation of the time intervals between the onsets of the prime and the target word of a pair (SOA), the time course of the single word contextual priming effect was further studied and compared in two between- and within-modality lexical decision studies (Anderson and Holcomb, 1995; Holcomb & Anderson, 1994). The stimuli in both studies used pairs of words that were either semantically related or unrelated and word-nonword pairs. The SOA of the prime and target in both studies was manipulated across three levels: simultaneous presentation (0ms), 200ms, and 800ms. What is different between these two studies is that in the between modality study (Holcomb & Anderson, 1994) there were two experiments: in the first, the prime word was visually displayed and the target item was spoken (visual/auditory); and, in the second, the prime was spoken and the target was written (auditory/visual). In contrast, in the within-modality study (Anderson & Holcomb, 1995), there were two experiments; one in which both the prime and the target were written (visual/visual) and the other in which both the prime and the target were spoken (auditory/auditory). It was predicted that if there were different mechanisms underlying semantic priming for spoken and written words, between- and within-modality priming should produce different patterns of N400 effects. In particular, if the SOA was set to zero, the priming effect should only be found in the auditory/visual and auditory/auditory experiments because the semantic priming in the auditory modality was assumed to be able to start even on the basis of partial acoustic information of the prime word.

Surprisingly, these two studies produced findings contradictory to what were predicted. In contrast with the significant priming effect that was observed at the 0 SOA in the visual/auditory and visual/visual experiments, there was no discernible N400 priming effect at the 0 SOA in the auditory/visual and auditory/auditory experiments. Unexpectedly, the N400 priming effect was only reliably obtained at the 800 ms SOA in these two latter auditory/visual and auditory/auditory experiments. More interestingly, at the 0 SOA condition in the auditory/auditory experiment, although no N400 priming effect was found, there was a discernible ERP contextual priming effect in the 200-300-ms time window after the word onset which had a reliable frontal distribution.

From these findings, it can be suggested that the early onset of the auditory ERP priming effect does not correspond to semantic processes characterized by the N400. Instead the effect seems to be associated with some other speech-related processes. By this account, it is logical to hypothesize that the early onset of the auditory priming effect may be the manifestation of a negative ERP component that overlaps with the N400 component and is sensitive to auditory-specific processes. It has become a common view that sound-based phonological encoding plays a critical role in access to semantic meaning of a spoken word (Frauenfelder & Tyler, 1987; Klatt, 1989; Pisoni & Luce, 1987). Given this account as well as the time course of the early contextual priming (right before semantic priming reflected by the N400), Connolly and colleagues (Connolly, Stewart, & Phillips, 1990; Connolly, Phillips, Stewart, et al., 1992) proposed that the relatively early onset of the aural contextual priming effect indicated the occurrence of a separate ERP component which was sensitive to phonological mechanisms. One of the theories underlying this proposal is that preceding semantic contextual information can pre-activate the corresponding word nodes in the mental lexicon (e.g., "The piano was out of" pre-activates the node of the word "tune"). In turn, the pre-activated node(s) spread the activation along connected links to the nodes that share similar or identical attributes in both semantic and the form-based (e.g., phonological) domains, resulting in a phonological priming effect. Indeed, Connolly and Phillips (1994) have provided explicit evidence in support of this proposal.

In order to successfully tease apart different cognitive processes reflected by the overlapping negative component and the N400 component, Connolly and Phillips (1994) directly manipulated both phonologically-sensitive and semantically-related features of sentence-ending words in a listening comprehension task. In this study, participants were presented with four types of medium to high constrained sentences that terminated with highest Cloze probability words (e.g., 'the piano was out of *tune*') (condition 1), with semantically anomalous words that had the same initial sound as the highest Cloze probability words (e.g., 'The gambler had a streak of bad *lu*ggage') (condition 2), with semantically appropriate but low Cloze words that did not have the same initial sound as the highest Cloze probability words (e.g., 'Don caught the ball with his *glove*'('hand' being the high Cloze probability ending for this sentence)) (condition 3), and with semantically anomalous words that did not have the same initial sound as the highest Cloze probability words (e.g., 'The dog chased our cat up the *queen*'('tree' being the high Cloze word in this example)) (condition 4). As expected, the initial sound mismatched sentence endings in the last two conditions were followed by an early negative component with a frontal distribution that peaked in the 250 and 300 ms time range. The N400 was seen in the last condition but not in condition 3. Interestingly, semantically anomalous sentence endings in the second condition failed to evoke such an early component but rather were characterized by a robust classical N400 component with a posterior maximum which peaked in the 400-500ms time range – significantly later than the N400 seen in the last condition. Given the sensitivity of this early component to phonological variations, Connolly and Phillips named it Phonological Mismatch Negativity (PMN).

It is interesting to note that examination of the ERP priming effect at the 0 SOA condition in the auditory/auditory experiment of Anderson and Holcomb (1995) suggests that the early priming effect may be the manifestation of the PMN without the presence of the N400. One of the possible accounts for this phenomenon is that phonological processing can begin on partial acoustic information of the prime word that can prime the target word processing in time. The reason why there is no discernible auditory N400 priming at the 0 SOA condition may be that acoustic information of a spoken word needs some time to spread out and the semantic analysis of a spoken word may not be able to proceed until adequate phonological information is available.

A number of ERP studies have provided further converging evidence that the PMN component is a reliable index of phonological processing during listening (e.g., Connolly, Byrne, & Dywan, 1995; Connolly, Service, D'Arcy, et al., 2001; Hagoort & Brown, 2000; van den Brink, Brown, & Hagoort, 2001). For example, Hagoort and Brown (2000) asked subjects to listen to sentences that terminated either with a semantically congruent or incongruent word that did not have the same initial sound as the congruent completion either. A robust biphasic negative deflection was observed in response to the semantically anomalous endings. The authors believe that two distinct components comprised this biphasic deflection: one component peaked around 250ms (the analogue to the PMN component), and was more sensitive to "the form-based activation of a lexical candidate" (p.1528) and the other was the N400.

1.4.3 Earlier occurring negative components and reading research

Several early-occurring ERP components (e.g., the N200 and the N270) associated with different aspects of reading processes have been recognized and described in the literature. It has been argued that these earlier occurring components reflect distinct aspects of language processing. In particular, they are believed to be associated with early parts of the language processing sequence which take place prior to the semantic analysis associated with the N400. The supporting evidence for this assumption is derived from their difference not only regarding their latency range but also in terms of their topographical distribution pattern. For example, the N400 has been reported to be more apparent over centro-parietal recording sites (Kutas & Van Petten, 1988) while the N270 has been repeatedly reported to have a fronto-temporal distribution (Connolly, Phillips, & Forbes, 1995; Forbes, 1993, 1998; Newman, 2000). In this review section, two earlier-occurring components, the N200 and the N270, will be fully delineated and discussed.

1.4.3.1 The N200/N170 and orthographic specificity

The N200 specific to orthographic stimuli was demonstrated in a study by Nobre, Allison, and McCarthy (1994). ERPs were recorded using intracranial implanted electrodes. Several tasks including a sentence reading task and a semantic priming task were manipulated in this study. In the sentence-reading task, subjects were required to judge whether the sentence-ending targets were semantically appropriate. The stimuli used in this task were composed of three types of sentences: those that ended with best completion words (contextually meaningful); those that ended with such stimuli as semantically incongruous words, pseudo-words and illegal nonwords; and, those that ended with non-orthographic visual stimuli such as human faces. In the semantic priming task, the authors presented a prime (e.g., chair) followed by a target (e.g., table) and asked subjects to identify the target. The stimuli used in this task consisted of orthographic and non-orthographic stimuli. In both tasks, ERPs to orthographic stimuli were compared with those to non-orthographic stimuli.

As expected, orthographically legal letter strings and incongruous real words elicited a robust N400. Prior to the N400, a negative component peaking about 200 ms after the stimulus onset (N200) was elicited by both orthographic and non-orthographic stimuli. Of interest, however, is that the intracranial distribution of the N200 varied according to the physical feature of presented stimuli: the N200 in response to orthographic stimuli dominated in distinct brain areas from that to non-orthographic complex visual targets (e.g., faces). Although both orthographic and non-orthographic stimuli induced activity in the posterior fusiform gyrus, the regions corresponding to these two types of stimuli did not appear to overlap within a subject (Allison, McCarthy, Nobre, et al., 1994). Meanwhile, it is worth noting that orthographic stimuli induced more activity in the left hemisphere while face stimuli produced either symmetrical response patterns or more activity in the right hemisphere.

Given these features of the N200, the authors concluded that the N200 reflects processes specific to letter string analysis. In particular, given that the intracranial distribution of the N200 did not distinguish between pronounceable and non-pronounceable nonwords, it was further argued that the N200 reflects visual form analysis independent of phonological analysis.

In an attempt to disentangle different aspects of visual word processes, Bentin, Mouchetant-Rostaing, Giard, et al. (1999) carried out a word identification study using ERP. The working hypothesis provided was that visual word recognition is a complex process that can be simplified into three distinct levels of cognitive operations which proceed in a sequential fashion: an orthographic level at which visually presented input is decoded and integrated to represent orthographic patterns; a lexical level at which the phonological representation of the input is transformed from orthographic patterns; and, a semantic level as the final processing point at which semantic meaning of the word input is successfully activated. It is argued that each level stated above can be intensified or constrained according to specific task manipulations.

With this hypothesis in mind, the authors manipulated a series of oddball tasks in which the distinction between designated target and non-target stimuli was manipulated to intensify the processing level of interest including orthographical, phonological, and semantic, in the meanwhile controlling other aspects of the processing. The ERPs to non-target stimuli were of particular relevance in this study.

In the orthographic processing task, the authors manipulated targets twice as large as the non-targets and asked participants to silently count the number of large targets. Stimuli used in this task consisted of orthographic stimuli (letter strings) and non-orthographic stimuli (alphanumeric symbols and ASCII forms). The authors hypothesized that orthographic processing should be automatically induced by orthographic stimuli in contrast to non-orthographic stimuli. In order to screen out the possibility that any ERP difference between orthographic and non-orthographic stimuli was in fact derived from phonological and/or semantic activation, the ERPs to words, pseudo-words, and unpronounceable strings of consonants were compared. The results of this experiment demonstrated that both orthographic and non-orthographic non-target stimuli elicited a negative component peaking about 170 ms that had a temporal-occipital distribution. It was further shown that the scalp distribution of the N170 distinguished between orthographic and non-orthographic stimuli: The N170 in response to orthographic stimuli was largest in amplitude over the left hemisphere and the N170 to non-orthographic stimuli was largest over the right hemisphere. Given its latency, scalp distribution and sensitivity to orthographic analysis, the N170 seems equivalent to the intracranial recorded N200 reported in the Nobre, et al.'s (1994) study.

Taking the findings from both of the above studies into consideration, it is evident that the N170/ N200 reflects early occurring processes specific to orthographic information. The sensitivity of the N170/ N200 to orthographic information further suggests that distinct brain regions in the visual system correspond selectively to discrete visual information (Bentin, Mouchetant-Rostaing, Giard, et al., 1999). In addition, given the insensitivity of the N170/N200 to phonological cues, aspects of reading processes manifested by this component do not seem to have involved phonological encoding.

1.4.3.2 The N270 and its associated reading processes

Forbes (1993) conducted an ERP study in an attempt to elucidate the influence of phonological information on written word processing. The stimuli in this study were made up of visually presented highly constrained English sentences the endings of which were either semantically expected words (Congruence) (e.g., 'The dove is a sign of **peace**'.) or semantically anomalous words. The anomalous ending words were either phonologically congruous but semantically incongruous to the sentence context (Foil) (e.g., 'The class baked a cake and everyone had a **peace** (piece)'.) or semantically inappropriate homophones that did not share phonological information with highly expected sentence ending words (Incongruence) (e.g., 'Many people eat peanut butter and jelly **wave** (waive)'.). ERPs were recorded to the sentence-ending words and subjects were asked to read sentences silently.

The findings of this study revealed that an N400 was reliably induced by the Incongruence condition but not by the Congruence and Foil conditions. A notable finding is that a negative component peaking about 270 ms after the terminal word onset (N270) was elicited in both Incongruent and Foil conditions but not in the Congruent condition. The author took the absence of the N400 in the Foil condition as converging evidence that there was phonological involvement at least when phonological cues were clearly provided during reading comprehension. Most importantly, the author proposed that the N270 reflected certain aspects of reading processing that were associated with orthographic expectation. Put differently, the occurrence of the N270 in both Incongruent and Foil conditions reflected the deviation from the orthographic representation of the highly expected sentence ending words.

In an attempt to clarify the issue of whether phonological involvement in English reading is indeed a default process or its presence is task-dependent, Forbes (1998) carried out a series of ERP studies to control the possibility that phonological activation is in fact derived from the advantage of participants' awareness of a phonologically based reading strategy. In a semantic judgment task, the author used the same stimuli as well as the same experimental manipulation as those employed in Forbe's (1993). In this task, subjects were required to pay close attention to the sentence ending word and judge via a button press whether the ending word was semantically correct. The author surmised that a semantic judgment task would make subjects pay close attention to semantic meaning instead of phonological cues. In this regard, subjects would presumably apply orthographic codes to constrain semantic interpretation of the terminal word.

Correspondingly, the Foil sentence-ending targets were expected to elicit a clear N400 component reflecting their semantic incongruity.

As expected, whereas an N400 was not observed in the Foil condition in Forbes (1993), a clear N400 was clearly invoked by the Foil targets in this study. The author therefore interpreted this finding as meaning that phonological facilitation was overshadowed by the semantic incongruity. Also, the results of this study showed that the Foil target words did not elicit a N270 although the N270 was steadily replicated in the Incongruent condition. This result was unexpected given the position that the N270 reflects orthographic expectancy, according to which there should have been an N270 in response to orthographic deviations of the Foil sentence ending targets.

In view of the conflicting results, the author proposed that the absence of the N270 in the Foil condition might have resulted from the similarity in the orthographic dimension between Foil targets and highly expected sentence endings (e.g., bear/bare). To test this hypothesis, Forbes (1997) carried out a second experiment in which the magnitude of orthographic similarity between Incongruent targets and highly expected sentence endings was manipulated. The stimuli used in this experiment included four types of sentences: those that ended with highly probable words (Congruence); those with semantically anomalous words which were orthographically similar to highly probable sentence endings (orthographic similarity (OS) condition); those with semantically anomalous words which were orthographically dissimilar to highly probable sentence endings (orthographic dissimilarity (OD) condition); and, those with homophone foils (FOIL condition). It was found that the N270 in response to both the Foil targets and the OS targets was reduced in amplitude in contrast with that to the OD targets. The author took this N270 finding as supporting evidence that the N270 is sensitive to orthographic expectations.

Unfortunately, the author did not further delineate and discuss the nature of the N270 elicited in the Foil condition which might have shed better light on the relationship between the phonological processing and the underlying mechanisms manifested by the N270. In an attempt to clarify the relationship between the processes manifested by the N270 and phonological activation in reading, Newman (2000) conducted an ERP study in which subjects were required to silently read sentences and decide whether they made sense. The stimuli used were visually presented English sentences ending with: words that were orthographically congruent (OC), phonologically congruent (PC), and semantically congruent (SC) to the high cloze probability endings (e.g., Larry writes with his left *hand*); pseudo-words that were orthographically incongruent (OI), PC, and SC to the high cloze probability endings (e.g., The ship disappeared into the thick *phog* (fog).); words that were OI, phonologically incongruent (PI), and semantically incongruent (SI) to the high cloze probability endings (e.g., The pizza was too hot to *sing* (eat).); and, pseudo-words that were OI, PI, and SI to the high cloze probability endings (e.g., The teacher wrote the problem on the *heet* (board).). The results of this study demonstrated that a clear N400 followed the SI targets while an N270 was observed in response to the OI targets. According to these results, Newman (2000) proposed that the N270 and the N400 reflect distinct aspects of language processing. Consistent with Forbes' proposition, she further suggested that the N270 is associated with orthographic expectation.

The results of Newman's (2000) study further revealed that the N270 in response to the OIPISI words, although showing no difference from that to the OIPISI pseudo-words, appeared significantly larger than the N270 to the OIPCSC pseudo-words. Facing this finding, a tempting question addressed is whether the neural processes characterized by the N270 involve phonological transformation mechanisms? Given limited data provided in the literature, the issue as to whether the N270 is associated with phonological activation in visual word recognition remains an open question.

In summary, the literature indicates that ERPs can be used to investigate discrete stages of a process that are engaged during language comprehension. It remains controversial as to what role phonology may play in speech processing and whether phonological cues contribute to reading comprehension. More importantly, almost all of the ERP literature conducted as of this date has been primarily devoted to semantic, orthographic, and/or segmental phonological analysis. There have been no ERP studies that explicitly sought to examine the role of word prosody in lexical processing and the relationship between word prosody and other form-based information (e.g., segmental phonology) in both speech and reading. The series of experiments in this thesis utilized ERP measures to probe the functional locus and time course of word prosodic and sub-lexical phonological analyses during word recognition in both English and Chinese Mandarin speech as well as silent Chinese reading. The amplitude and latency of language related components (e.g., PMN, N270 and N400) were taken as dependent measures to elucidate the mechanisms underlying the processes of lexical prosody (lexical stress in English and lexical tone in Chinese), syllable-sized sub-lexical phonology, and segmental (onset and rime) phonology in language comprehension. In addition, the ERP findings obtained from the present research were also used to complement the performance data observed in behavioral studies and offer a better insight into the validity of some spoken and written word recognition models derived from behavioral studies.

Chapter Two:

Experiment 1: The influence of lexical stress and syllable-sized sublexical phonology in auditory multi-syllabic English word processing: An ERP investigation

2.1 Objectives:

In the first experiment of this thesis, I wish to shed light, primarily and empirically, on the neural correlates of word prosodic processing and sub-lexical phonological encoding during auditory English word recognition. The two specific issues that were focused on in this experiment are as follows. First, I wished to establish whether word onsets or primary stressed information of a spoken English word plays a predominant role in initiating lexical processing. Second, I wished to identify what processes are engaged during listening comprehension so that semantic representations can be successfully accessed. At the same time, the other relevant goal of the first experiment was to define the role that sub-lexical phonology plays in spoken word recognition. More specifically, the syllable unit of phonological information in multi-syllabic English words (such as the syllable-sized phonology /ka:/ in the word 'cartoon') was targeted and the issue of what effects syllable-sized sub-lexical phonology exerts on multi-syllabic word processing was addressed and studied in this experiment.

As ERPs provide empirical information about the strength, timing and neural basis of cognitive processes engaged during speech comprehension, they can address the aforementioned issues that have been impervious to behavioral investigations. Thus, those issues were approached in this experiment by investigating the sensitivity to certain experimental manipulations of the PMN and the N400, the two ERP components that are respectively sensitive to phonological processing and semantic analysis. To date, no neuroimaging studies have addressed the issues of the influence of word prosody and sub-lexical phonology and their interactive function on word recognition. This experiment then will provide the first examination of these issues using ERP. Using this measure, unique information regarding the influence of word prosody and sub-lexical phonological information in the underlying cognitive architecture of the language system will be obtained. The viability of the sequential, goodness-of-fit, and stress-guided hypotheses can be thereby tested in a manner impossible with purely behavioral approaches.

In this experiment, a sentence judgment task was used in which participants were required to attend to binaurally- presented English sentences and make a judgment whether the sentence heard made any sense. Five experimental conditions were employed that varied according to stress pattern, sublexical phonology, and/or semantic appropriateness of the sentence terminating word to the sentence context (see Table 1).

For the purposes of the present experiment, word prosody is operationally defined in terms of lexical stress division. The rationale for using this division is that a good number of reports demonstrate that English listeners are sensitive to subtle stress differences in speech, which may imply the functioning of lexical stress patterns on speech segmentation and word identification (e.g., Mattys, 2000; Vroomen & De Gelder, 1997; Vroomen, Tuomainen, & de Gelder, 1998). As stated in the introduction, the Metrical Stress Segmentation (MSS) model (Cutler & Norris, 1988) is based on liberal metrical prosody. The disadvantage of this liberal stress division, however, is that the dependence of metrical stress may result in a great number of erroneous detections of word boundaries (Mattys, 2000). For instance, on the basis of this metrical prosodic

Table 1

Experimental Design used in Experiment 1.

Stimulus Conditions	Sentence Example	Number of Trials
CONGRUENT	Nicole was killed in a major highway ***accident***	40
LIGHT/PO**LITE**	She told the lost tourist to turn right at the traffic	40
(Stress non-initial-Prime non-initial)	po**lite** (light)	
ONE/**WON**DER	Eight minus seven equals ***won***_der (one)_	40
(Stress initial – Prime initial)		
EYE/I**DEA**	The pirate wore a patch over his i**dea** (eye)	40
(Stress non-initial – Prime initial)		
INCONGRUENT	Janice went outside to get some fresh ***dic***_tionary (air)_	40

Note: 120 fillers were sentences ending with high Cloze probability words.

A total of 320 trials were binaurally presented in Experiment 1.

segmentation strategy, false mid-word segmentation is bound to take place in at least some bi-syllabic words (e.g., “migraine”), many tri-syllabic words (e.g., “generate”), and eventually all longer words (e.g., “cosmopolitan”). In contrast, lexical stress division gives the priority of word segmentation to primary stressed syllables (Cutler & Carter, 1987). Such restriction correspondingly reduces the false segmentation at the mid-word. The empirical evidence (Mattys, 2000) demonstrates that listeners can discriminate between primary and secondary stressed syllables that may guide spoken word recognition. In fact, another line of research indicates that not only adults but also infant listeners are able to exploit subtle stress differences to segment words from continuing speech (Mattys, Jusczyk, Luce, et al., 1999; Morgan, 1996; Vroomen & de Gelder, 1997). By this account, stress patterns of sentence terminal words in this experiment were defined according to the lexical stress division (instead of the metrical division): two types of sentence endings were used based on whether or not their initial syllables were primarily stressed.

2.2 Hypotheses:

In line with the sequential hypothesis underscoring the processing primacy of word onsets in spoken word recognition, it was expected that the lexical initiation of target words in the ONE/**WON**DER (in which the initial-stress sentence ending word was semantically incongruent, but its initial stressed syllable was phonologically congruent, to the sentence context.) and EYE/I**DEA** (in which the noninitial-stress sentence ending was semantically inappropriate, but its initial unstressed syllable was phonologically congruent, to the sentence context) conditions would be facilitated given that the initial information of target words in these two conditions can be highly expected by its preceding sentential contextual information. Correspondingly, it was further expected that the PMN component in response to target words in these two conditions should appear smaller than the PMN in response to target words in the LIGHT/PO**LITE** (in which the noninitial-stress sentence ending was semantically incongruent, but its noninitial stressed syllable was phonologically congruent, to the sentence context) and INCONGRUENT (in which the initial-stress sentence ending was semantically and phonologically inappropriate to the sentence context) conditions. In contrast, the PMN to ONE/**WON**DER and EYE/**IDEA** ending words should show no difference from the PMN elicited in the CONGRUENT (in which the initial-stress sentence ending was the highest Cloze probable word to the sentence). Otherwise, the refining of the Cohort model should be taken into consideration.

If the “goodness of fit” hypothesis holds a grain of truth, it was expected that lexical processing of a target word should be facilitated as long as a syllable of the word is primed or expected by its preceding sentential context. The consequence of such a facilitation effect is that the PMN component in response to the CONGRUENT, ONE/**WON**DER, EYE/**IDEA** and LIGHT/PO**LITE** target words should appear smaller than the PMN elicited in the INCONGRUENT condition. Most critically, the position of the “goodness of fit” hypothesis would give rise to the expectation that the PMN component elicited in the three ONE/**WON**DER, EYE/**IDEA,** and LIGHT/PO**LITE** conditions should appear similar to each other in amplitude due to their equal weighting of sub-lexical syllabic priming according to their preceding contextual cues. Otherwise, this hypothesis needs to be justified in order to fit well in with word recognition model.

If the stress guide hypothesis reflects the truth of the nature of lexical processing, it was hypothesized that lexical processing primacy would not be given to word onset acoustic information but to that which is primarily stressed. Based on this stress guide hypothesis, it was expected that only those targets whose primary stressed information regardless of its word position is highly expected by their preceding contextual cues, can be eased during lexical initiation. That is, the phonological contextual priming effect varies with the stress status of the primed syllabic information instead of its word physical position. In line with this expectation, the PMN component elicited in the ONE/**WON**DER and LIGHT/PO**LITE** conditions was expected to appear smaller than the PMN in the EYE/I**DEA** condition. Meanwhile, given the deemphasizing of word onsets in the stress guide hypothesis, it was further proposed that the PMN elicited in the EYE/I**DEA** condition should show no difference from the PMN in the INCONGRUENT condition.

Furthermore, if the retroactive strategy is the true account of the processing of non-initial-stress words, it was expected that the onset and time course of the N400 corresponding to non-initial-stress targets in the EYE/I**DEA** and LIGHT/PO**LITE** conditions should be delayed and prolonged in contrast to those in response to initial-stress target words in the ONE/**WON**DER and Incongruent conditions. Otherwise, the viability of the retroactive strategy would be questioned and the nature of the processing of noninitial stress words would need to be further explored.

As for the effects of sub-lexical phonology on semantic constrains of the target English word: if sub-lexical phonology plays a facilitation role in semantic analysis of English words, the amplitude of the N400 should be reduced when the sub-lexical phonology of the target can be expected by its preceding sentential context; if sublexical phonology exerts inhibitory effects on whole word recognition, the corresponding N400 component should present a broad shape and the time course of the N400 should be prolonged due to the extra effort to compete with other candidates at the lexical level. An alternative possibility was also expected that the sub-lexical phonological effect only takes place at an earlier stage of phonological encoding and it may have no impact on the semantic representation of the whole word. In this regard, the amplitude and the time course of the N400 should only present its sensitivity to the semantic manipulation: the N400 elicited in the CONGRUENT condition appears much smaller than that in the other four conditions between of which should show no differentiation in amplitude and time course domains.

2.3 Method:

2.3.1 Participants

Fifteen right-handed first year undergraduate Psychology students (8 male and 7 female) were recruited and awarded course credit points for their participation in this study. Their mean age was 19.67 (SD = 0.82) years. In order to constrain potentially confounding variables, only individuals whose native language was English were accepted into the experiment. All of the participants had normal hearing and normal or corrected-to-normal vision. Furthermore, all participants were screened with a self-report measure for previous audiological, psychiatric, and/or neurological disorders. They were fully informed of the EEG recording and experimental task procedures and signed a consent form prior to the onset of the experiment.

2.3.2 Stimuli and experimental conditions

The 320 English sentences used in this experiment were recorded by a male native Canadian English speaker. Stimuli were digitized at 10 KHz (12 bit resolution) and timing marks (inaudible to participants) were placed at the onset of the sentence-ending word. To identify word onset, the waveform of the terminal portion of each sentence was viewed on a computer screen using successive windowing until the onset of the terminal word had been identified. Visually identified word onsets were further confirmed by repeated auditory presentation of the windowed area around word onset. The timing marks were used to initiate sampling of the electroencephalographic activity (EEG). The sentences varied in length from 5 to 13 words, most of which were composed of 6-9 words (For details, see Appendix A). Terminal word length was balanced across conditions.

The five experimental conditions used sentences (40 per condition) in which the sentence endings varied according to stress pattern, sub-lexical phonology, and/or semantic appropriateness (See Table 1 for the stimulus sample in each condition). In the CONGRUENT condition, the terminal word of each sentence was a word whose primary stress was on the initial syllable and which was the highest Cloze probability word of the sentence (e.g., "Three people were killed in a major highway ***accident***"). In the LIGHT/PO**LITE** (stress noninitial-prime noninitial) condition, the terminal word of each sentence was an unexpected word, which was semantically anomalous regarding the sentence context, whose stress sat on the second syllable, and whose stressed syllable was the (homophone of) highest cloze probability word for that sentence (e.g., "She told the lost tourist to turn right at the traffic ***polite*** (light)"). In the ONE/**WON**DER (stress initial-prime initial) condition, the terminal word of each sentence was a semantically anomalous word to the sentence context, whose stress was on the initial syllable and whose stressed (initial) syllable was the (homophone of) highest Cloze probability word for that sentence (e.g., "Eight minus seven equals ***wonder*** (one)"). In the EYE/**IDEA** (stress noninitial-prime initial) condition, the terminal word of each sentence was a semantically anomalous word regarding the sentence context, whose stress was on the first syllable and whose initial unstressed syllable was the (homophone of) highest cloze probability word for that sentence (e.g., "The pirate wore a patch over his i***dea*** (eye)"). In the **Incongruent** condition, the terminal word of each sentence was a semantically anomalous word to the sentence context whose stress was on the initial syllable (e.g., "The spider sat in its web awaiting a ***closet*** (fly)"). 120 fillers were sentences ending with the highest Cloze probability words (e.g., "The little girl has long brown **hair**"). The conditions were presented pseudo-randomly with the restriction that no condition was presented successively more than three times.

2.3.3 Procedure

Participants were seated comfortably in a sound attenuated room adjoining a room containing the recording equipment. Participants were instructed to attend to the sentences binaurally presented to them and to judge, as accurately and as quickly as possible, whether each sentence made any sense by pressing one of two response buttons. Given that the behavioral judgment demand was only aimed to draw participants' attention to aurally presented stimuli, the performance data were hereby not recorded and studied. Participants were also told to avoid blinking during sentence presentation. They were told that they might be asked some questions about the stimuli at the end of the experiment. The experimenter went through the study trials with participants to ensure that they fully understood the task demands of the experiment. Frequent breaks were given to reduce participant fatigue.

2.3.4 EEG recording

EEG was recorded from 17 sites using tin electrodes with linked ears as the reference and a mid-forehead ground. Recording locations were placed according to the standard international 10-20 system (Jasper, 1958). Midline recording sites (Fz, Cz, Pz) were used, along with lateral electrode pairs over frontal (F3, F4, F7, F8), central (C3, C4), parietal (P3, P4), temporal (T3, T4, T5, T6), and occipital (O1, O2) sites. The electro-oculogram (EOG), with electrodes placed above and below the right eye, recorded vertical eye movements (VEOG). Horizontal eye movements were recorded with electrodes placed over the outer canthi of the right and left eye. The electrode impedance was kept at or below 5 k Ω. Analogue EEG recordings were made with a half amplitude bandpass of 0.01Hz to 100 Hz and digitally sampled at 500Hz. Off-line filtering was set between 0.1Hz and 30Hz. The sample duration began 100ms before the stimulus onset and continued until 1000ms after stimulus onset. The stimulus onset was defined as the beginning of the sentence-ending word. The EEG data were baseline corrected. Trials contaminated by EOG greater than $\pm 75\mu v$ or any other artifacts (e.g., muscle activity and frequent α waves) were excluded from the analysis. The remaining EEG trials were then averaged by experimental conditions. The individual ERPs were also averaged together to create grand average waveforms for the various conditions.

2.3.5 Data analysis:

The statistical analyses of the ERP data were conducted with a repeated measures analysis of variance (ANOVA) and Greenhouse-Geisser corrections to the degrees of freedom were used when appropriate (Greenhouse & Geisser, 1959). The primary analyses of the PMN and N400 responses included the Condition (5 levels), Time (3 levels for PMN and 6 levels for N400), and Site (17 levels) factors. The five levels of the Condition factor were CONGRUENT, LIGHT/PO**LITE**, ONE/**WON**DER, EYE/**IDEA**, and INCONGRUENT conditions. The time factors were adjusted for each component analysis as follows: the PMN was examined with the 200-350ms window (three, 50ms intervals); and, the N400 was examined within the 350-650ms window (six, 50ms intervals). The 17 levels for the Site factor were Fz, F3, F4, F7, F8, Cz, C3, C4, Pz, P3, P4, T3, T4, T5, T6, O1, and O2. In addition, a secondary analysis was conducted to examine further the spatial distribution of components. In this analysis, Condition (5 levels) and Time (3 levels) remained as factors, but the recording sites were divided into two factors: Region (frontal, central, parietal, temporal, occipital) and Hemisphere (left, right). The frontal level combined F3 and F7 for a left hemisphere frontal value and F4 and F8 for the corresponding right hemisphere value. A similar approach was conducted with temporal sites combining T3 and T5 for the left and T4 and T6 for the right. The central level consisted of C3 and C4 values while the parietal level consisted of the P3 and P4 values. The occipital level is composed of O1 and O2 values. For the sake of brevity, results from the Condition/Time/Site analyses are presented first, with those from the Condition/Time/Region/Hemisphere analyses being presented only when relevant (e.g., if providing new information).

Significant main effects and interactions were submitted to further post hoc analyses using the Tukey Honestly Significant differences (HSD) test ($p<0.05$). In addition, differences in scalp distribution across conditions were submitted to normalization procedures to control for spurious interactions involving the Site factor (c.f., McCarthy & Wood, 1985). An alpha level of $p < .05$ was required for statistical significance.

2.4 Results

2.4.1 Waveforms

The grand average waveforms (Figure 1A) display clear differentiation amongst the five experimental conditions. The waveforms elicited in five conditions separated at about 200 ms post-stimulus. This difference was characterized by two sequential negative-going components (PMN and N400). The N400 is clearly seen in the EYE/**IDEA,** ONE/**WON**DER, and INCONGRUENT conditions with the largest being in the INCONGRUENT condition. Among these three conditions, the peak of the N400 in the ONE/**WON**DER condition appears delayed relative to that in the other two conditions. There is a discernible N400 in the LIGHT/PO**LITE** condition, which is nonetheless much smaller than that in the above three conditions. The characteristic positive-going waveform in the N400 time window is clearly observed in the CONGRUENT condition. The amplitude of N400 in the EYE/**IDEA** condition is much larger than that in the ONE/**WON**DER condition in the 350-450 ms time window. However, there is no N400 difference between the two conditions in the 450-650 ms time window (Figure 1C). Immediately before the N400, a negative component (PMN) can be observed in the

Figure Caption

Figure 1A. Grand average waveforms at 17 recording sites from 15 individuals for the CONGRUENT, LIGHT/PO**LIGHT**, ONE/**WON**DER, EYE/**IDEA**, and INCONGRUENT stimulus conditions presented in Experiment 1.

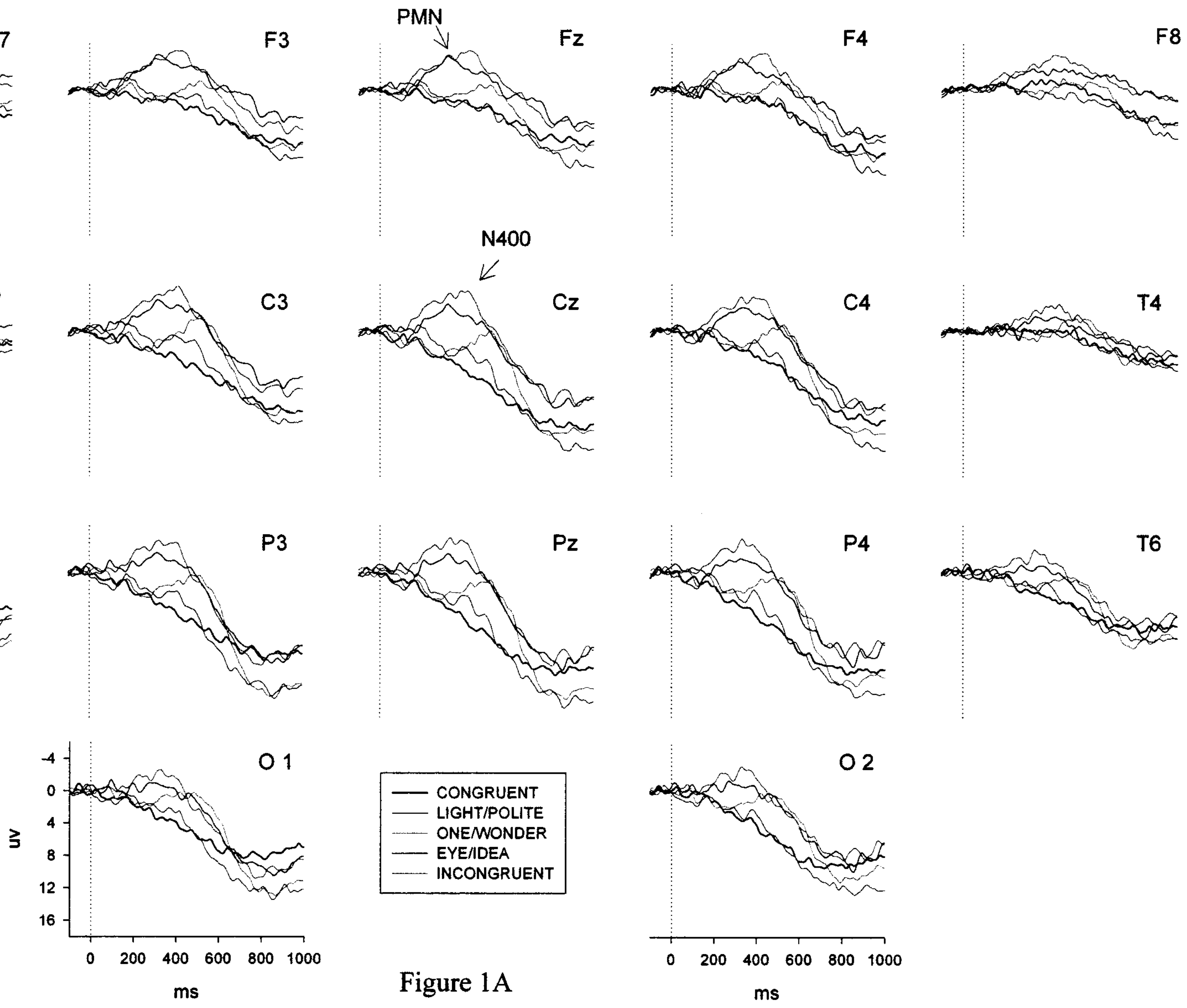
PMN
N400
F7
F3
Fz
F4
F8
T3
C3
Cz
C4
T4
T5
P3
Pz
P4
T6
O 1
O 2
uv
-4
0
4
8
12
16
0
200
400
600
800
1000
ms
CONGRUENT
LIGHT/POLITE
ONE/WONDER
EYE/IDEA
INCONGRUENT

Figure 1A

Figure Caption

Figure 1B. Grand average waveforms at Fz from 15 individuals for the CONGRUENT, LIGHT/PO**LIGHT**, ONE/**WON**DER, EYE/**IDEA**, and INCONGRUENT stimulus conditions presented in Experiment 1.

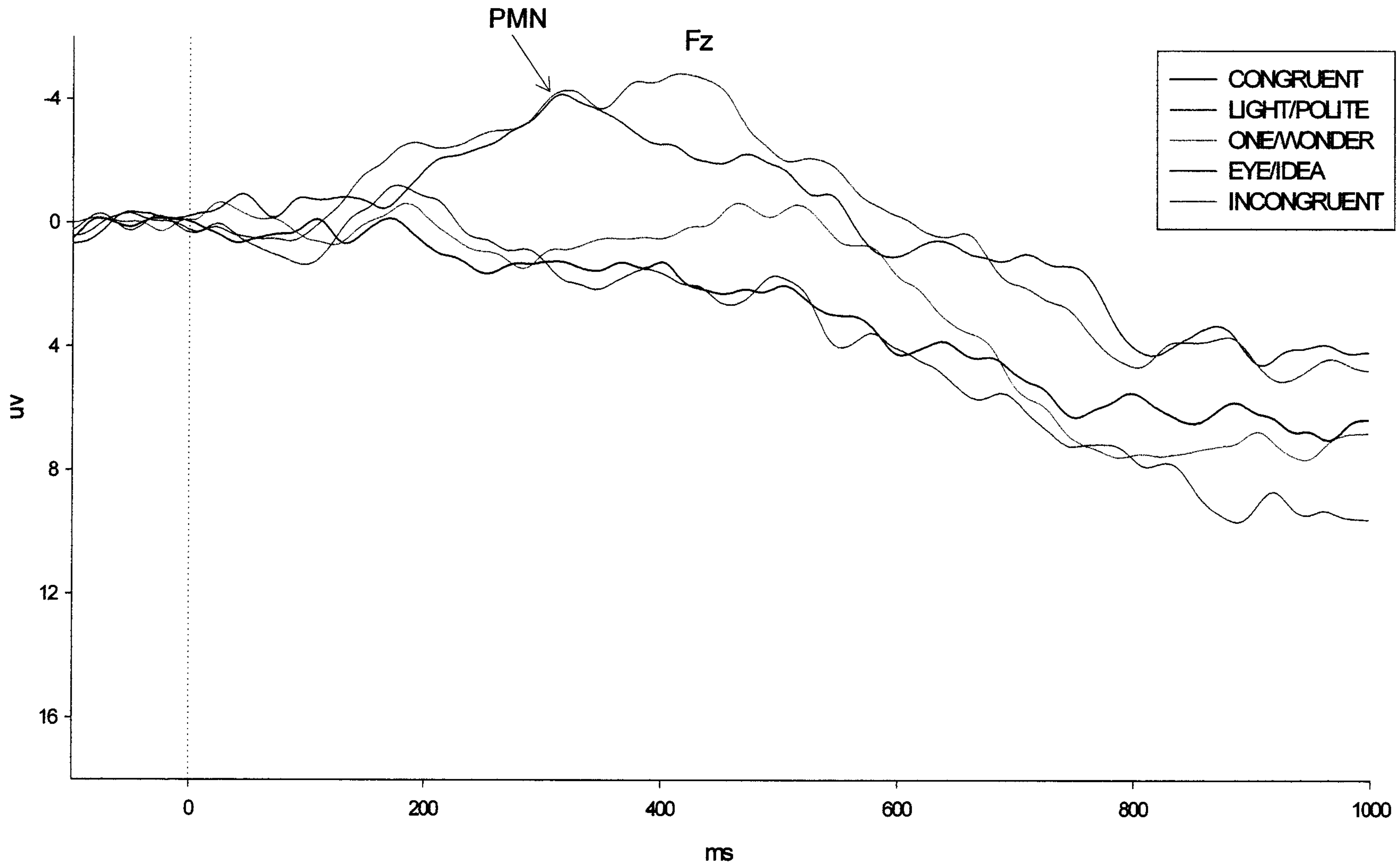
PMN
Fz
CONGRUENT
LIGHT/POLITE
ONE/WONDER
EYE/IDEA
INCONGRUENT
-4
0
4
8
12
16
uv
0
200
400
600
800
1000
ms

Figure 1B

Figure Caption

Figure 1C. Grand average waveforms at Cz from 15 individuals for the CONGRUENT, LIGHT/PO**LIGHT**, ONE/**WON**DER, EYE/**IDEA**, and INCONGRUENT stimulus conditions presented in Experiment 1.

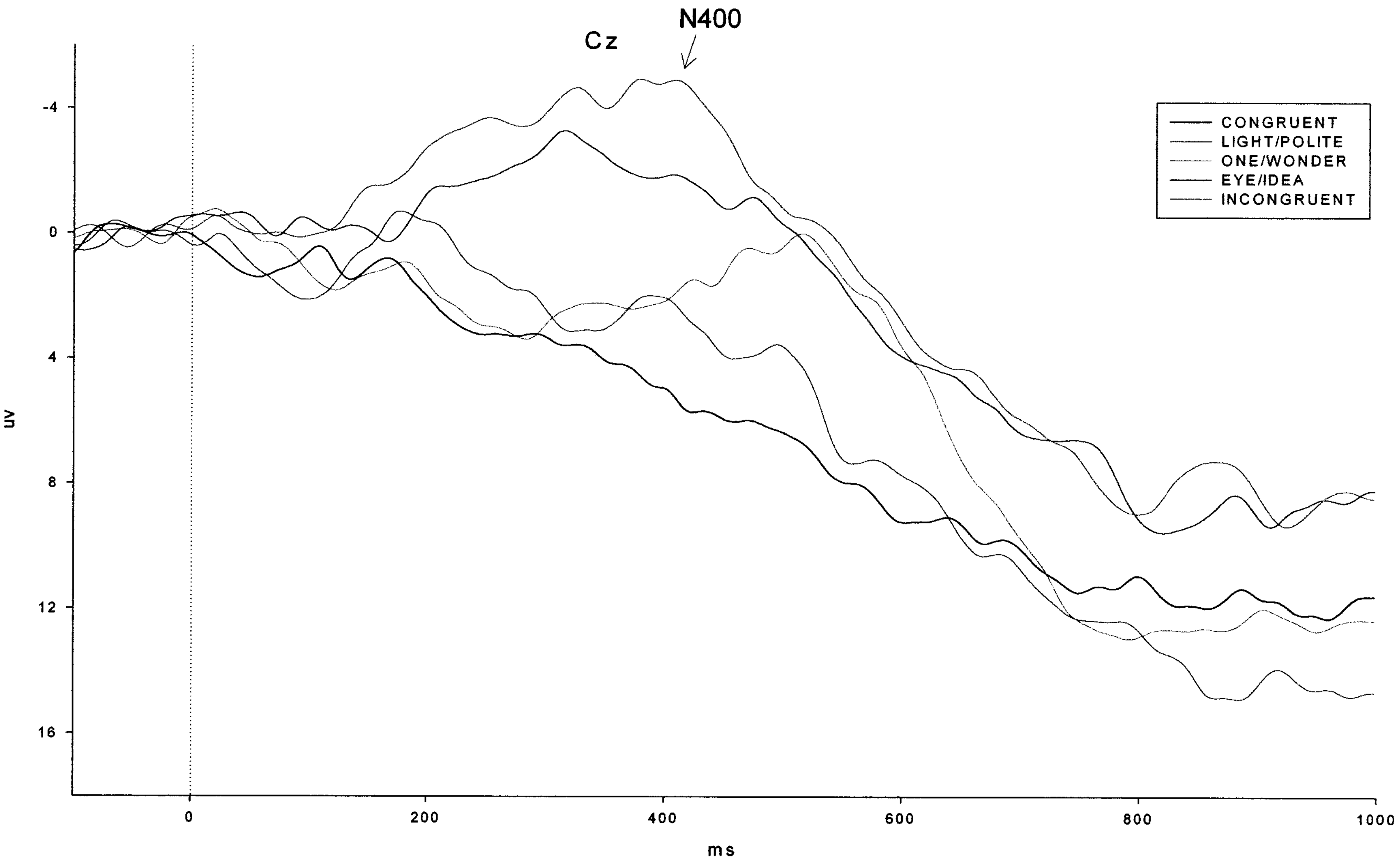
Cz
N400
CONGRUENT
LIGHT/POLITE
ONE/WONDER
EYE/IDEA
INCONGRUENT
-4
0
4
8
12
16
uv
0
200
400
600
800
1000
ms

Figure 1C

INCONGRUENT and EYE/**IDEA** conditions that appears similar in the two conditions over frontal recording sites (Figure 1B). In the corresponding time window (200-350ms after word onsets), a positive deflection is clearly observed in the LIGHT/PO**LITE,** ONE/**WON**DER and CONGRUENT conditions with little apparent difference between the three conditions, which is identifiable in each individual participant ERP waveforms. Figure 12 demonstrates the waveforms of a representative participant.

2.4.2 Statistical analyses

Phonological Mismatch Negativity (PMN)

Analysis of the PMN demonstrated its sensitivity to the stress status of the primed syllable, regardless of its physical position in the target word. There was a main effect of condition (F (4, 56)=30.57, p<0.001, ε=0.67) (See Appendix B Table B1) and subsequent post-hoc analyses indicated that PMN amplitudes in the INCONGRUENT condition (M = -2.72uv, SE=0.44) were larger than those in the EYE/**IDEA** (M = -1.60uv, SE=0.39), LIGHT/PO**LITE** (M = 0.89uv, SE= 0.37), ONE/**WON**DER (M = 1.21uv, SE= 0.52), and CONGRUENT (M = 1.77uv, SE=0.42) conditions (See Figure 2). Among the last four conditions, the mean amplitudes in the EYE/**IDEA** condition were statistically larger than those in the other three conditions among which no significant differences were found. There was a significant Condition x Time interaction (F (8, 112) = 7.02, p<0.001, ε=0.59), and subsequent analyses demonstrated that the mean amplitudes in the EYE/**IDEA** and INCONGRUENT conditions were significantly larger than those in the other three conditions in the 300-350ms time interval. In addition, differences in amplitude were

Figure Caption

Figure 2. Graph depicting the amplitude of the PMN response as a function of the stimulus condition in Experiment 1.

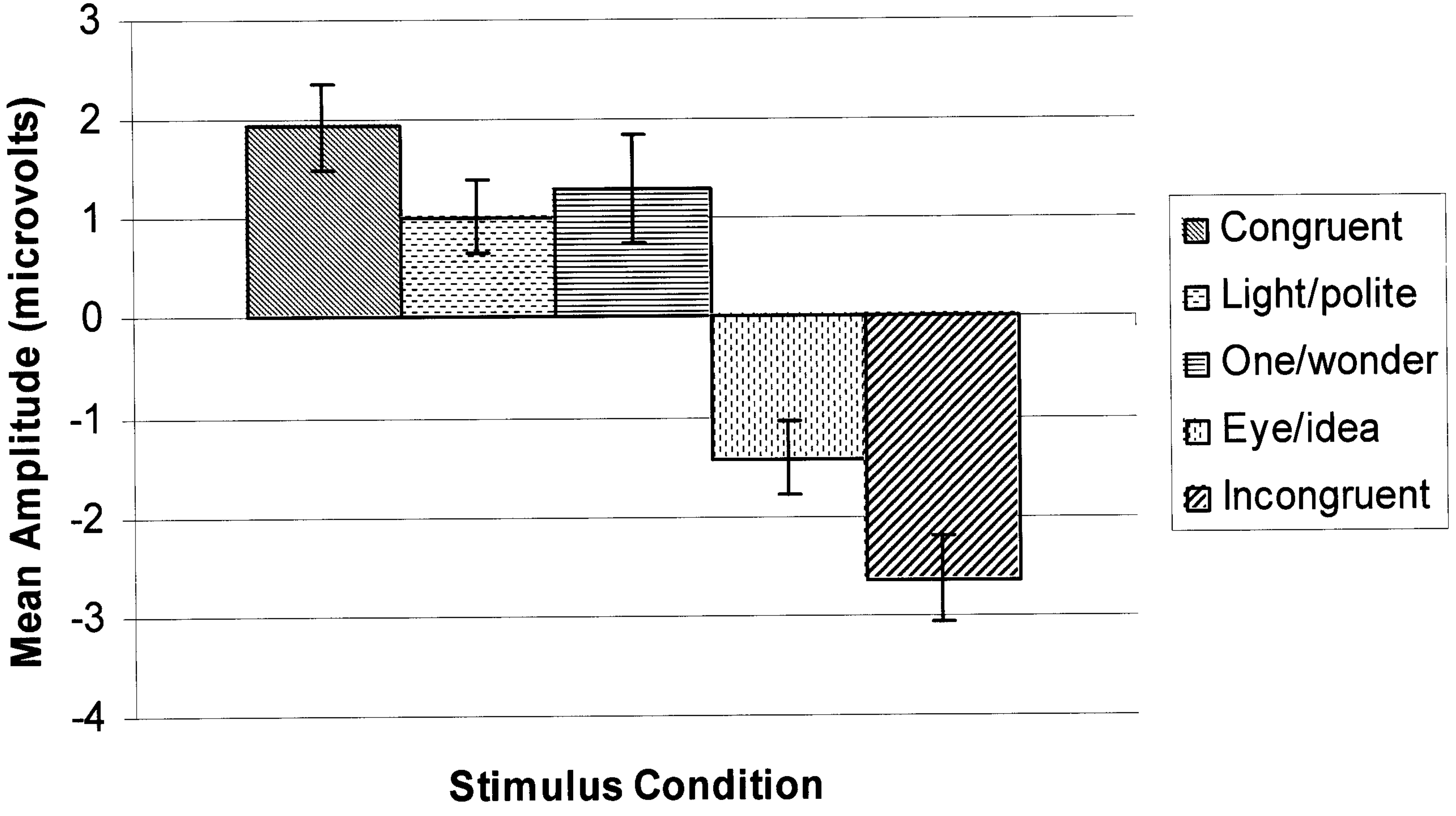
3
2
1
0
-1
-2
-3
-4
Mean Amplitude (microvolts)
Stimulus Condition
Congruent
Light/polite
One/wonder
Eye/idea
Incongruent

Figure 2

Figure Caption

Figure 3. Graph depicting the significant interaction between the Condition and Time factors for PMN amplitude data in Experiment 1.

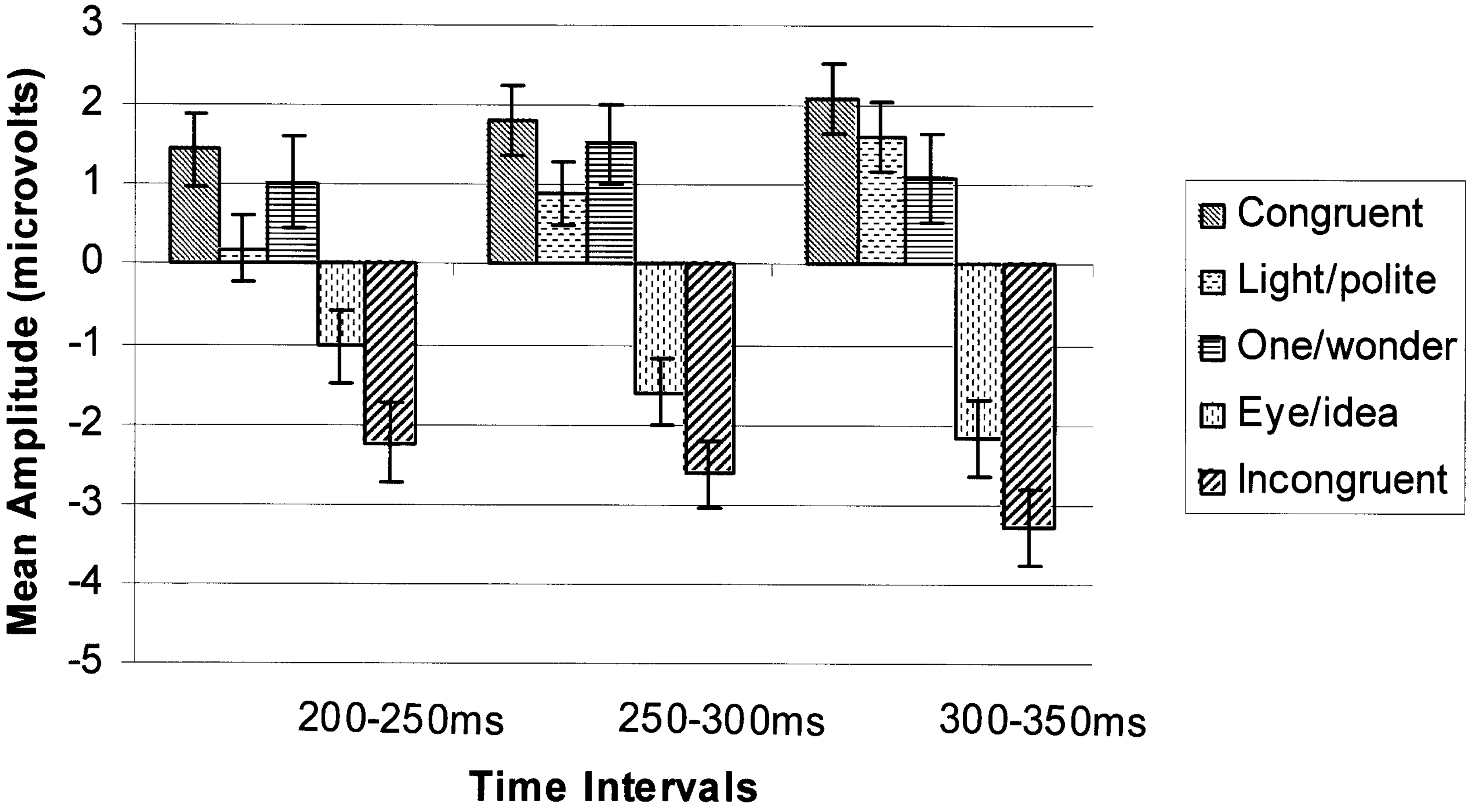
3
2
1
0
-1
-2
-3
-4
-5
Mean Amplitude (microvolts)
200-250ms
250-300ms
300-350ms
Time Intervals
Congruent
Light/polite
One/wonder
Eye/idea
Incongruent

Figure 3

found between the INCONGRUENT and EYE/**IDEA** conditions in the 200-250ms and 300-350ms time intervals (See Figure 3).

There were significant Condition x Site (F (64, 896) = 4.19, $p<0.001$, $\varepsilon=0.11$) and Time x Site (F (32, 448)= 2.89, $p<0.001$, $\varepsilon= 0.15$) interactions. The significant main effects and interactions are best interpreted within the significant three way Condition x Time x Site interaction (F (128, 1792) = 2.60, $p<0.001$, $\varepsilon=0.06$). The PMN displayed a left central distribution in the INCONGRUENT condition throughout the three time intervals (See Figure 7). In the EYE/**IDEA** condition, there was a fronto-central distribution in the 250-300ms time interval but a left frontal distribution in the 200-250ms and 300-350ms time intervals (See Figure 8; also see Appendix B Table B2). There was a fronto-temporal distribution in the CONGRUENT, LIGHT/PO**LITE** and ONE/**WON**DER conditions in the three time intervals (See Appendix B Table B2). Secondary analyses found no Hemisphere or Condition x Hemisphere effects on the PMN (See Appendix B Table B3).

N400

Analyses of the N400 confirmed its sensitivity to the lexical stress and sub-lexical phonology manipulations (F (4, 56)= 27.30, $p<0.001$, $\varepsilon=0.79$) (See Appendix B Tables B4, B5): the mean amplitudes in the INCONGRUENT condition (M = -0.93uv, SE=0.51) were significantly larger than those in the EYE/**IDEA** condition (M = -0.03uv, SE = 0.44) and the ONE/**WON**DER condition (M = 0.74uv, SE = 0.62) which were, in turn, larger than those in the LIGHT/PO**LITE** condition (M = 3.10uv, SE=0.55) and the CONGRUENT condition (M = 3.77uv, SE = 0.49). There was a main effect of Time

Figure Caption

<u>Figure 4.</u> Graph depicting the amplitude of the N400 response as a function of the stimulus condition across time in Experiment 1.

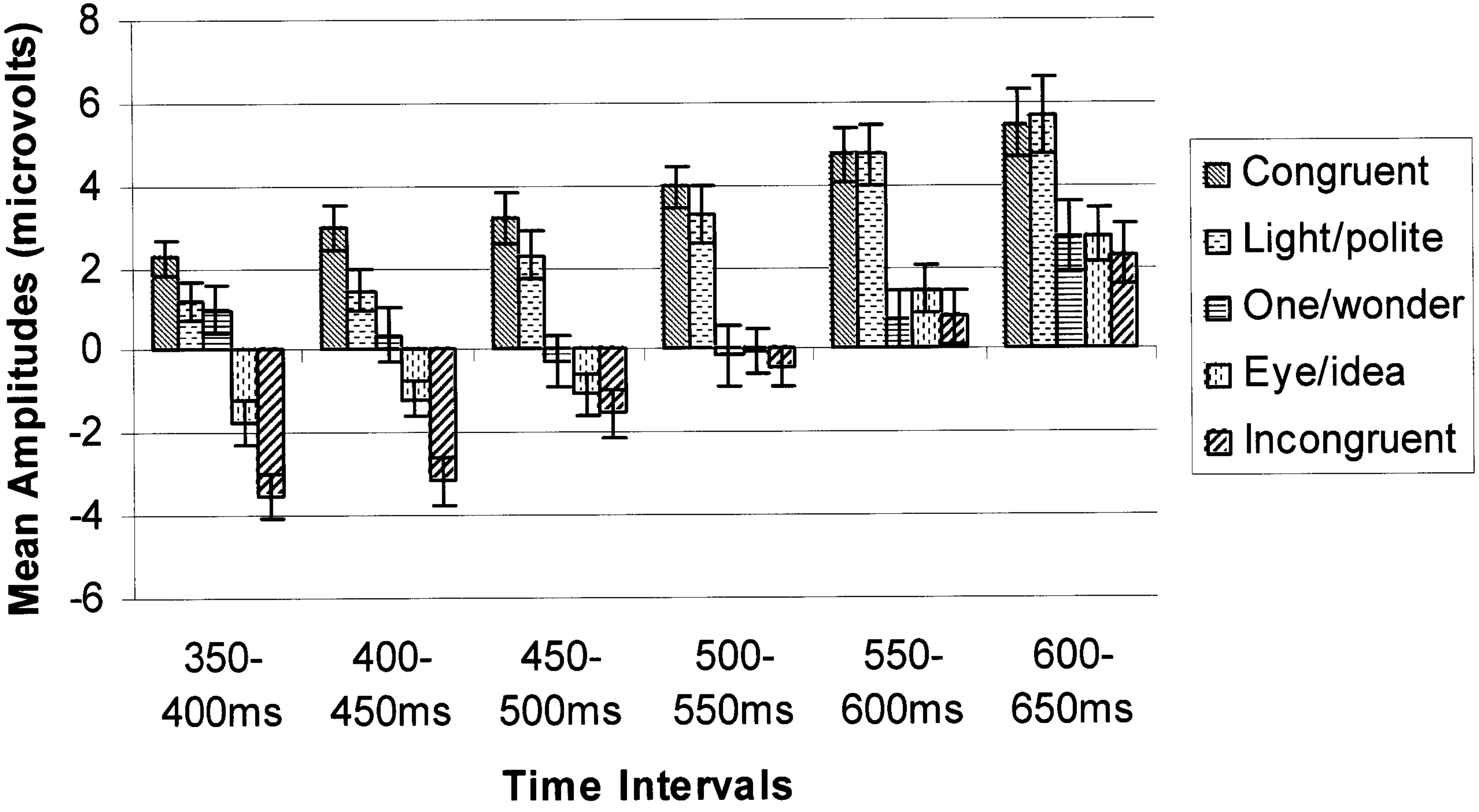
Mean Amplitudes (microvolts)
8
6
4
2
0
-2
-4
-6
350-400ms
400-450ms
450-500ms
500-550ms
550-600ms
600-650ms
Time Intervals
Congruent
Light/polite
One/wonder
Eye/idea
Incongruent

Figure 4

Figure Caption

<u>Figure 5.</u> Graph depicting the significant interaction between the Condition and Region factors for N400 amplitude data in Experiment 1.

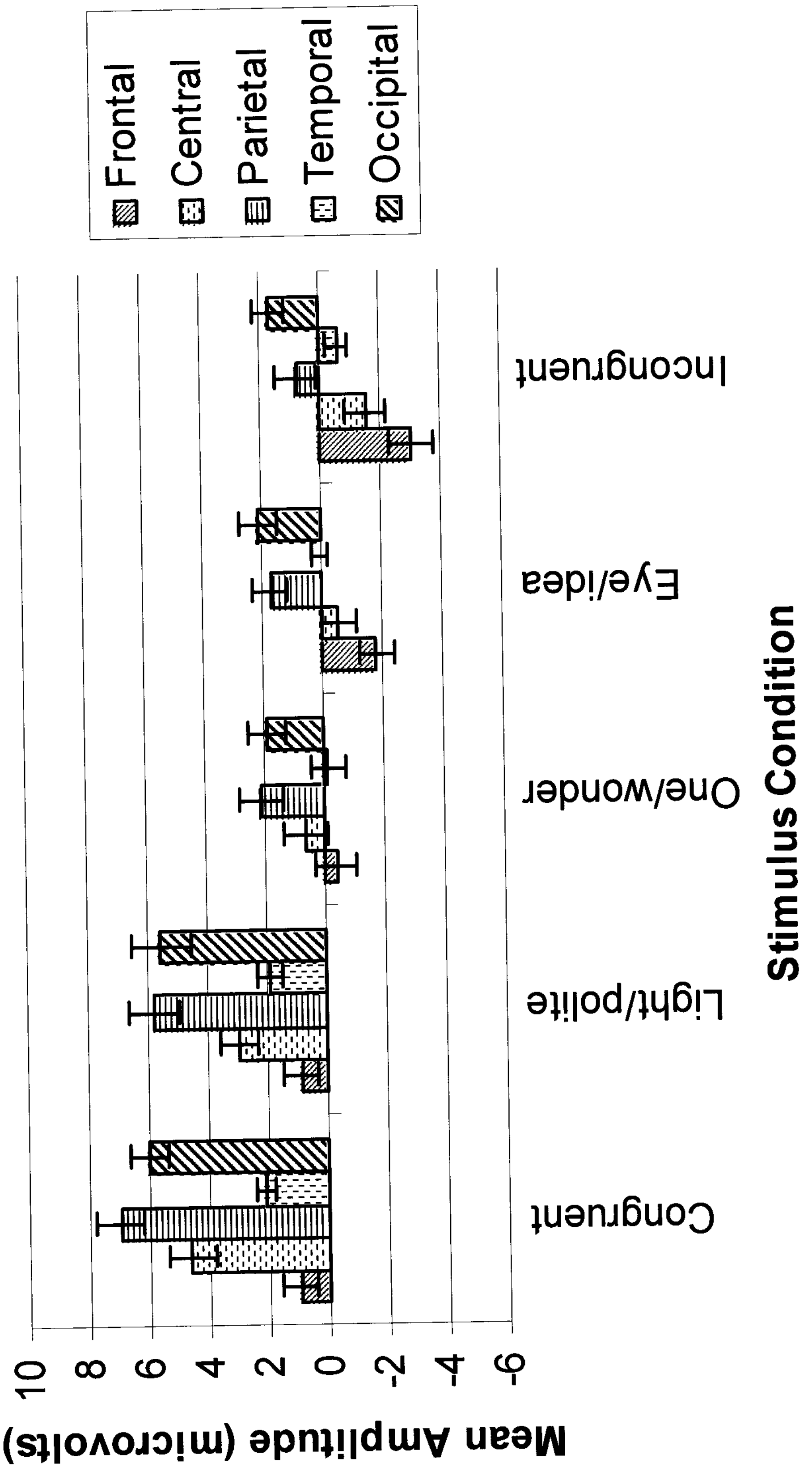
Mean Amplitude (microvolts)
10
8
6
4
2
0
-2
-4
-6
Congruent
Light/polite
One/wonder
Eye/idea
Incongruent
Stimulus Condition
Frontal
Central
Parietal
Temporal
Occipital

Figure 5

Figure Caption

Figure 6. Graph depicting the significant interaction between the Region and Hemisphere factors for N400 amplitude data in Experiment 1.

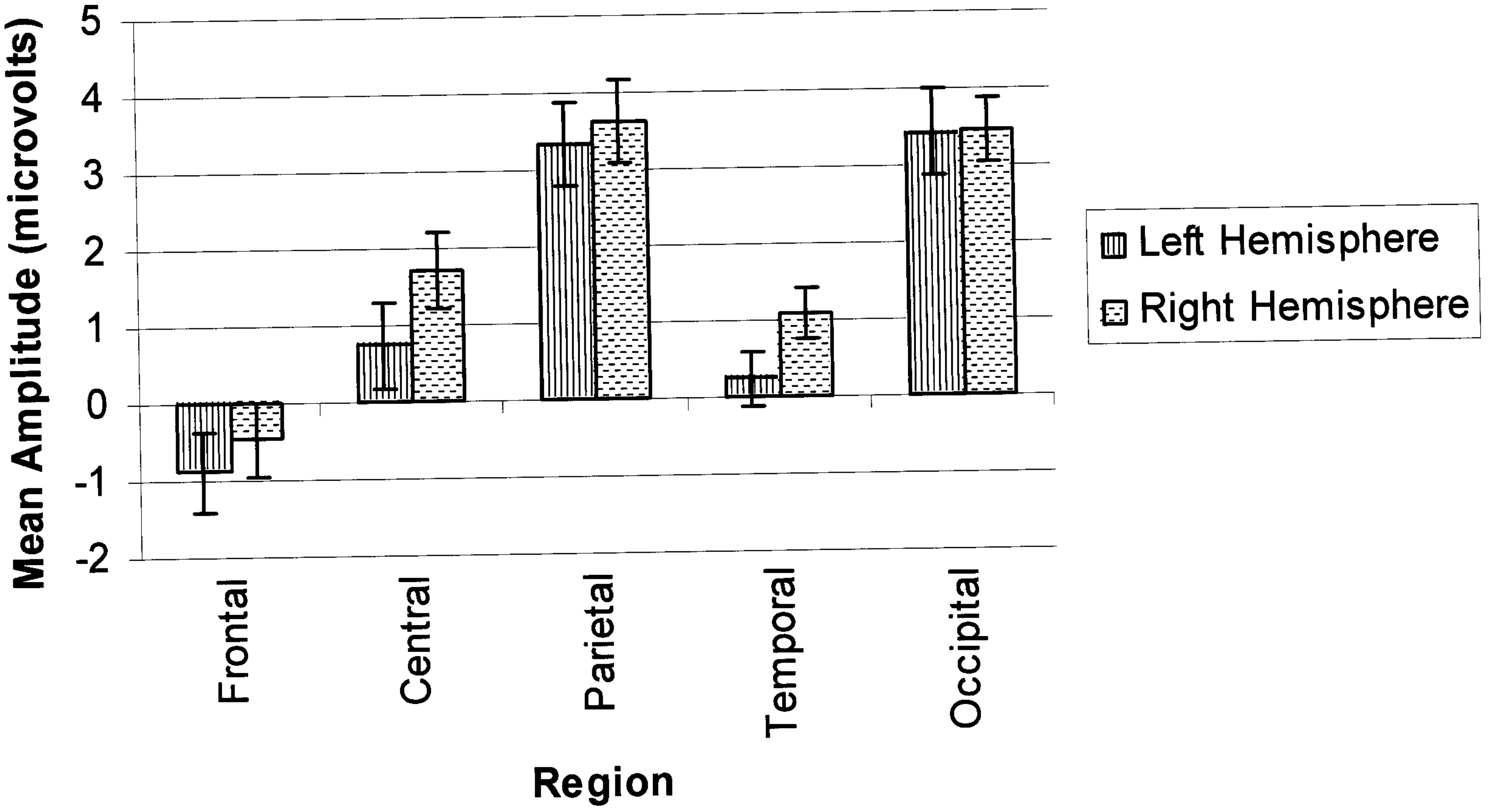
Mean Amplitude (microvolts)
5
4
3
2
1
0
-1
-2
Frontal
Central
Parietal
Temporal
Occipital
Region
▥ Left Hemisphere
▨ Right Hemisphere

Figure 6

(F (5, 70)= 34.44, $p<0.001$, $\varepsilon = 0.26$) that subsequent post hoc analyses revealed were due to larger amplitudes in the 350-450 ms time range. Also, a main effect of Site was found (F (16, 224) = 31.01, $p<0.001$, $\varepsilon = 0.17$) reflecting a left fronto-central scalp distribution.

There was a significant Condition x Time interaction (F (20,280) = 6.28, $p<0.001$, $\varepsilon= 0.30$) (See Figure 4) which subsequent post hoc analyses indicated were due to larger N400 amplitudes in the EYE/**IDEA** than ONE/**WON**DER condition in the 350-450 ms time range. However, no amplitude differences in the 450-650ms time range between these two conditions were found. There were significant Condition x Site (F (64,896) = 3.93, $p<0.001$, $\varepsilon = 0.12$) and Time x Site (F (80, 1120) = 22.93, $p<0.001$, $\varepsilon = 0.04$) interactions. The significant main effects and interactions are best interpreted within the significant three way Condition x Time x Site interaction (F (320, 4480) = 2.63, $p<0.001$, $\varepsilon= 0.02$) (See Appendix B Table B4). In the INCONGRUENT condition, there was a left central distribution in the 350-450 ms time interval but a fronto-temporal distribution in the 450-650 ms time interval (See figure 7). In the EYE/**IDEA** condition, there was a left fronto-central distribution in the 350-450ms time range but a clear left fronto-temporal distribution in the 450-650 ms time range (See figure 8). In the ONE/**WON**DER condition, there was a left temporal distribution in the 350-500ms time range but a clear left fronto-temporal distribution in the 500-650ms time range (See figure 9). In the LIGHT/PO**LITE** condition, there was a left fronto-temporal distribution in the 350-500ms time interval but a left temporal distribution in the 500-650 ms time interval (See figure 10; also see figures 5, 6). In the CONGRUENT condition, there was a fronto-temporal distribution across the time intervals (Figure 11). Secondary analyses further

Figure Caption

Figure 7. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the INCONGRUENT condition at 10 different latencies between 200 and 650 ms post stimulus in Experiment 1. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

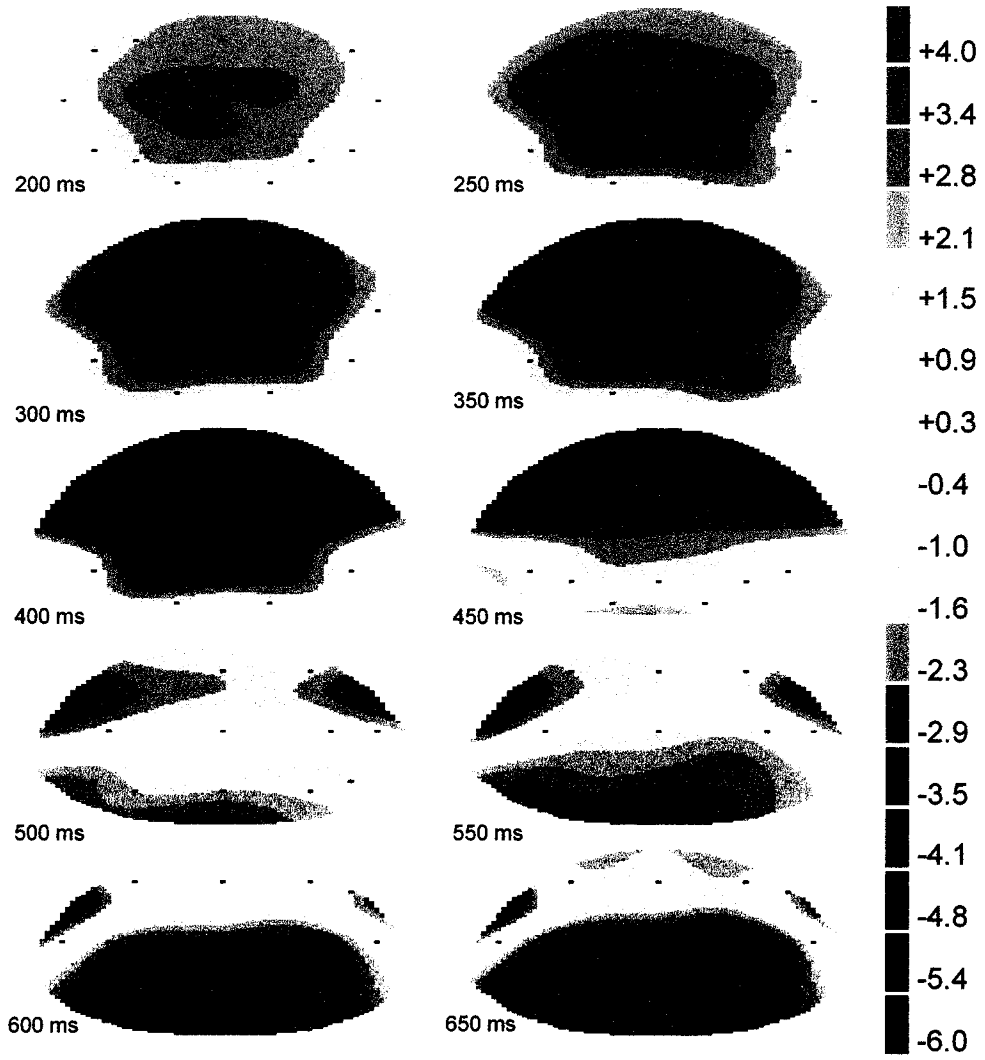
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
600 ms
650 ms
+4.0
+3.4
+2.8
+2.1
+1.5
+0.9
+0.3
-0.4
-1.0
-1.6
-2.3
-2.9
-3.5
-4.1
-4.8
-5.4
-6.0

Figure 7

Figure Caption

Figure 8. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the EYE/**IDEA** condition at 10 different latencies between 200 and 650 ms post stimulus in Experiment 1. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

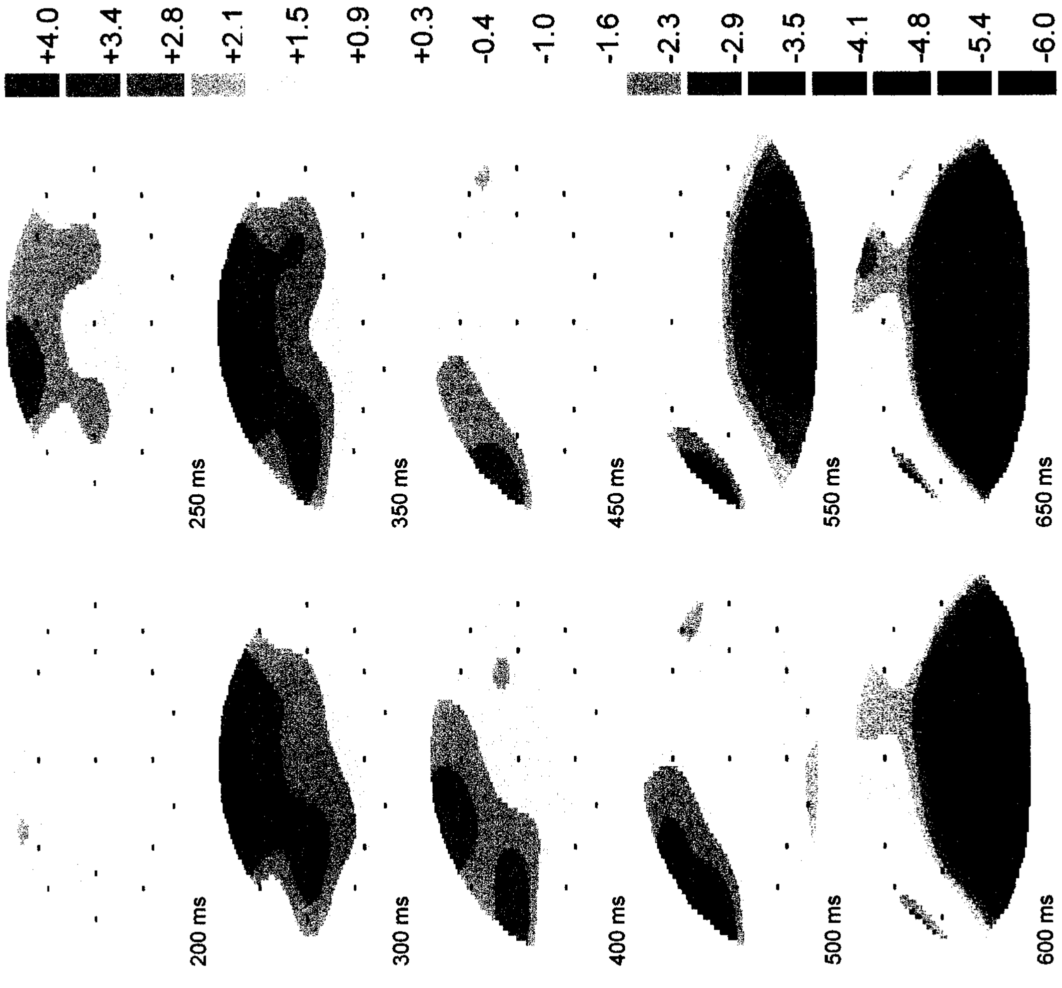
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
600 ms
650 ms
+4.0
+3.4
+2.8
+2.1
+1.5
+0.9
+0.3
-0.4
-1.0
-1.6
-2.3
-2.9
-3.5
-4.1
-4.8
-5.4
-6.0

Figure 8

Figure Caption

<u>Figure 9.</u> Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the ONE/**WON**DER condition at 10 different latencies between 200 and 650 ms post stimulus in Experiment 1. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

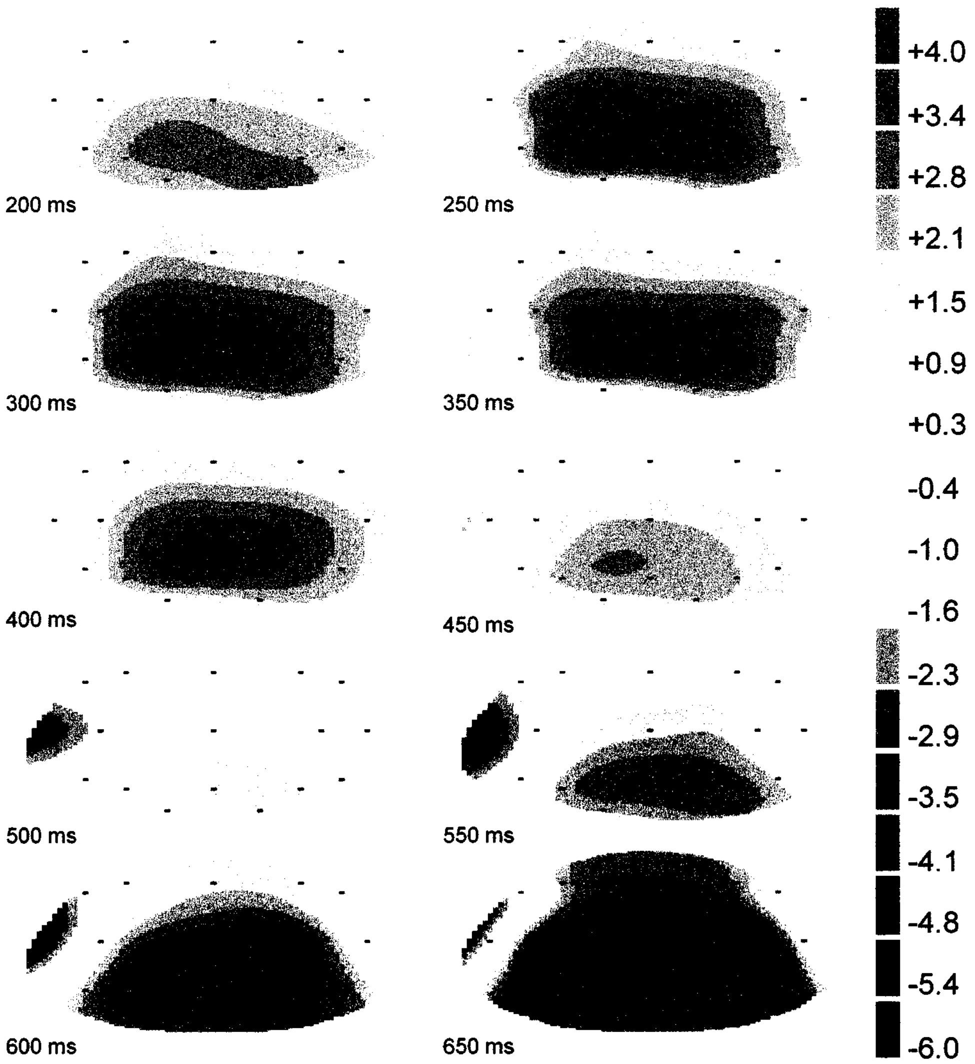
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
600 ms
650 ms
+4.0
+3.4
+2.8
+2.1
+1.5
+0.9
+0.3
-0.4
-1.0
-1.6
-2.3
-2.9
-3.5
-4.1
-4.8
-5.4
-6.0

Figure 9

Figure Caption

Figure 10. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the LIGHT/PO**LITE** condition at 10 different latencies between 200 and 650 ms post stimulus in Experiment 1. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

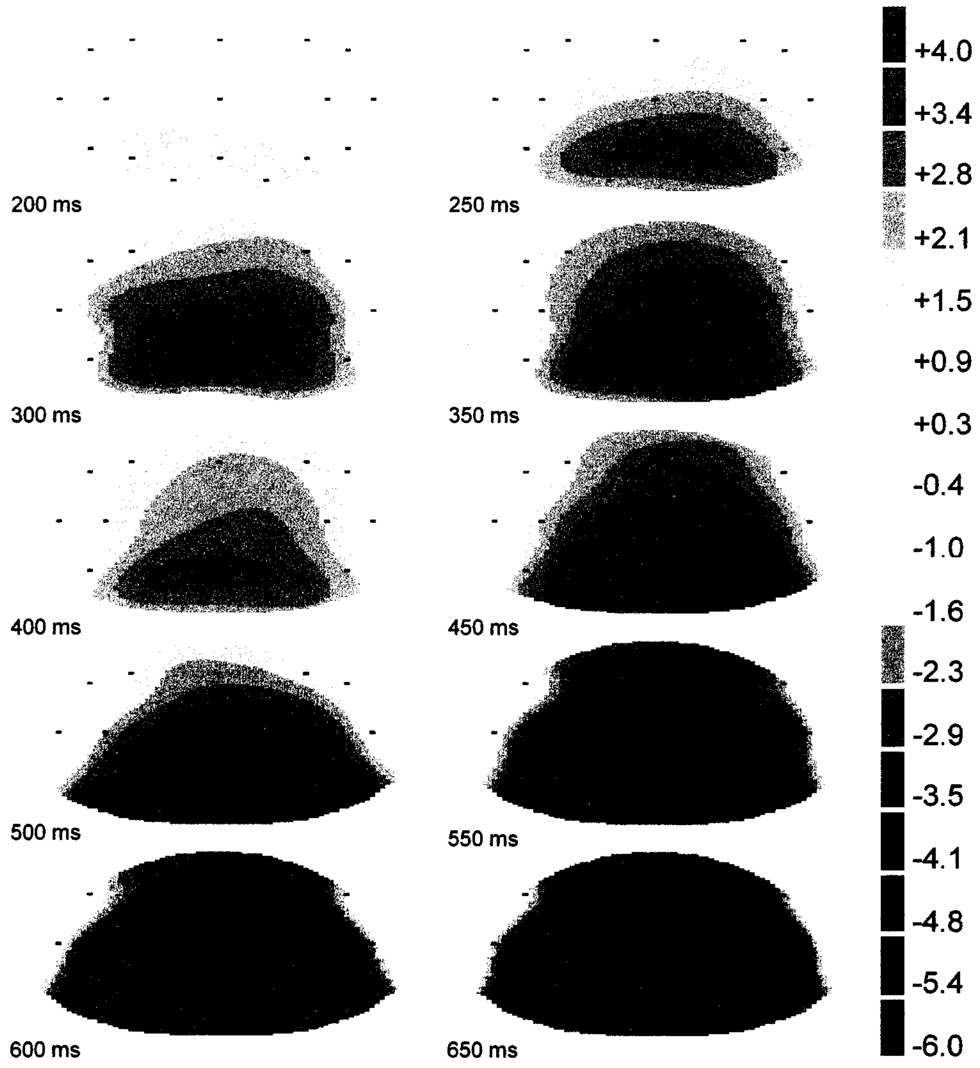
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
600 ms
650 ms
+4.0
+3.4
+2.8
+2.1
+1.5
+0.9
+0.3
-0.4
-1.0
-1.6
-2.3
-2.9
-3.5
-4.1
-4.8
-5.4
-6.0

Figure 10

Figure Caption

Figure 11. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the CONGRUENT condition at 10 different latencies between 200 and 650 ms post stimulus in Experiment 1. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

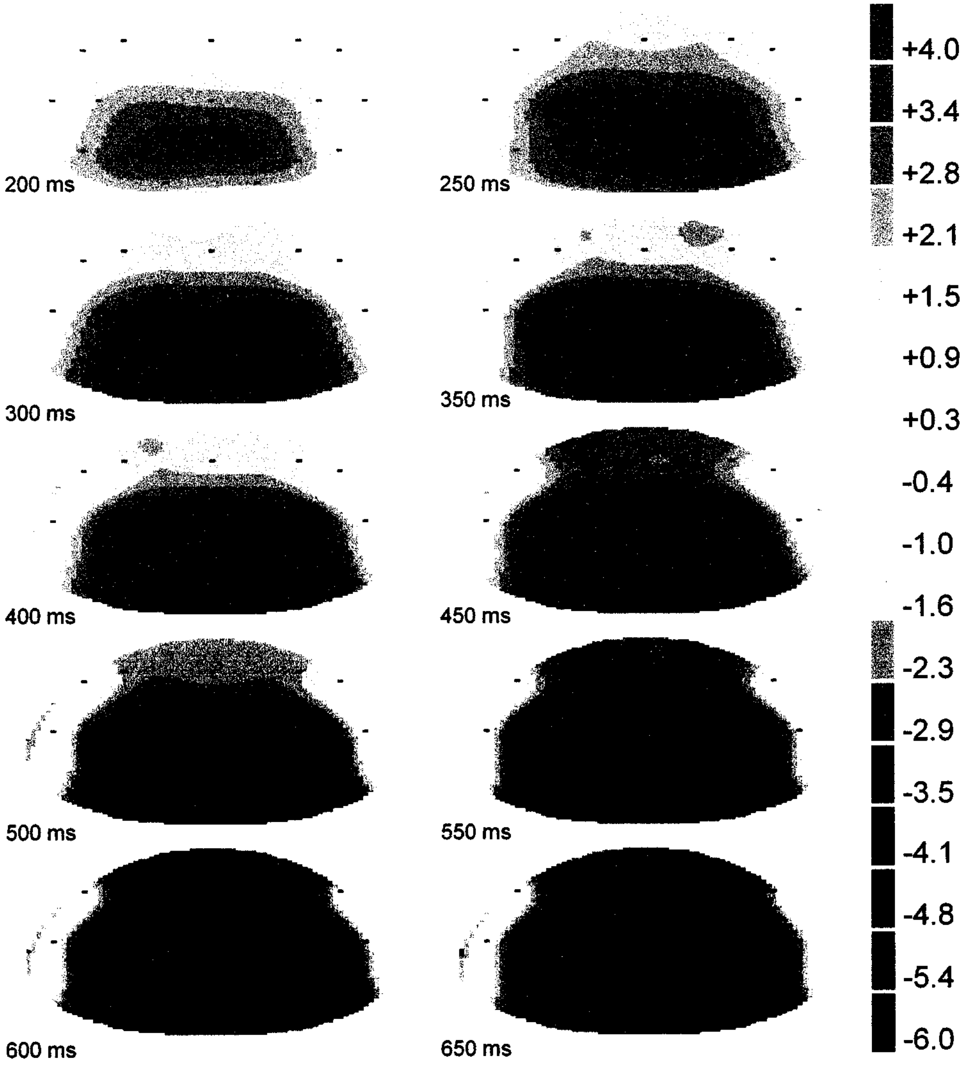
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
600 ms
650 ms
+4.0
+3.4
+2.8
+2.1
+1.5
+0.9
+0.3
-0.4
-1.0
-1.6
-2.3
-2.9
-3.5
-4.1
-4.8
-5.4
-6.0

Figure 11

Figure Caption

Figure 12. Individual waveforms of a representative participant at 17 recording sites for the CONGRUENT, LIGHT/PO**LIGHT**, ONE/**WON**DER, EYE/**IDEA**, and INCONGRUENT stimulus conditions presented in Experiment 1. For the display purpose, the low pass filter was set at 15 Hz for the ERP wave forms in this figure.

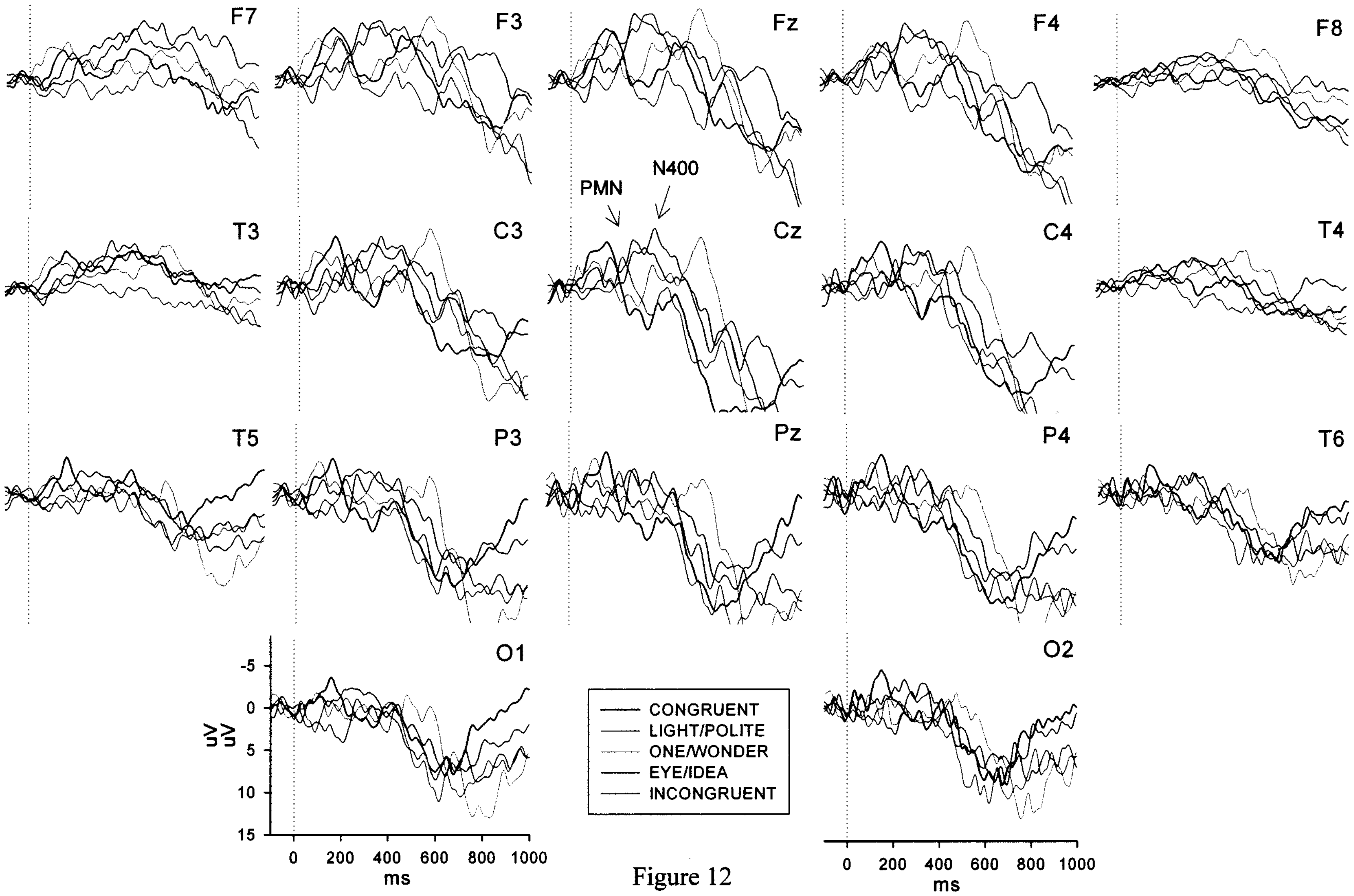
F7
F3
Fz
F4
F8
PMN
N400
T3
C3
Cz
C4
T4
T5
P3
Pz
P4
T6
O1
O2
-5
0
5
10
15
uV
0
200
400
600
800
1000
ms
CONGRUENT
LIGHT/POLITE
ONE/WONDER
EYE/IDEA
INCONGRUENT

Figure 12

indicated that there was a Hemisphere main effect (F (1, 14) = 5.03, $p<0.05$) reflecting a left hemisphere distribution (See Appendix B Tables B6, B7).

2.5 Discussion

The primary goals of the present experiment were to examine: first, whether word stress, regardless of its temporal position in an aurally presented word, or word onset plays a predominant role in initiating lexical processing; and second, to determine the effects that syllable-sized sub-lexical phonology exerts on lexical access to semantic representations of spoken words. The ERP findings obtained in this experiment demonstrate that ERP waveforms are modulated by primary stressed information (regardless of whether or not it was positioned on the word onset) as well as syllable-sized sub-lexical phonological manipulations. Two major findings emerge from this study, which will be fully delineated and interpreted and whose implications for spoken word recognition models will be discussed in detail.

The first major finding in this study is that an early negative component that peaks in the 250-350 ms time range after the target word onset and has a frontal distribution presents a striking differentiation between the five conditions according to the stress status of contextually expected syllable-sized sub-lexical information. It was revealed that the early negative component in response to the target word "**idea**" in the EYE/I**DEA** condition appeared almost as large as that elicited in the INCONGRUENT condition even though the word initial phonology /ai/ in the target "**idea**" was highly expected by its preceding sentential context "The pirate wore a patch over his ___". It is noteworthy that virtually no early negative activity in the 200-350 ms time range was seen in the

CONGRUENT, LIGHT/PO**LITE**, and ONE/**WON**DER conditions. Given its frontal scalp distribution, its time course (the 200-350ms time window), as well as its sensitivity to phonetic-phonological cues, this early negative component is clearly the phonological mismatch negativity (PMN) that was first explicitly identified by Connolly and Phillips (1994).

Most contemporary ERP researchers have been in general agreement that the amplitude of the PMN reflects the ease with which acoustic-based abstract representations are processed (phonological encoding) in an attempt to establish contact between the acoustic input and semantic meaning of a spoken word (e.g., Connolly & Phillips, 1994). When the information that initiates phonological processes is expected (primed) by preceding contextual cues, the amplitude of the PMN component would be correspondingly reduced. This phenomenon is called the phonological priming effect of the PMN. In line with this view, the reduced amplitudes of the PMN to the target "po**lite**" and the target "**won**der" may suggest a clear phonological priming effect that was apparently derived from preceding contextual expectations respectively on the phonological information /lait/ (in the target "po**lite**") and the information /wΛn/ (in the target "**won**der"). It is worth noting that these phonological priming effects on the PMN can be construed as strong evidence for the processing primacy of primary stress instead of word onset information for lexical initiation. The logic behind this suggestion is that the only common feature of these primed syllables was their primary stress status instead of their temporal position in the words.

As a matter of fact, the PMN data obtained in the EYE/**IDEA** condition further corroborates this view. Although the word initial phonological information /ai/ in the target “**idea**” was highly expected by its preceding sentential context, the “**idea**” PMN nonetheless did not demonstrate any phonological contextual priming effect: particularly over the frontal sites, where it appeared as large as the PMN elicited in the INCONGRUENT condition. A possible explanation for the minimal priming effect on the “**idea**”-elicited PMN is that word onset phonological information /ai/ is not used at the early stage of lexical processing to initiate phonological representations of the target “**idea**” given its unstressed prosodic status. By this account, it seems reasonable to argue that lexical analysis does not exclusively start with the temporally earlier-occurring input such as word onset information. Rather, lexical processes give priority to primary stressed cues regardless of their temporal position in a word. Put differently, primary stress is of crucial relevance to provide the access code to the mental lexicon. In line with this view, the present PMN findings do not seem compatible with what was hypothesized based on either the “goodness of fit” or the “sequential” hypothesis.

Firstly, the models on the basis of the “goodness of fit” hypothesis (e.g., shortlist, TRACE) (McClelland & Elman, 1986; Norris, 1994) have no way of predicting the PMN findings in this experiment. According to the TRACE model, for example, lexical access attempts, in principle, can be initiated at every point of acoustic input, which leads to the assumption that any partial phonological cues about a word that can be expected or primed by its preceding context would facilitate the process of lexical initiation. Based on this model, the PMN should not be sensitive to the discrete stress status of the primed sub-lexical phonological information of a target word. That is, the amplitudes of PMN in response to the target “**idea**” in which the syllable /ai/ was primed, the target “**won**der” in which the syllable /wΛn/ was primed, and the target “po**lite**” in which the syllable /lait/

was primed should show no difference from each other. As well, the PMN to these three syllable-sized sublexical phonologically primed targets should all appear smaller than that to the INCONGRUENT target word. However, the present data show that this is not the case. Although the theory could manage to interpret the PMN waveform in response to the targets "**won**der" and "po**lite**", it cannot account for the small phonological priming effect on the PMN to the target "**i**dea". To this point, the "goodness of fit" hypothesis, at least at the stage of lexical access, is not tenable.

Secondly, the present PMN findings also run counter to the Cohort model on spoken word recognition, which proposes a strict time-shadowing of the acoustic stream leading to lexical access (e.g., Marslen-Wislon, 1980, 1987). At a first glance, the priming effect on the PMN observed in the ONE/**WON**DER condition is easily implemented in the Cohort model since the PMN effect can be construed as the consequence of the preceding contextual influence on the word initial phonology /wΛn/ during the phonological processing of the target "**won**der". In this view, the priming effect on the "**won**der" PMN can be taken as compelling evidence that lexical processing at an early stage starts with the initial sound information of a word. Furthermore, the processing primacy on the early-occurring phonological cue /wΛn/ in "wonder" appears to reinforce the Cohort's view that lexical access is critically dependent on reception of temporally early sounds of a word (word onset phonemes or syllables).

These facts notwithstanding, the sequential Cohort model meets great difficulty in accounting for the PMN effects in the LIGHT/PO**LITE** and EYE/**I**DEA conditions. If the strictly sequential processing reflects the true nature of the time course of spoken word recognition, activation of phonological encoding should begin on the initial sounds of any

English word. By this account, the cognitive system should start phonological analysis with the initial input /ai/ in the target word “**idea**” and the initial sound /p∂/ in the target word “po**lite**”. Given that the phonological information /ai/ in the target “**idea**” was highly expected by its preceding contextual information while the information /p∂/ in the target “po**lite**” was not, the PMN to the word ‘**idea**’ should appear smaller in contrast to the PMN to the target ‘po**lite**’. Simply put, according to the sequential hypothesis, the phonological encoding of the word “po**lite**” that was manifested by the PMN component should not present any phonological priming effect that was in fact derived from the preceding contextual expectation on the phonological information /lait/; yet such a priming effect on the PMN is precisely what has been demonstrated in this experiment.

In an attempt to elucidate mental processes manifested by the PMN waveforms in the five conditions, an explanation for the PMN findings on the basis of the stress guide hypothesis is proposed. In point of fact, the evident phonological priming effects on the PMN to both the targets “po**lite**” and “**won**der” can be taken as converging evidence that primary stressed information plays a predominant role in initiating lexical processing. More importantly, using the “stress guide” hypothesis, the lack of a contextual facilitation effect on the “**idea**” PMN can be also appropriately accounted for. The “stress guide” hypothesis claims that lexical search is not necessarily initiated from word onsets but from primary stressed information instead (Mattys & Samuel, 2000). In line with this view, it is logical to argue that phonological encoding for the target “**idea**” was initiated with the primary stressed information /di∂/ which was, however, not expected or primed by its preceding contextual information. Although word initial information /ai/ is highly expected by its preceding contextual information, this facilitation effect on the word onset did not draw significant attention from the speech cognition system given its unstressed status and therefore appeared of secondary importance during phonological encoding of the whole target word. This interpretation would account for the finding that the PMN to the target "**idea**" appeared almost as large as that elicited in the INCONGRUENT condition.

Meanwhile, what needs to be kept in mind is that the significant PMN contextual effect observed in the ONE/**WON**DER condition cannot be thought of merely as ambiguous data having the ability to distinguish the validity of the stress guide hypothesis from the "sequential" or Cohort hypothesis. Rather the data from the ONE/**WON**DER condition is further evidence of the more general mechanisms of stress guide models.

Although there has been little direct empirical evidence for the role of lexical stress in lexical processing, several lines of behavioral and clinical research have provided indirect but convincing evidence in support of the stress guide hypothesis. In an influential study done by Taft (1984), bi-syllabic stimuli either with initial stressed patterns (e.g., lettuce, let us) or with non-initial stressed patterns (e.g., assign, a sign) were presented to native English listeners who were required to judge whether the presented string was one or two words. It was found that listeners tend to consider the first pattern as one word ("lettuce" rather than "let us") and two words for the second pattern ("a sign" instead of "assign"). This finding was interpreted as strong evidence for the argument that it is stress that is used to segment the speech stream and that stress is important at an earlier stage of lexical processing to specify a set of potential candidates. Further support for the stress guide hypothesis comes from Cutler and Norris (1988) who required native English listeners to detect a word (e.g., mint) in a nonsense sequence that had either two stressed syllables (e.g., mintayve) or a stressed and an unstressed syllable (e.g., mintesh). They hypothesized that if spoken word recognition is a stress-guided process, detection of the word (mint) would meet with difficulty in the first case because the nonsense string is segmented into two units based on two stressed syllables (min/tayve). The findings supported this prediction: listeners indeed used stressed information to segment the speech signal. Cutler and Butterfield (1992) demonstrate that the majority of lexical juncture misperceptions induced by experimental manipulations came from the erroneous insertion of a word boundary before a stress and this stress effect was observed in about 70% of the mis-segmentations of continuous speech found in corpora of spontaneous "slips of the ear".

Another line of evidence for the importance of lexical stress on spoken word recognition has been obtained by examining the influence of stress pattern violation on word recognition. It was hypothesized that the "stress guide" strategy should mislead the system to segment speech if the acoustic correlates to stress were not accurately picked up. Indeed, the mis-placing of lexical stress on words has been shown to significantly interfere with recognition (Bond & Small, 1983; Cutler & Clifton, 1984; Slowiaczek, 1987,1990). For instance, Slowiaczek (1990) used three kinds of tasks to examine the effects of lexical stress on the processing of spoken words; word identification, shadowing, and lexical decision. Stimuli were divided into two types: one type used words with correct stress patterns and the other included items where the primary stress was moved to a normally unstressed syllable (e.g., **au**dience was changed to au**dience).** The results demonstrated that correctly stressed items were identified (in the identification task), produced (in the naming task), and classified (in the lexical decision task) faster and more accurately than incorrectly stressed items. This evidence indicates that correct stress locus of a spoken English word plays a critical role at an early stage of lexical processing to initiate lexical search and thereby activate a set of potential candidates.

The PMN findings in the present study, taken together with the behavioral literature, provide convincing support for the "stress guide" hypothesis emphasizing that lexical search starts with stressed acoustic input regardless of its temporal position in words. In contrast, word initial but unstressed information does not play any predominant role in computing phonological codes of a word. It would be likely that the application of primary stressed information to lexical initiation is a default strategy in auditory English word processing.

In evaluating the stress guide hypothesis, an interesting but critical question emerges: how are non-initial-stress words (e.g., "po**lite**") accurately processed and recognized if the cognitive system relies on non-initial primary stressed acoustic input (e.g., /lait/ in the word "po**lite**") to initiate lexical search. In practice, all the participants in this experiment did not show any difficulty in recognizing the words whose primary stress was not placed at the beginning. To remain faithful to the processing primacy on stressed information in lexical search, proponents of the hypothesis propose that the stress guide processing can be complemented with a subsidiary operation by which a non-initial-stress word can still in some fashion be successfully processed through their stressed syllables (Cutler, 1989). The subsidiary operation allows the cognitive system to remedy incorrect lexical access via some repair operations such as retroactivity (Bradley, 1980; Cutler, 1989; Gow & Gordon, 1993). For example, the cognitive system can repair incorrect stress-guided lexical search by backtracking to the initial information occurring before the stressed syllables, known as the "retroactive" strategy (Mattys & Samuel, 2000).

Although there has been little direct empirical evidence for this speculation, retroactive strategy has been demonstrated repeatedly and utilized as a true mechanism underlying non-initial stressed word processing. It is widely believed that retroactive processing occurs to undo the errors from the lexical attempts using primary stressed syllables. When the stress-guided lexical search does not achieve any results, lexical access is then attempted using the unstressed syllable immediately preceding the stressed syllable. If this attempt also fails, lexical access is then attempted through the unstressed syllable preceding the one used in the last search (Mattys & Samuel, 2000). Given the processing "cost" of this "retroactive" strategy, it is worth noting that extra time or at least extra work is needed when non-initial-stress words are processed in contrast to initial stressed words. As a matter of fact, in an attempt to support the viability of this retroactive strategy, some researchers monitored participants' response time in speech perceptual studies and took the delayed responses for non-initial-stress targets that were observed as compelling evidence supporting this retroactive strategy (Mattys, 1997).

When taking the present ERP findings into consideration, it is logical to argue that if there is a "cost" to the processing of non-initial-stress words, the processing of both the target "po**lite**" and the target '**idea**' should give rise to a delayed time course of lexical access to semantic representations of target words. As a consequence, there should be a delay in the occurrence of the N400 component in response to the targets "po**lite**" and "**idea**" in contrast to the timing of the N400 to the stress-initial target words in the INCONGRUENT condition. However, there is no sign of any delayed N400 components in the present study. Instead, N400 latencies in both the LIGHT/PO**LITE** and EYE/**IDEA** conditions were similar to those found in the INCONGRUENT condition. Thus, in this respect, the present findings do not support a model that relies on a retroactive strategy.

Before shedding further light on the mental processes underlying non-initial-stress word processing, it is worth emphasizing that both PMN and N400 data in the present study present a firm stand on the view that primary stress plays an important role not only in phonological processing reflected by the PMN component but also in semantic search analysis reflected by the N400 component. As well as the PMN data modulation of stress cues, the early part of the N400 wave form (in the 350-450ms time window[1]) likewise shows sensitivity to the stress status of the contextually primed syllable of the target word: the amplitude of the N400 in response to the target whose primed syllable is primarily stressed, regardless of its physical position in the word (see the N400 waveforms in the LIGHT/PO**LITE** and ONE/**WON**DER conditions) is significantly reduced; in contrast, the amplitude of the N400 to the target whose primed syllable is prosodically unstressed (in the EYE/**IDE**A condition) is not reduced. To illustrate, the N400 to the target "po**lite**" appeared much smaller than that to the INCONGRUENT target. In addition, although the amplitude of the N400 elicited by the word "**idea**" seems smaller than that to the INCONGRUENT word, this priming effect is nonetheless much smaller than the effect on the "po**lite**" N400. This finding indicates that the word initial information /ai/ does not play any leading role in semantic access due to its unstressed status. Data to this point

---

[1] The demonstration of the ERP in this time range is not meant to imply another component other than the N400; rather, the time label is simply adopted for clarification purposes.

thus suggest that when taking the present N400 findings into consideration, the extent to which contextually primed syllabic information exerts its effect on word processing is largely determined by its stress status in a word. With this in mind, it seems reasonable to argue that primary stressed and unstressed information receives different weighting during the course of lexical processing.

Based on the preceding information, one can speculate that a strategy for non-initial-stress word recognition must deal with the two following points: first, primarily stressed information initiates lexical analysis and plays a predominant role in constraining semantic representations of spoken words; and second, as Grosjean and Gee (1987) first stated, concurrent with the process dominated by primarily stressed syllables, there is a process dealing with initial unstressed information. Processing of the initial unstressed portion of a word requires additional memory storage so that the temporarily unprocessed word initial information can be stored and retrieved when the primary analysis of the stressed syllable starts. In point of fact, several memory models have included this type of function. According to Nooteboom's theory (1979, 1981), there are two kinds of memory to hold acoustic-phonetic information during spoken word recognition: a short-lived physical-acoustical buffer and a more perceptual-auditory short term memory. It is demonstrated that the auditory information stored in memory in general persists long enough to create acoustic or perceptual "rehearsal". In all, although the parallel processing proposal is still sketchy and general, it is more compatible with the present ERP findings than the models favoring "retroactive" strategy.

In addition, it is worth keeping in mind that the parallel processes proposal requires an ability to isolate the initial unstressed syllable from the word preceding the target word so that the system senses the initial unstressed syllable as potential part of the target word. There may be a fast backward scanning system that can locate this unstressed part. Another possibility is that the listener's knowledge of phonotactic and morphophonemic rules (e.g., functors and affixes) also exerts some effects on signaling the initial unstressed syllable. Unfortunately, the means and procedures through which initial unstressed acoustic input for non-initial-stress words can be processed are as yet unclear. The attempts to clarify this issue may be the impetus for future research.

The second major finding in this experiment, which has been grossly stated above, is that there are clear sub-lexical phonological priming effects on the N400. The N400 elicited in the ONE/**WON**DER, EYE/**IDEA,** and LIGHT/PO**LITE** conditions appear smaller than that in the INCONGRUENT condition, which suggests a clear syllable-sized sub-lexical facilitation effect on the N400. More relevant, the N400 in the ONE/**WON**DER and LIGHT/PO**LITE** conditions are smaller than the N400 in the EYE/**IDEA** condition. It indicates that the magnitude of sub-lexical phonological priming effects on lexical access to semantic representations is constrained by the stress status of primed syllables. Given that the unstressed information /ai/ plays an ancillary role in the processing of the word "**idea**", the primed but unstressed syllable /ai/ does not facilitate the whole word processing as much as that of the primed and stressed syllables /wΛn/ in "**won**der" and /lait/ in "po**lite**".

The sub-lexical phonological effects on the N400 in this study raise a question of whether the preceding contextual priming merely on partial word phonological information facilitates or inhibits recognition of the whole spoken word. This issue is currently the focus of considerable interest and debate in psycholinguistics, which remains controversial both as an empirical result and as a theoretical issue.

Compatible with the present ERP findings, a large body of behavioral studies demonstrate that there is a clear facilitation effect when the prime and target words share merely sub-lexical phonological information (Goldinger, Luce, Pisoni, & Marcario, 1992; Nusbaum & Slowiaczek, 1983; Slowiaczek, Nusbaum, & Pisoni, 1987; Slowiaczek & Hamburger, 1992). For example, Slowiaczek and colleagues (1987) conducted three phonological priming experiments in which participants were asked to identify primed and unprimed target words embedded in noise. In this study, an auditory prime was presented which, 50 ms after its offset, was followed by the auditory target. Four-phoneme monosyllabic target words were used and the extent of phonological overlap between primes and targets was varied, which ranged from trials without any overlapping phonemes between primes and targets to trials with fully shared phonemes. A priming facilitation effect was readily observed across three experiments as the probability of identifying a word increased with the degree of phonological overlap between primes and targets. Compelling support for this facilitation claim is also provided in another influential study. Slowiaczek and Hamburger (1992) conducted six auditory single-word shadowing experiments in which participants repeated an aurally presented target item as fast and accurately as possible upon hearing it. Word or nonword prime items were presented either aurally or visually. The findings in this study replicate and extend the results of the Slowiaczek et al.'s study by demonstrating that partial word phonology overlap between the prime and target mono-syllabic words facilitates auditory word recognition.

However, this facilitation view on sub-lexical phonology has spawned a lively controversy over the generality of the above sub-lexical facilitation findings. Some argue that this facilitation effect of phonological priming is difficult to replicate in some other lexical decision studies (Radeau, Morais, & Dewier, 1989; Slowiaczek & Pisoni, 1986). For instance, Slowiaczek and Pisoni (1986) did two speeded auditory lexical decision experiments and found that there was a reliable facilitation effect on lexical decisions only for identical phonological priming (repetition). However, the facilitation effect did not hold up for trials with one, two, or three overlapping phonemes between target and prime items. In a similar vein, performance data contradictory to the facilitation view have also been reported (Goldinger, Luce, & Pisoni, 1989). Similar to the study of Slowiaczek et al. (1987), Goldinger et al. examined perceptual identification of primed words in noise. In contrast to previous work (Slowiaczek et al., 1987), however, they examined priming effects when phonetic information of the prime and the target which shared no common phonemes, became confusable when presented in noise. The results of this study suggested an inhibitory priming effect: the success rate of target identification became decreased when the target items followed phonetically similar primes than when they followed phonetically unrelated primes. All this evidence has been interpreted as compelling evidence for the argument that sublexical phonological priming plays an inhibitory or interference role in whole word recognition.

As a matter of fact, some influential word recognition models, such as TRACE, have been developed as strong theoretical support for this sub-lexical phonological inhibition view. Proponents of the TRACE model (Elman & McClelland, 1986; McClelland & Elman, 1986) claim that, in word recognition processes, there are several levels of representation for features, phonemes, and words. On each level there are lateral inhibitory connections among nodes. This lateral inhibition between lexical candidates activated by the acoustic input during spoken word recognition renders possible that better fitting or more active candidates can directly suppress the activation level of less fitting competitors. For example, when the prime "word" and the target "work" pair are presented, lateral inhibition would exert its effect on the target word processing due to shared phoneme features between the prime and the target. This seems convincing to account for the aforementioned inhibition performance data. However, caution needs to be exercised before embracing this conclusion. It is commonly argued that the validity of a word recognition model is confirmed only if the method with which the model is reinforced taps on a pre-lexical or lexical stage during spoken word processing. If the method instead taps on a post-lexical stage, the phonological effects obtained would not imply a true picture of phonological top-down influence of preceding contextual information on lexical processing but post-lexical decision bias (Goldinger, Luce, Pisoni, & Marcario, 1992). Most of the performance data described above resulted from lexical decision tasks that are believed to include considerable postlexical sub-processes (Balota & Chumbley, 1984; Forster, 1981; Jakimik, Cole, & Rudnicky, 1985; Seidenberg, Waters, Sanders, & Langer, 1984; West & Stanovich, 1982). This may account for the discrepancy between the facilitation data and the inhibition data. Setting this problem to one side, it is worth noting that although the behavioral studies above address the same question on the potential effect that sub-lexical phonology may exert in spoken word recognition, these studies manipulated different dimensions of word processing. For example, stimuli used in the Goldinger et al. (1989) study were in fact manipulated in the phonetic dimension, the result of which was difficult to accept as evidence relevant to any discussion about the nature of phonological processing. This may be another explanation for the discrepant findings in the behavioral literature.

A particularly important point is that the TRACE model runs counter to the present ERP findings. According to TRACE, when the syllable /lait/ in the word "po**lite**" was played, the candidates sharing this syllable such as "light", "polite", and "delight" all would be activated. Given that the preceding sentential context in this study made the candidate "light" more expected, the candidate "light" should have shown more pre-activation than the target candidate "po**lite**". On the basis of the lateral inhibition theory in TRACE, the more strongly the item "light" is activated, the more it inhibits the level of activity of other candidates such as "po**lite**". In this case, the processing of the word "po**lite**" would be delayed or require extra effort to override the enhanced inhibitory force from the primed candidate "light". This effort in fact should make the "po**lite**" N400 larger and last longer even than that to the INCONGRUENT target word. However, that is not what was found. The peak of the "po**lite**" N400 was virtually observed in the same time range (350-450 ms post stimulus) as that of the "Incongruent" N400. Also, the "po**lite**" N400 appeared much smaller than the "Incongruent" N400.

In contrast with the position of the TRACE model, the Cohort model fares relatively well with the sub-lexical facilitation data on spoken word recognition. According to the classical Cohort model, word-initial information (e.g., /wΛn/) activates a cohort of words, such as "one" and "**won**der", which correspond with this acoustic input. Because the Cohort model does not involve lateral inhibition, the activation of each candidate does not compete with or inhibit the other. Viewed in this light, the enhanced activation of the syllable "one" due to preceding contextual expectation would facilitate all candidates' processes instead of being a factor to disfavor some of the candidates. When the later-occurring information /d∂/ is heard, a robust positive effect on the activation level of the candidate "wonder" is enhanced. Importantly, the simultaneous activation of "**won**der" does not exert any negative effects on the activity in the candidate node "one". According to the Cohort model, "competition" occurs in lexical activation but it comes into play only at the post-lexical decision stage (Marslen-Wilson & Warren, 1994). Meanwhile, the more strongly its competitors are activated, the longer the process takes until the optimal candidate is selected (Marslen-Wilson, Moss, & Halen, 1996). By this logic, it is easy to interpret the delayed decision time data and the facilitation data in the above behavioral studies. When taking the present PMN findings favoring the stress guide hypothesis into consideration, if we elaborate the Cohort model by replacing the emphasis of word onset with the primacy of primary stressed information, this refined model can account for the sub-lexical phonological priming effect on the N400 in the present experiment.

In turn, the present N400 findings, in conjunction with behavioral facilitation evidence, can be taken as strong evidence supporting the proposal that sub-lexical phonological priming has a facilitation instead of an inhibitory effect on spoken word recognition. It is reasonable to argue that the preceding contextual information boosts the activation levels of sub-lexical phonological processing, the consequence of which facilitates spoken word recognition. In line with this view, although contextually primed candidates (such as /lait/ during the target "po**lite**" recognition) may be highly activated, the enhanced activity of the candidate "light" does not interfere with or inhibit the recognition of whole word "po**lite**" at the lexical level. Rather, the elevated activation of the primed syllable /lait/ gives rise to a facilitation effect on the word "po**lite**" processing.

It is worth mentioning that the present evidence that primary stress plays a predominant role in both phonological and semantic processes raises another issue. Is lexical prosody represented in the mental lexicon?

As stated in the introduction, a good number of speech studies have emphasized that continuous speech input needs to be explicitly segmented into discrete units prior to lexical access of single spoken words and lexical prosody helps the English listener parse the speech input into discrete words (e.g., Culter, 1992; Luce, 1986). In this vein, it seems likely that stress exerts its effect at least on speech perception.

The present ERP findings provide strong evidence that word prosody is also of crucial relevance in lexical processing. As demonstrated in the present experiment, the processing primacy on stressed syllables during phonological encoding and the differential weighting of sub-lexical priming effects on semantic constraint due to the different stress status of primed syllables, suggests a great deal of sensitivity of the language system to the lexical prosody of an English word. In fact, although indirect, several lines of behavioral research have provided converging evidence underscoring the importance of lexical prosody in both speech comprehension and production (Cutler, 1992; Levelt, 1989, 1992; McQueen, Norris, & Cutler., 1994; Norris, McQueen, & Cutler.,1995; Windfield, Goodglass, & Lindfield, 1997).

For example, it has been argued that the presentation of a spoken word with a specific stress pattern activates only those candidates in the lexicon that bear this stress pattern (Connine, Clifton, & Cutler, 1987). In an attempt to examine the viability of this argument, Lindfield and colleagues (1999) employed a word-onset gating technique and required English listeners to identify words according to word onsets alone, word onsets with word duration information, or word onsets with full word prosody (both duration and stress pattern) cues. It was found that when lexical prosody was present, word recognition was enhanced with significantly less phonological onset information. In contrast, the presence of only the duration of the word excluding stress pattern provided no better help than the simple presence of word onset phonological information. The authors concluded that lexical prosody must exert its effects alongside segmental information on the lexical constraints during spoken word recognition.

Taken together, it seems reasonable to claim that word prosody is represented in the mental lexicon and plays an important role in lexical constraints. In line with this argument, the role of stress information in ongoing English speech processing further brings another interesting issue to the researcher's attention. Is the role of word prosody in spoken English word processing universal or specific to stress-based languages? Unfortunately, due to our limited knowledge, it still remains an open question. In an attempt to shed more light on the nature of word prosodic processing, the second experiment was conducted to investigate whether word prosody (lexical tone) in tone-based languages (e.g., Chinese) plays any role in spoken word recognition; if it does, to what extent do tones contribute to spoken Chinese character processing. In addition, the second experiment was designed to unravel the nature of segmental (onset and rime) processing as well as the relationship between word prosody and segmental phonology during ongoing linguistic processing.

Chapter Three:

Experiment 2: The influence of lexical tone, segmental phonology, and meaning in the comprehension of spoken Chinese Mandarin: an ERP investigation

3.1 Objectives:

The main intentions of this study were to arrive at a clearer understanding of the neuro-physiological bases of the processing of spoken Chinese Mandarin and to help build a better model of prosodic and segmental phonological processing in Chinese. In particular, the present experiment was critically aimed to examine potential effects of lexical tones, onsets, rimes, and meaning on the recognition of spoken Chinese characters and determine the relationship between tone, onset and rime processes.

In order to elucidate the interactive locus of segmental and tonal processes and the time course of spoken Chinese character processing, experimental methods and designs need to be sensitive to both temporal and spatial domains. In this regard, the ERP method was thereby used. The correlated components of ERPs (the PMN and the N400 in particular) were studied to define the roles of segments and tones in spoken Chinese Mandarin processing.

Four-character Chinese proverbs were chosen as the experimental stimuli. Chinese proverbs are popular, stereotyped phrases of long-standing use characterized by simplified but strict structural form and penetrating meaning which produce high-constraint contexts. For example, “塞翁失马”, whose literal meaning is "the old frontiersman losing his mare—a blessing in disguise" [1]. None of the four characters could be replaced -- even with a synonym -- and the order of the characters is strictly fixed.

---

[1] This proverb is from the parable in the Huai Nan Zi about the old frontiersman whose strayed horse returned to its master accompanied by a better horse.

With these characteristics and their highly constrained contextual features, Chinese four-character proverbs and the systematic manipulation of Chinese proverbs provide an ideal method of isolating different component features in neural responses to language comprehension. In addition, a Chinese character generally has a simple monosyllabic structure that consists of an initial consonant or consonant cluster (onset) followed by either a simple vowel or by a diphthong or vowel combination (rime). A vowel followed by a nasal consonant is commonly represented by one symbol in Chinese phonetic presentation. Meanwhile, Chinese is a tonal language. Onset and rime information does not provide adequate information to identify concrete Chinese characters unless tonal cues are simultaneously provided. This study manipulated the dimensions (e.g., tone, rime, onset, & meaning) of proverb-ending characters and hereby examined the ERP manifestation, in an attempt to further understand the nature of Chinese Mandarin language processing. See Table 2 for examples of Chinese four-character proverbs from the stimulus conditions.

As seen in the prior English experiment in which the decision-making was required, the late positive component (LPC) that is tightly associated with post-lexical decision related processes (Kutas & Hillyard, 1989; Kutas & Van Petten, 1988) was found to a large degree overlapping with the N400 component which is indeed relevant to the addressed issues of interest. It is acknowledged that there is an avoidable problem in determining variations of the descending leg of the N400 component according to experimental manipulations when there is overlap with the LPC component. According to the pilot data in the present experiment, however, it was found that this confound (the

Table 2

Experimental Design used in Experiment 2 &3.

Stimulus Conditions	Proverb Example	Phonetic forms	+ semantic meanings of ending characters	number of Trials
Congruent	入乡随俗	/su2/	custom	100
Onset Change	鸡飞蛋塔(**打**)	/ta3/ (**/da3/**)	tower (**break**)	40
Rime Change	爱屋及挖(**乌**)	/wa1/ (**/wu1/**)	dig (**black**)	40
Tone Change	不期而鱼(**遇**)	/yu2/ (**/yu4/**)	fish (**meet**)	40
Incongruent	引狼入毡(**室**)	/zhan1/ (**/shi4/**)	blanket (**room**)	40

Note: Words with boldface type were correct proverb ending characters and were not presented in the experiment.

A total of 260 trials were either binaurally presented in Experiment 2 or visually presented in Experiment 3.

LPC) can be more likely to be eliminated or drastically reduced if the experimental task demand steers away from any requirement of decision-related mental processes. In this view, given that the issues addressed in the present experiment were mainly focused on mental processes at the lexical other than post-lexical stage, I used a simple Chinese proverb comprehension task in which native Chinese participants were asked to carefully listen to aurally presented proverbs; and, when the fourth character was presented, they were required to think of its meaning as accurately and concretely as they could.

3.2 Hypotheses:

Several possible patterns of results might emerge according to prior behavioral research. Based on the argument that lexical tones are represented in the mental lexicon, it was hypothesized that the N400 component would vary according to tonal manipulations. Thus, ERP results were expected to reflect that the N400 in response to the **Tone** change target character (which had the same onset and rime as, but different tone and meaning from, the correct proverb-ending character) would appear larger than that to the **Congruent** character (the correct fourth character of a proverb). Moreover, based on the performance data that there is right ear advantage in lexical tone identification and the English ERP literature that the speech N400 presents left hemisphere dominance (e.g., Connolly & Phillips, 1994), it was further hypothesized that the N400 in the **Tone** condition would have a left hemisphere distribution.

To challenge the argument that tonal information has to be translated into vowels and become part of vowels' phonological configuration at the stage of phonological representation, it was hypothesized that the PMN component that reflects phonological representation in language processing would present either a different time course or scalp distribution, or both, between the **Tone** condition and the **Rime** condition (in which the target character had the same tone and onset as, but different rime and meaning from, the correct proverb-ending character). Meanwhile, it was also hypothesized that sub-lexical segmental phonology has a facilitation effect on monosyllabic Chinese character recognition. Confirmation of this hypothesis would be demonstrated if the N400 components in response to both **Rime** change and **Onset** change (which had the same rime and tone as, but different onset and meaning from, the correct proverb-ending character) targets appeared smaller than that to the **Incongruent** target (whose tone, onset, rime, and semantic were inappropriate to the proverb context) but larger than that to the **Congruent** target.

3.3 Method:

3.3.1 Participants

Participants (6M/9F) were native speakers of Mandarin Chinese studying at Dalhousie University as international students. All of the participants were right-handed with no left-handed relatives in their immediate families, had normal or corrected-to-normal vision, reported no history of hearing or speech disorders, and were 19-28 years of age ($\underline{M}$ = 23.5). Informed consent was obtained from all participants and they were financially reimbursed for their participation.

3.3.2 Stimuli and experimental conditions

The stimuli in this study were 260 Chinese four-character proverbs that were selected from the Little Dictionary of Chinese Proverbs, which was compiled for elementary, and middle school students in the mainland of China. All stimuli were recorded by a female native speaker of Chinese Mandarin. Stimuli were digitized at 10 KHz (12 bit resolution) and timing marks (inaudible to participants) were placed at the onset of the proverb-ending character. To identify the word onset, the waveform of the terminal portion of each proverb was viewed on a computer screen using successive windowing until the onset of the terminal character had been identified. Visually identified word onsets were confirmed by repeated auditory presentation of the windowed area around word onset. The timing marks were used to initiate sampling of the electroencephalographic activity (EEG).

The five experimental conditions used proverbs in which the terminal character varied according to onset, rime, tone, and/or semantic appropriateness (See Appendix C for proverb stimuli). In the **Congruent** condition, the ending character of each proverb was the correct fourth character of the proverb. For example, “入乡随俗/su2/” , in which the terminal character ‘俗’ means custom. The literal meaning of this proverb is ‘Wherever you are, follow local customs’ which is similar to the proverb ‘When in Rome, do as the Romans do.’ In the **Onset** condition, the terminal character of each proverb had the same rime and tone as, but a different onset and meaning from, the correct fourth character of the proverb, for instance, ‘鸡飞蛋塔/ta3/’, in which the terminal character should have been ‘打/da3/’ ‘break’. However, the character ‘塔’ used means ‘tower’, which does not semantically fit the context of the proverb whose literal meaning is ‘The hen has flown away and the eggs in the coop are broken’, comparable to the English proverb ‘Come out empty-handed.’ In the **Rime** condition, each proverb ended with a character that shared the same onset and tone as, but different rime and meaning from, the correct terminal character of the proverb. For example, “爱屋及挖/wa1/”, in which the terminal character should have been ‘乌/wu1/’ ‘black (crow)’ but was, in fact, ‘挖’ ‘dig’. The literal meaning of the proverb is ‘Love for a person extends even to the black crows on his roof’ similar to ‘Love me, love my dog.’ In the **Tone** condition, the terminal character of the proverb had the same onset and rime as, but different tone and meaning from, the correct fourth character of the proverb. Thus, ‘不期而鱼/yu2/’, is a proverb in which the terminal character should have been’遇/yu4/’ ‘meet’ but was in fact, ‘鱼/yu2/’ ‘fish’ and has a literal meaning of ‘Meet unexpectedly or have a chance encounter’ similar to ‘Bump into.’ In the **Incongruent** condition, the onset, rime, tone, and meaning of the terminal character were completely different from those of the correct fourth character of the proverb. For example, “引狼入毡/zhan1/” should have ended with the character ‘室/shi4/’ meaning ‘room’ but in fact ended with the character ‘毡’ meaning ‘blanket’ The proverb’s literal meaning is ‘Invite a wolf into the house’ which means ‘Set the wolf to keep the sheep.’. 160 fillers were proverbs ending with correct fourth characters. The conditions were presented pseudo-randomly with the restriction that no condition was presented successively more than three times.

3.3.3 Procedure

Participants were seated in a comfortable, padded chair in an electrically isolated and sound-attenuated room adjoining a room containing the recording equipment.

Participants were instructed to attend to the proverbs and to avoid blinking during proverb presentation. When the fourth character (target) was presented, they were asked to think of its meaning as accurately and clearly as they could. Participants were told that they might be asked some questions about the stimuli at the end of the experiment. Participants were given time to practice until they fully understood the task demands of the experiment. Frequent breaks were taken to reduce participant fatigue.

3.3.4 EEG Recording

The electroencephalogram (EEG) was recorded from 17 sites using tin electrodes with linked ears as the reference and a mid-forehead site as ground. Recording locations were located on the basis of the standard international 10-20 system (Jasper, 1958) at Fz, F3, F4, F7, F8, Cz, C3, C4, Pz, P3, P4, T3, T4, T5, T6, O1, and O2. The electro-oculogram (EOG) recorded vertical eye movements with electrodes placed above and below the right eye. Horizontal eye movements were recorded with electrodes placed over the outer canthi of the right and left eye. Electrode impedance was kept below 5 K Ω. The analogue EEG recordings were obtained with a half-amplitude bandpass of 0.01-100 Hz (digitally sampled at 500 Hz, low pass filtered off-line at 30 Hz and high pass filtered off-line at 0.1 Hz). The sample duration began 100 ms before the stimulus onset and continued until 1000 ms after stimulus onset. The stimulus onset was defined as the beginning of the terminal character of the proverb. Trials contaminated by EOG greater than ±75 μv or any other artifacts (e.g., muscle activity and frequent α waves) were excluded from the analysis. The remaining EEG trials were then averaged by experimental condition. The individual average ERPs were also averaged together to create grand average waveforms for the various conditions.

3.3.5 Data Analysis

ERPs that were collected from each electrode site for each individual were obtained by averaging the EEG activity that was recorded to the proverb-ending characters within each stimulus condition. The ERPs for each individual were then averaged together to obtain grand average waveforms that reflected the general ERP pattern for the entire group of participants in the different stimulus conditions. Data were analyzed several ways in order to obtain as good a view as possible of the negative components observed in the present experiment in the 200-550 ms time range. The main assessment approach used the interval scoring method by taking the mean amplitudes of 50 ms epochs successively from 200-350 ms after the stimulus onset for the PMN component and from 350-550 ms post stimulus onset for the N400 component. In addition, in order to better isolate the variations in the PMN peak latency across varied conditions, the peak latency data for the PMN component were obtained using the peak identification method. The PMN was scored as the most negative point in the 200-350 ms time range. Peak latency was defined as the time from stimulus onset to the point scored as the most negative within the latency range.

The statistical analyses of the ERP data were conducted with a repeated measures analysis of variance (ANOVA) and Greenhouse-Geisser corrections to the degrees of freedom were used when appropriate (Greenhouse & Geisser, 1959). As for the first time interval approach: in the first analysis, the analyses of the PMN and N400 responses included the Condition factor (5 levels: **Incongruent**, **Onset**, **Rime**, **Tone**, and **Incongruent**), a Time factor (for the PMN, the time intervals were 200-250, 250-300, 300-350 ms, and for the N400, the time intervals were 350-400, 400-450, 450-500, 500-550ms), and a Site factor (17 levels). The second analysis was conducted with Condition (5 levels) as one factor and the recording sites divided into two factors: Hemisphere (left, right) and Region (5 levels: frontal, central, parietal, temporal and occipital). The frontal level combined F3 and F7 for a left hemisphere frontal value and F4 and F8 for the corresponding right hemisphere value. A similar approach was conducted with temporal sites combining T3 and T5 for the left and T4 and T6 for the right. The central level consisted of C3 and C4 values while the parietal level consisted of the P3 and P4 values. The Time x Condition x Region x Hemisphere analysis was done as a way of simplifying the assessment of scalp topography differences. Results from the Time x Condition x Site analyses are presented first and those from the Time x Condition x Region x Hemisphere analyses are presented only if they provide new information or clarification of previous results. As for the peak scoring method for examining PMN latency, the first analysis of the data included Condition (5 levels) and Site (17 levels) as factors in the ANOVA. The second analysis was conducted with condition (5 levels) as one factor and the recoding sites divided into Hemisphere (2 levels) and Region (5 levels) two factors. This approach facilitated the assessment of topographical differences by investigating the PMN latency data associated with specific brain areas and hemispheres.

Significant main effects and interactions were submitted to further post hoc analyses using the Tukey Honestly Significant difference (HSD) test. In addition, for data involved in interactions between conditions and sites or regions and hemispheres, a normalization procedure was conducted to equalize amplitudes across conditions so that interactions between conditions and scalp distribution would reflect real topographical differences amongst conditions and not an artifact of absolute amplitude differences amongst conditions (McCarthy & Wood, 1985). An alpha level of $p < .05$ was required for statistical significance.

3.4 Results

3.4.1 Waveforms

Observation of the grand average waveforms (Figure 13A) reveals clear differentiation amongst the five experimental conditions. The largest amplitude N400 is seen in the **Incongruent** condition (i.e., in which onset, rime, tone and semantic features of the final character were incongruous to the proverb context) while the characteristic late positive component is observable in the **Congruent** condition (i.e., in which the correct ending was presented). The **Onset** (i.e., in which the terminal character had the incorrect onset and semantic features) and **Rime** (i.e., in which the terminal character had the incorrect rime and semantic features) conditions elicited a clear N400 whose amplitude is larger than that observed in the **Tone** condition (i.e., in which the tonal and semantic features of the ending character were incorrect with regard to the sentence context) but smaller than that observed in the **Incongruent** condition. In addition, the morphology of the putative-N400 in **Onset**, **Rime**, and **Tone** conditions is not only characterized by lower amplitude but by a wider, less pronounced shape strongly suggestive of the presence of other condition-related ERP components. The PMN can be

Figure Caption

Figure 13A. Grand average waveforms at 17 recording sites from 15 individuals for the **Congruent**, **Onset**, **Rime**, **Tone**, and **Incongruent** stimulus conditions presented in Experiment 2.

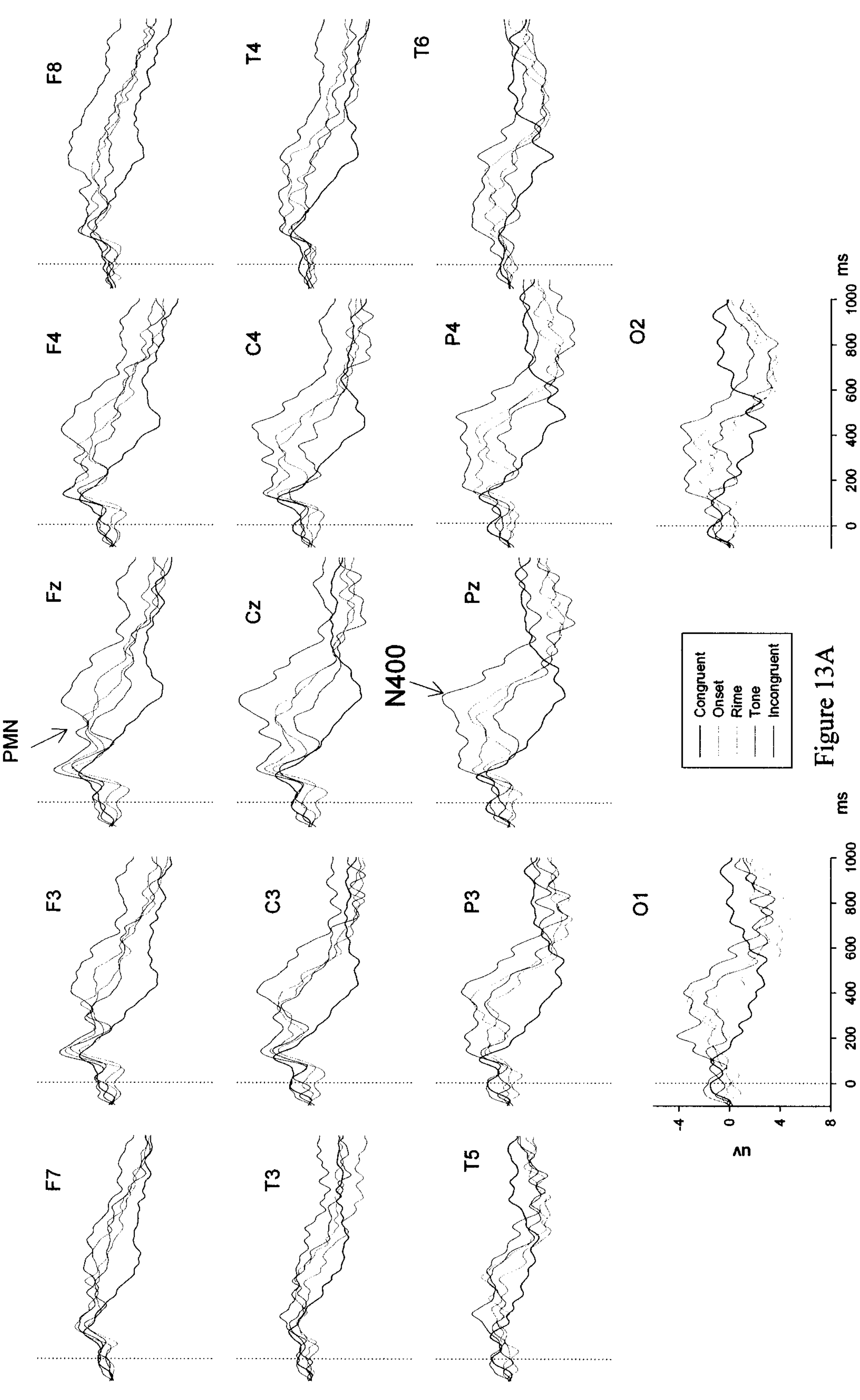
F7
F3
Fz
F4
F8
T3
C3
Cz
C4
T4
T5
P3
Pz
P4
T6
O1
O2
PMN
N400
uV
-4
0
4
8
0
200
400
600
800
1000
ms
Congruent
Onset
Rime
Tone
Incongruent

Figure 13A

Figure Caption

Figure 13B. Grand average waveforms at Fz from 15 individuals for the **Congruent**, **Onset**, **Rime**, **Tone**, and **Incongruent** stimulus conditions presented in Experiment 2.

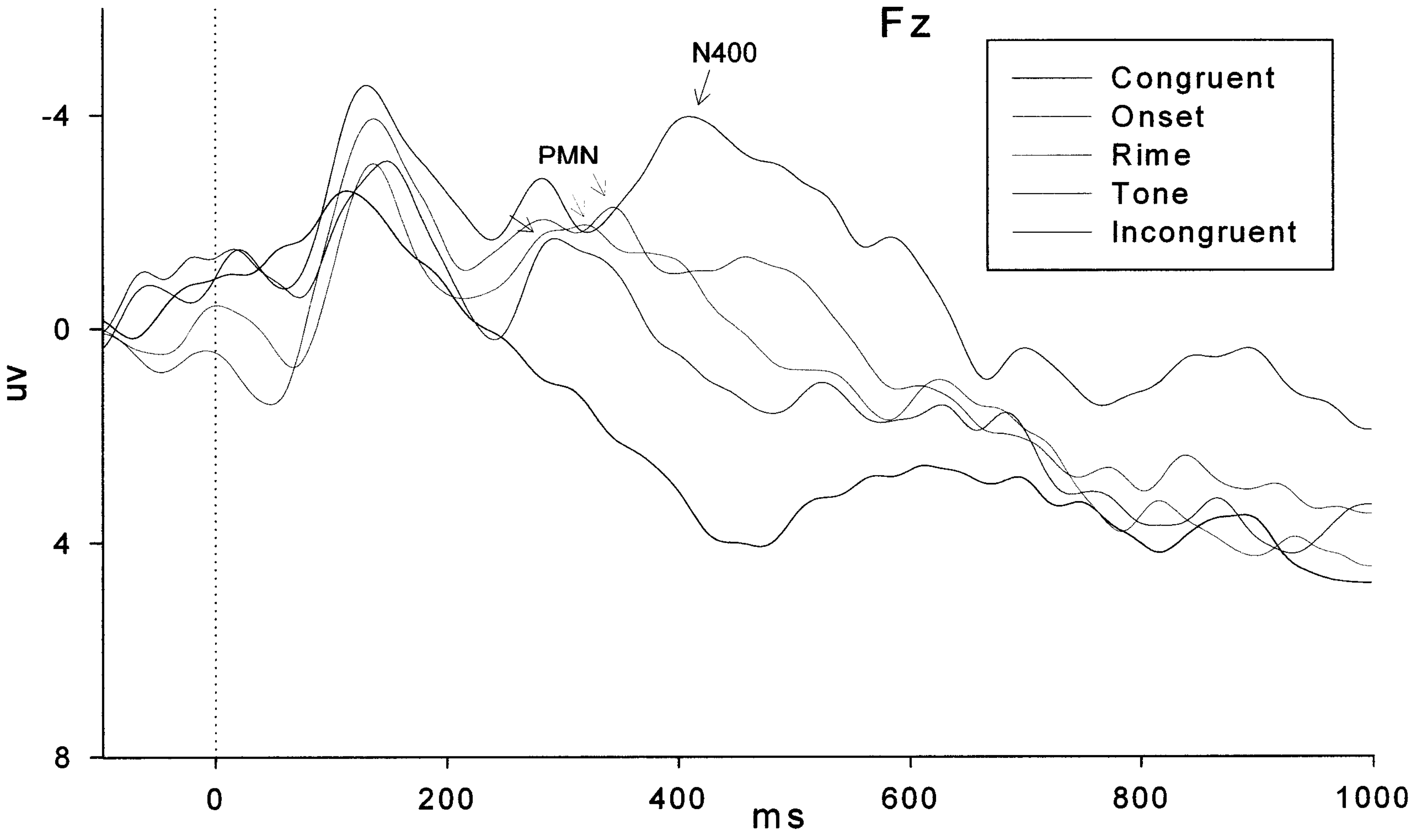
Fz
N400
PMN
Congruent
Onset
Rime
Tone
Incongruent
-4
0
4
8
uv
0
200
400
600
800
1000
ms

Figure 13B

Figure Caption

Figure 13C. Grand average waveforms at F3 from 15 individuals for the **Congruent**, **Onset**, **Rime**, **Tone**, and **Incongruent** stimulus conditions presented in Experiment 2.

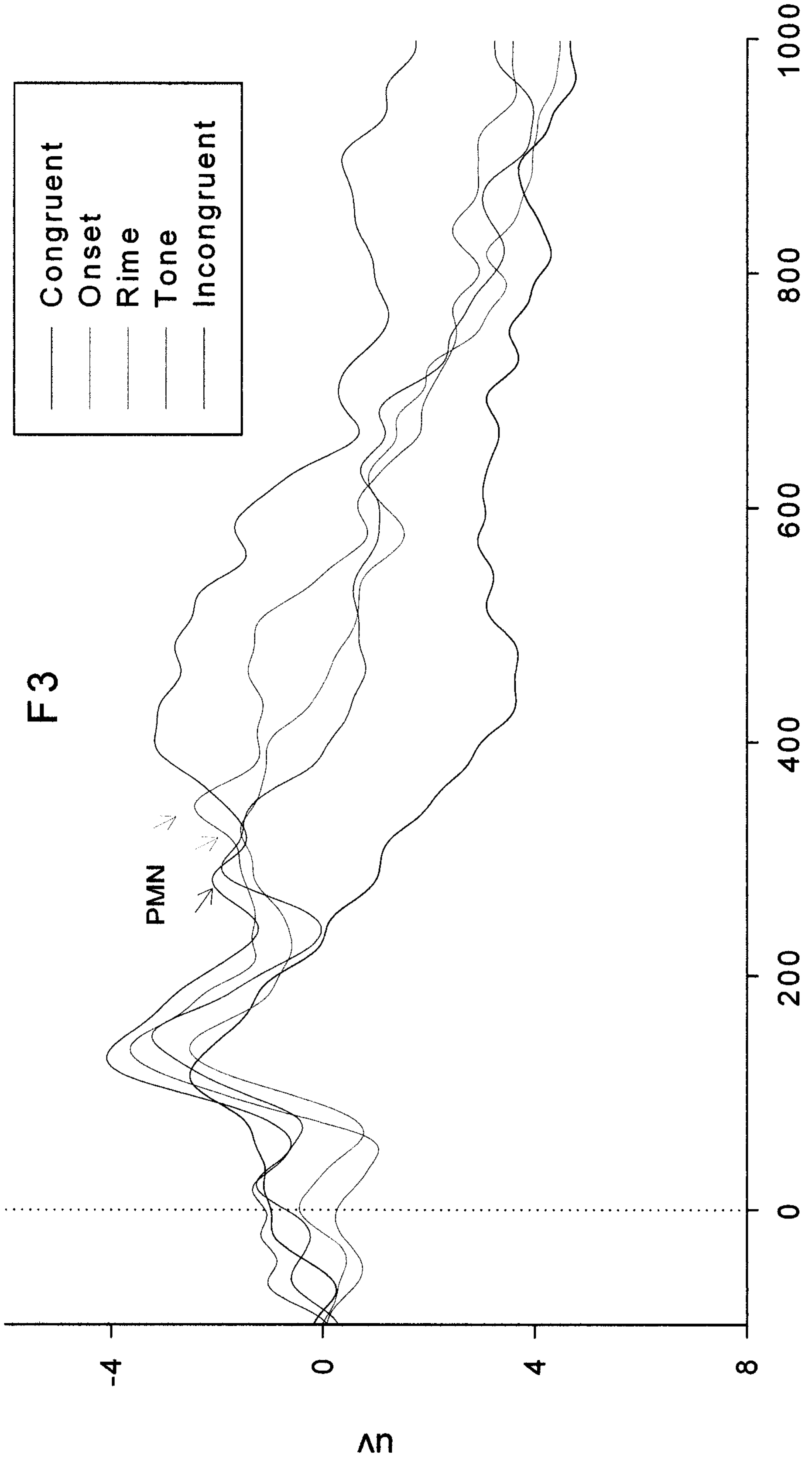
F3
PMN
Congruent
Onset
Rime
Tone
Incongruent
-4
0
4
8
uV
0
200
400
600
800
1000
ms

Figure 13 C

observed in the **Incongruent** and **Onset** conditions most clearly but is also discernable in the **Rime** and **Tone** conditions. The PMN appears to peak differently in between the **Onset, Tone** and **Rime** conditions. This phenomenon seems more evident over frontal recording sites (Figures 13B, C). It appears that the PMN in the **Tone** condition peaks early relative to those in the other two conditions; and, the peak of the PMN in the **Rime** condition seems earlier that that in the **Onset** condition. These phenomena observed in the grand average waveforms are identifiable in each individual participant ERP waveforms. Figure 24 displays the waveforms of a representative participant.

3.4.2 Statistical analyses

Phonological Mismatch Negativity (PMN)

The interval scoring analyses:

Analysis indicated the sensitivity of the PMN to the different manipulations of the phonological units (onset, rime and tone) (See Appendix D Table D1). A main effect of condition ($F(4, 56) = 18.38$, $p < 0.0001$, $\varepsilon = 0.61$) was found (See Figure 14). Subsequent post-hoc analyses indicated that PMN amplitudes to semantically **Incongruent** characters ($\underline{M} = -2.30$ μv, SE = 0.52) appeared larger than those in the **Tone** condition ($\underline{M} = 0.22$ μv, SE = 0.59) and the **Congruent** condition ($\underline{M} = 1.54$ μv, SE = 0.39) but showed no statistical difference from those in the **Onset** ($\underline{M} = -1.80$ μv, SE = 0.31) and **Rime** ($\underline{M} = -1.33$ μv, SE = 0.42) conditions. Moreover, PMN amplitudes in the **Tone** condition were statistically different from those in the **Congruent** condition. The Condition x Time interaction ($F(8, 112) = 4.04$, $p < 0.01$, $\varepsilon = 0.49$) and subsequent post-hoc analyses revealed that the highest amplitudes of the PMN to **Rime** change and **Tone** change

Figure Caption

Figure 14. Graph depicting the amplitude of the PMN response as a function of the stimulus condition in Experiment 2.

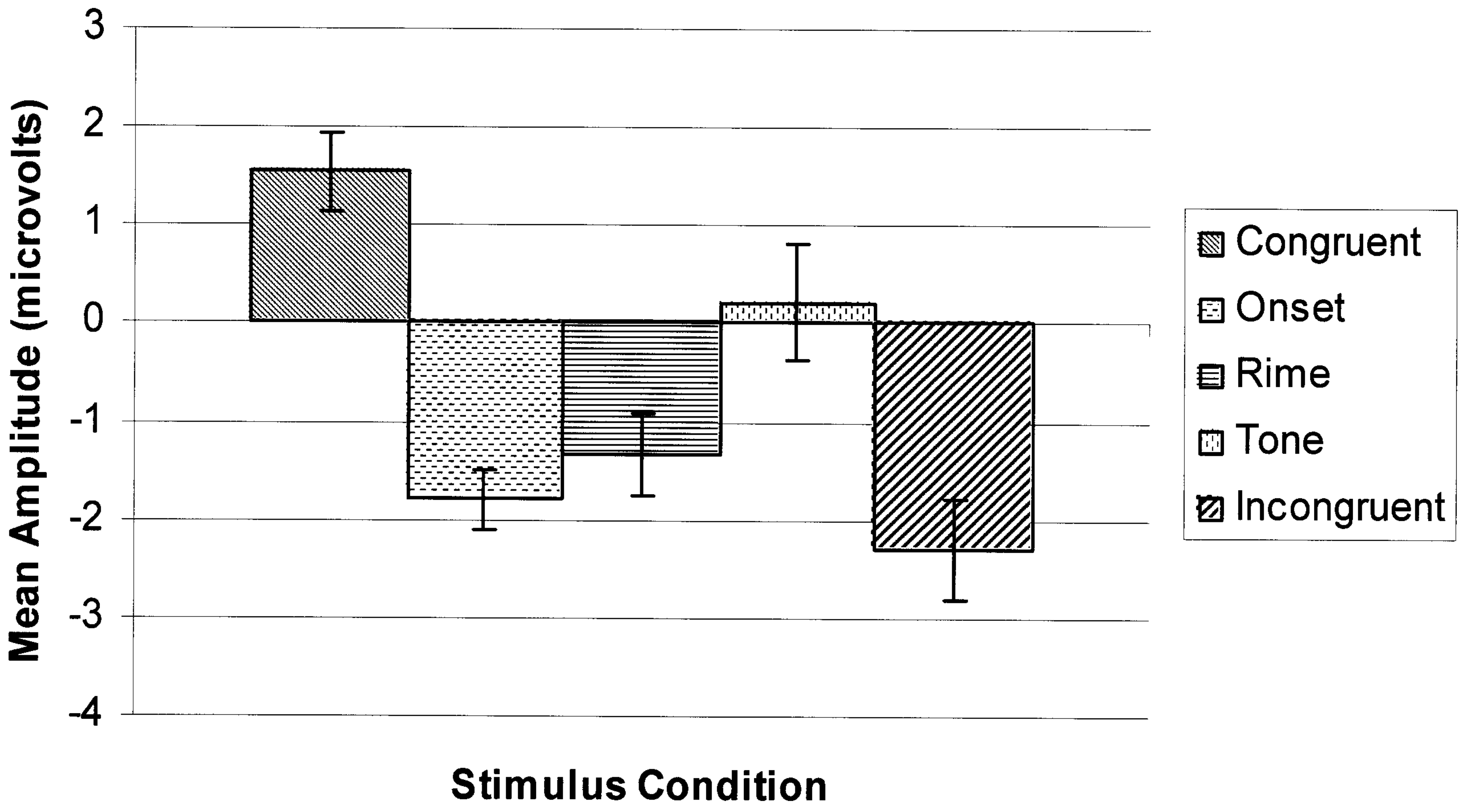
3
2
1
0
-1
-2
-3
-4
Mean Amplitude (microvolts)
Stimulus Condition
▨ Congruent
▤ Onset
▤ Rime
▦ Tone
▨ Incongruent

Figure 14

Figure Caption

Figure 15. Graph depicting the significant interaction between the Condition and Time factors for PMN amplitude data in Experiment 2.

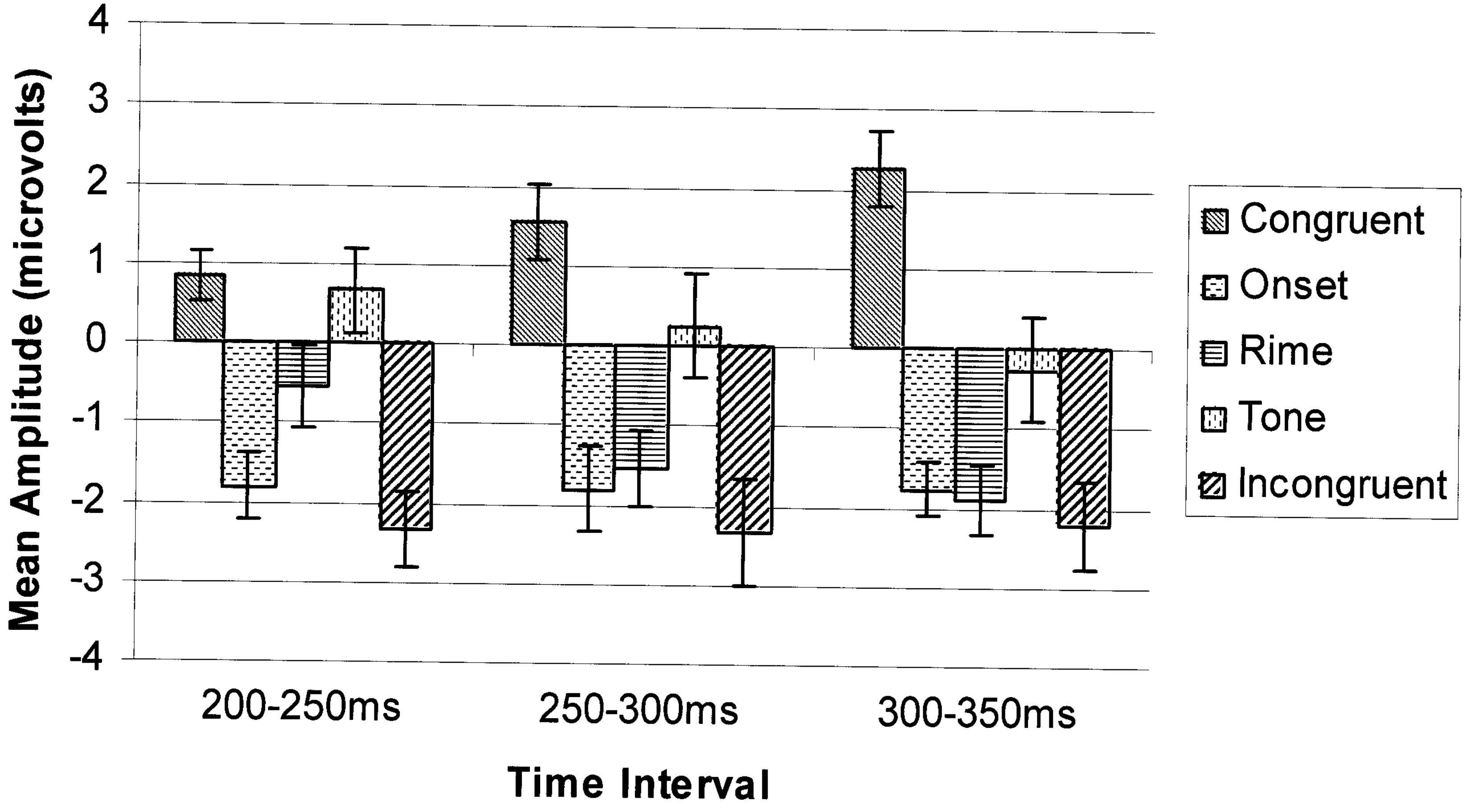
Mean Amplitude (microvolts)
4
3
2
1
0
-1
-2
-3
-4
200-250ms
250-300ms
300-350ms
Time Interval
▨ Congruent
▨ Onset
▤ Rime
▨ Tone
▨ Incongruent

Figure15

Figure Caption

Figure 16. Graph depicting the significant interaction between the Condition and Region factors for PMN amplitude data in Experiment 2.

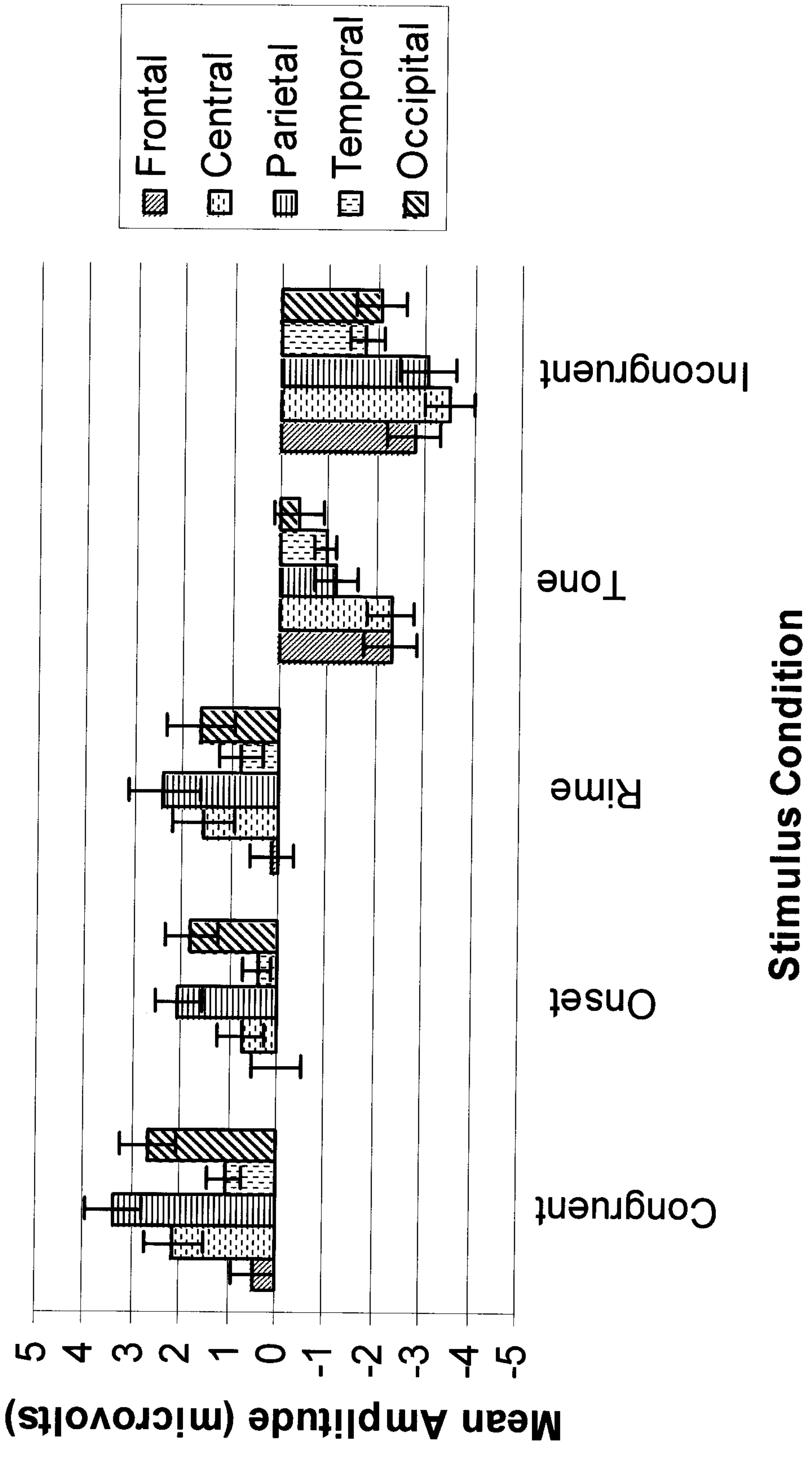
Mean Amplitude (microvolts)
5
4
3
2
1
0
-1
-2
-3
-4
-5
Congruent
Onset
Rime
Tone
Incongruent
Stimulus Condition
Frontal
Central
Parietal
Temporal
Occipital

Figure 16

characters appeared in the 250-350 ms time windows whereas there was no statistical difference in the mean amplitudes of the PMN among the 200-250 ms, 250-300 ms, and 300-350 ms three time intervals in the **Onset** and **Incongruent** two conditions (See Figure 15).

There were significant Condition x Site (F (64,896) = 3.32, $p<0.01$, $\varepsilon = 0.09$) and Time x Site (F (32,448) = 4.01, $p<0.01$, $\varepsilon = 0.10$) two-way interactions. Subsequent analyses indicated that the amplitudes in different sites varied across the conditions in each time interval (See Appendix D, Tables D2, D4; Figure 16). In the time window of 200-300 ms, there was a right fronto-central distribution in the **Onset** condition (Fig. 20) while the dominant distribution was present over frontal sites in the **Tone** condition (Fig. 22) and centro-parietal extending to right frontal sites in the **Incongruent** condition (Fig. 19). No significant scalp distribution effects were seen in the **Rime** condition although the amplitudes at the right temporal sites appeared higher than those over other sites (Fig. 21). In the time window of 300-350 ms, there was a central characteristic extending to frontal sites in the **Onset** condition. However, amplitudes in the **Incongruent** condition were larger over the centro-parietal midline than the frontal and temporal sites. There was a centro-parietal characteristic that extended to the occipital area in the **Rime** condition (Fig. 21). The dominant distribution was present over left central in the **Tone** condition (Fig. 22).

In the meanwhile, there was no main effect of hemisphere and no interactive effect of Time x Condition x Hemisphere or Condition x Hemisphere, which indicates there was no clear specific hemisphere dominance across five conditions (See Appendix D, Table D3).

Figure Caption

Figure 17. Graph depicting the peak latency of the PMN response as a function of the stimulus condition in Experiment 2.

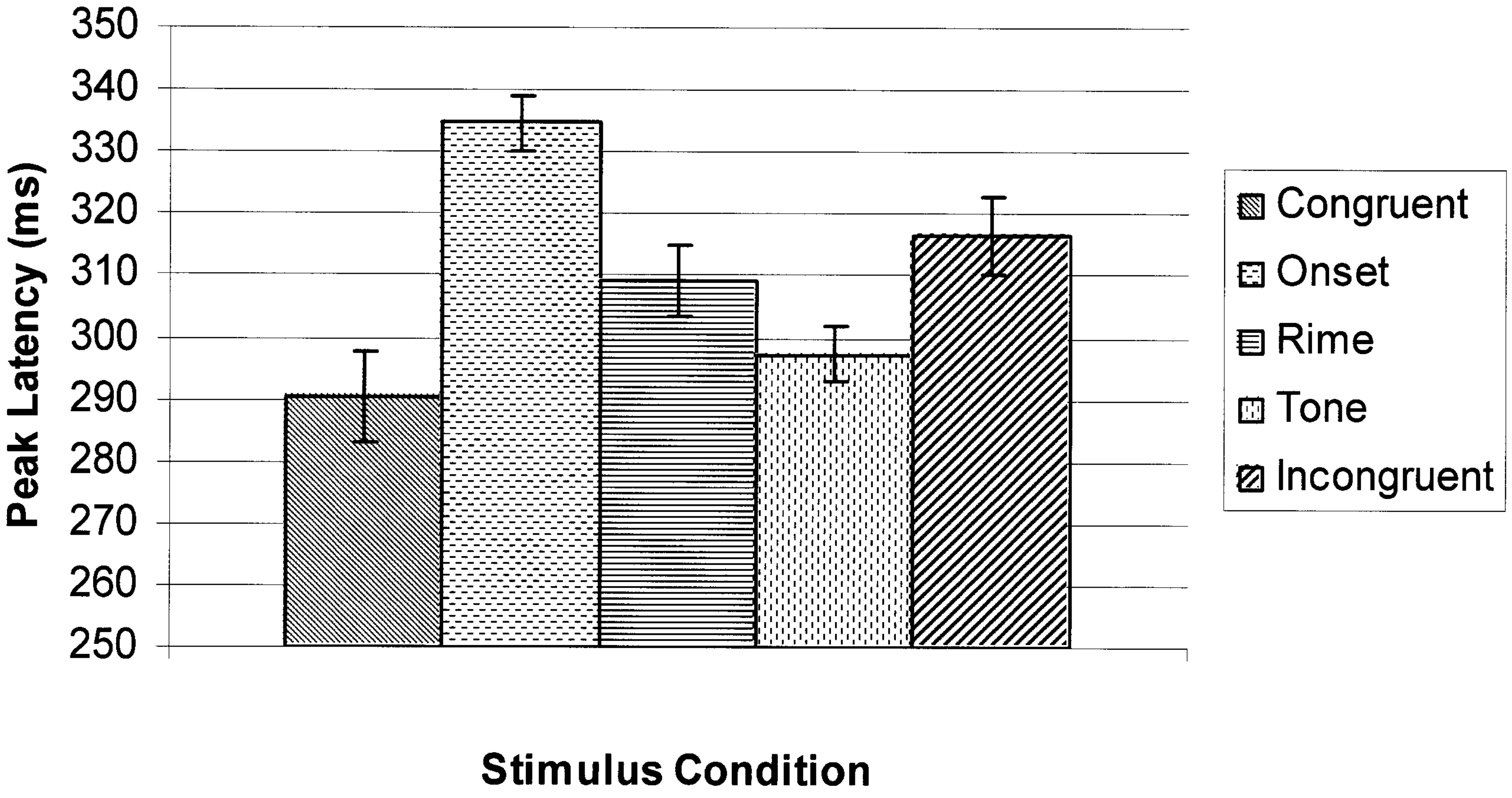
350
340
330
320
310
300
290
280
270
260
250
Peak Latency (ms)
Stimulus Condition
▨ Congruent
▤ Onset
▤ Rime
▥ Tone
▨ Incongruent

Figure 17

The peak scoring analyses:

The results of the peak scoring analysis showed that the peak latency of the PMN response varied as a function of the stimulus condition (F (4, 56) = 9.78, p<0.001, ε =0.77) (See Appendix D, Table D5). Subsequent post-hoc analyses revealed that the PMN occurred earlier to **Tone** characters (297ms, SE = 4.44) than to **Rime** (309ms, SE = 5.72) and **Incongruent** (316ms, SE = 6.49) characters which in turn appeared earlier than to **Onset** characters (335ms, SE = 4.56) (See Figure 17).

N400

Analysis confirmed the sensitivity of the N400 amplitude to the experimental conditions (F (4, 56) = 36.78, p <0.0001, ε =0.58) (See Appendix D, Table D6). The Condition x Time interaction (F (12, 168) = 3.74, p<0.02, ε = 0.30) and subsequent post-hoc analyses suggested that N400 amplitude to **Incongruent** characters was reliably larger than that in any of the other four conditions across the three time intervals. N400 amplitude to **Onset** characters (M = -0.57μv, SE = 0.46) appeared smaller than that to **Rime** characters (M = -1.32μv, SE = 0.49) in the 400-450 ms time interval. In the 500-550 ms time interval, the mean amplitudes in the **Onset** condition (M = 0.74μv, SE = 0.51) were larger (more negative) than those in the **Rime** condition (M = 1.03μv, SE = 0.47). The amplitudes in these two conditions were further larger than those in the **Tone** and **Congruent** conditions across three time intervals, between of which differences in the mean amplitudes were observed. The largest mean amplitudes in the **Incongruent** condition were observed in the 400-450 ms time interval while the mean amplitudes in the 400-450 ms and 450-500 ms time intervals (between which no difference in

Figure Caption

<u>Figure 18.</u> Graph depicting the significant interaction between the Condition and Time factors for N400 amplitude data in Experiment 2.

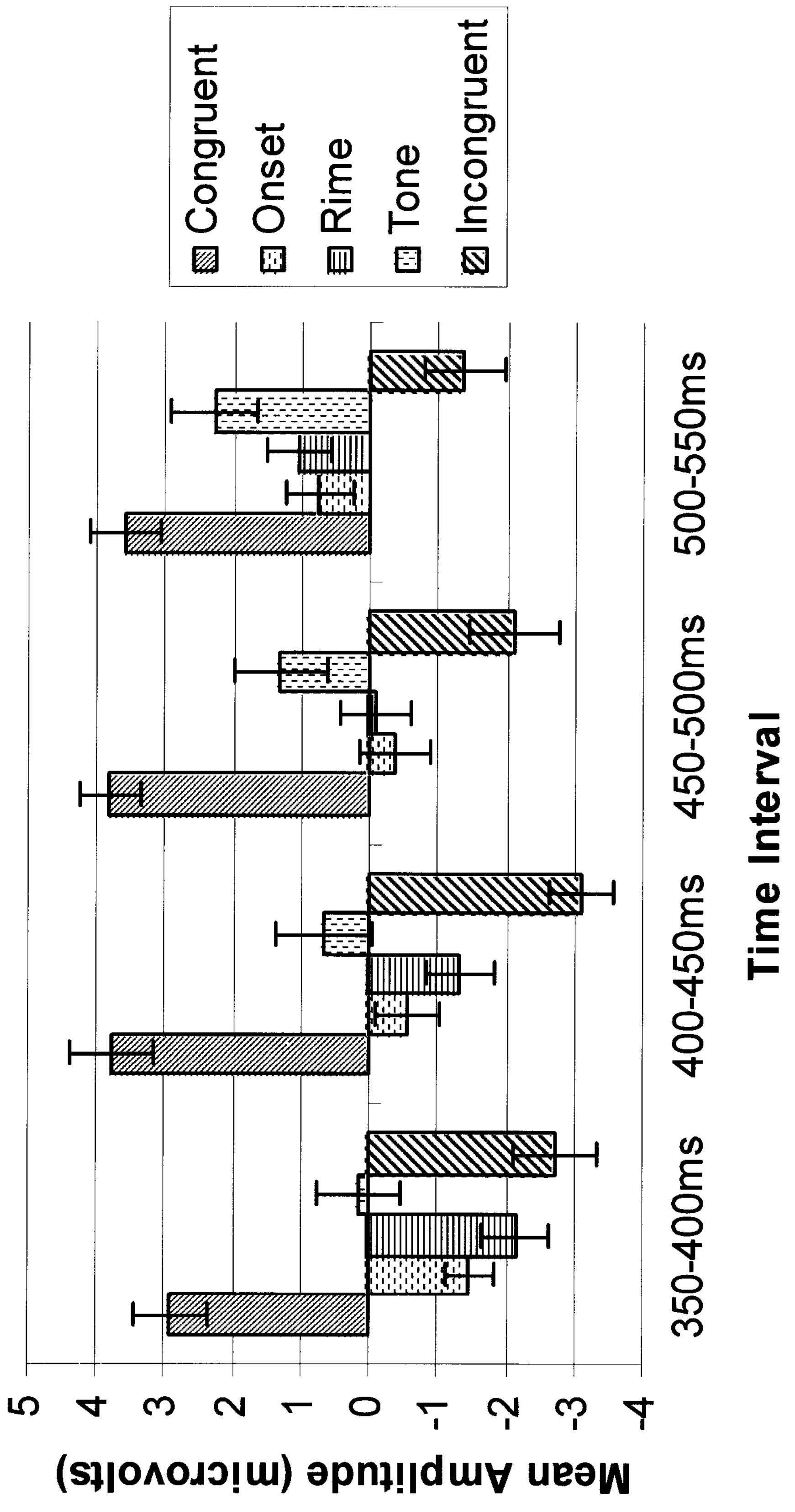
Mean Amplitude (microvolts)
5
4
3
2
1
0
-1
-2
-3
-4
350-400ms
400-450ms
450-500ms
500-550ms
Time Interval
▨ Congruent
▨ Onset
▥ Rime
▨ Tone
▨ Incongruent

Figure 18

Figure Caption

Figure 19. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Incongruent** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 2. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

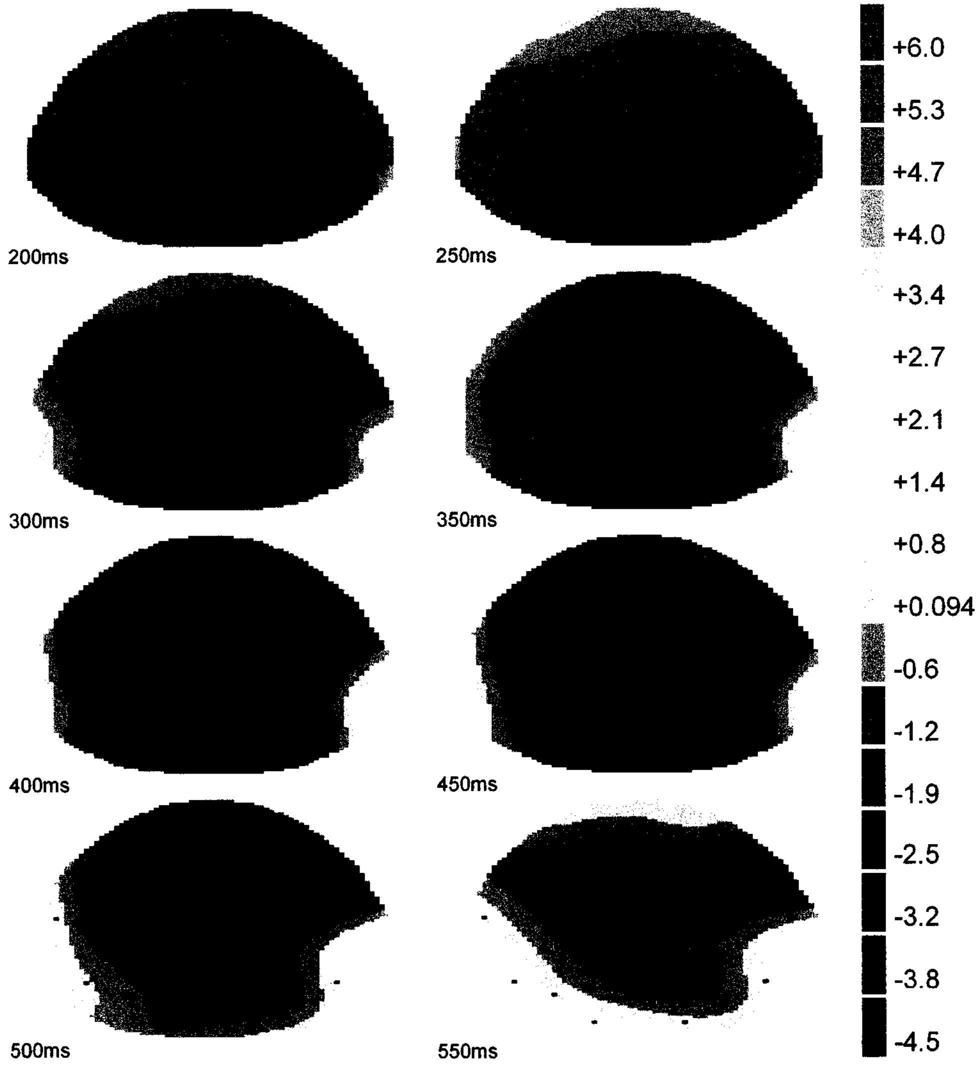
200ms
250ms
300ms
350ms
400ms
450ms
500ms
550ms
+6.0
+5.3
+4.7
+4.0
+3.4
+2.7
+2.1
+1.4
+0.8
+0.094
-0.6
-1.2
-1.9
-2.5
-3.2
-3.8
-4.5

Figure 19

Figure Caption

Figure 20. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Onset** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 2. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

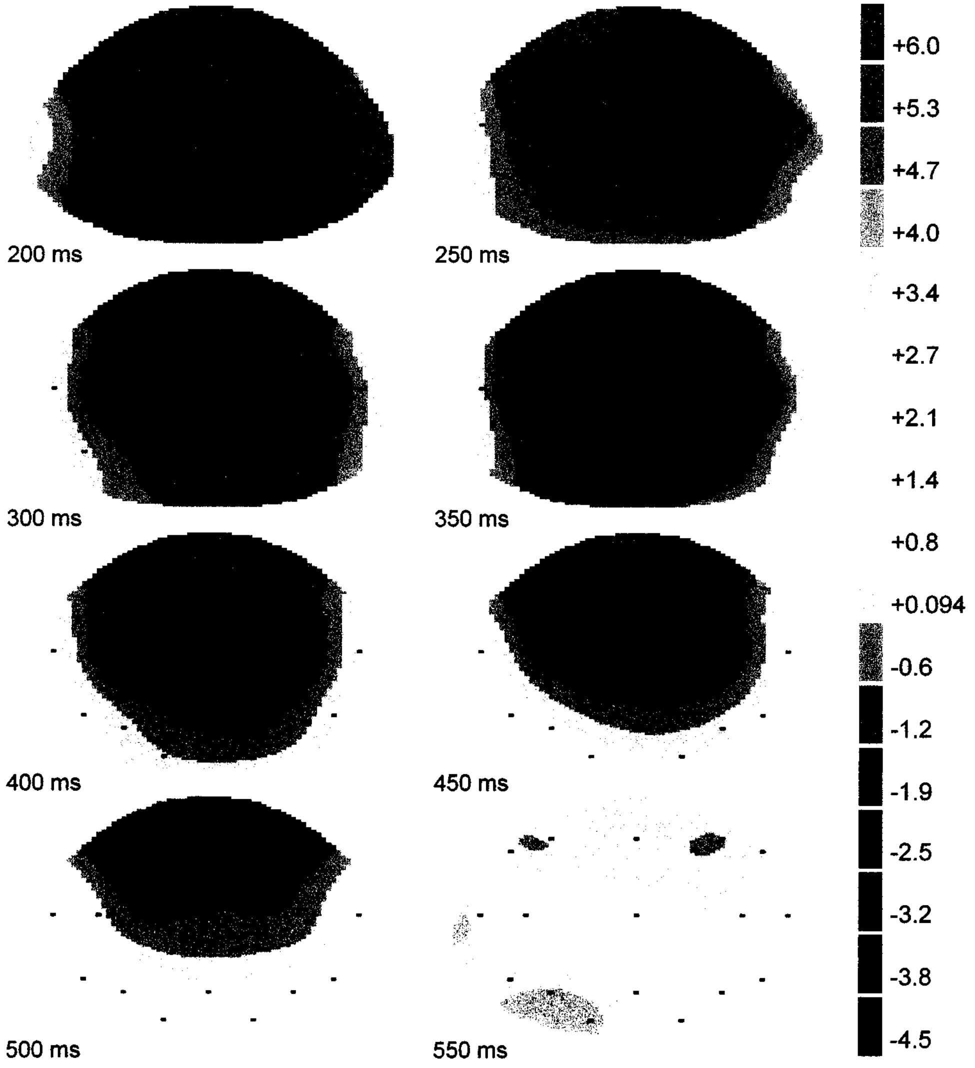
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.3
+4.7
+4.0
+3.4
+2.7
+2.1
+1.4
+0.8
+0.094
-0.6
-1.2
-1.9
-2.5
-3.2
-3.8
-4.5

Figure 20

Figure Caption

Figure 21. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Rime** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 2. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

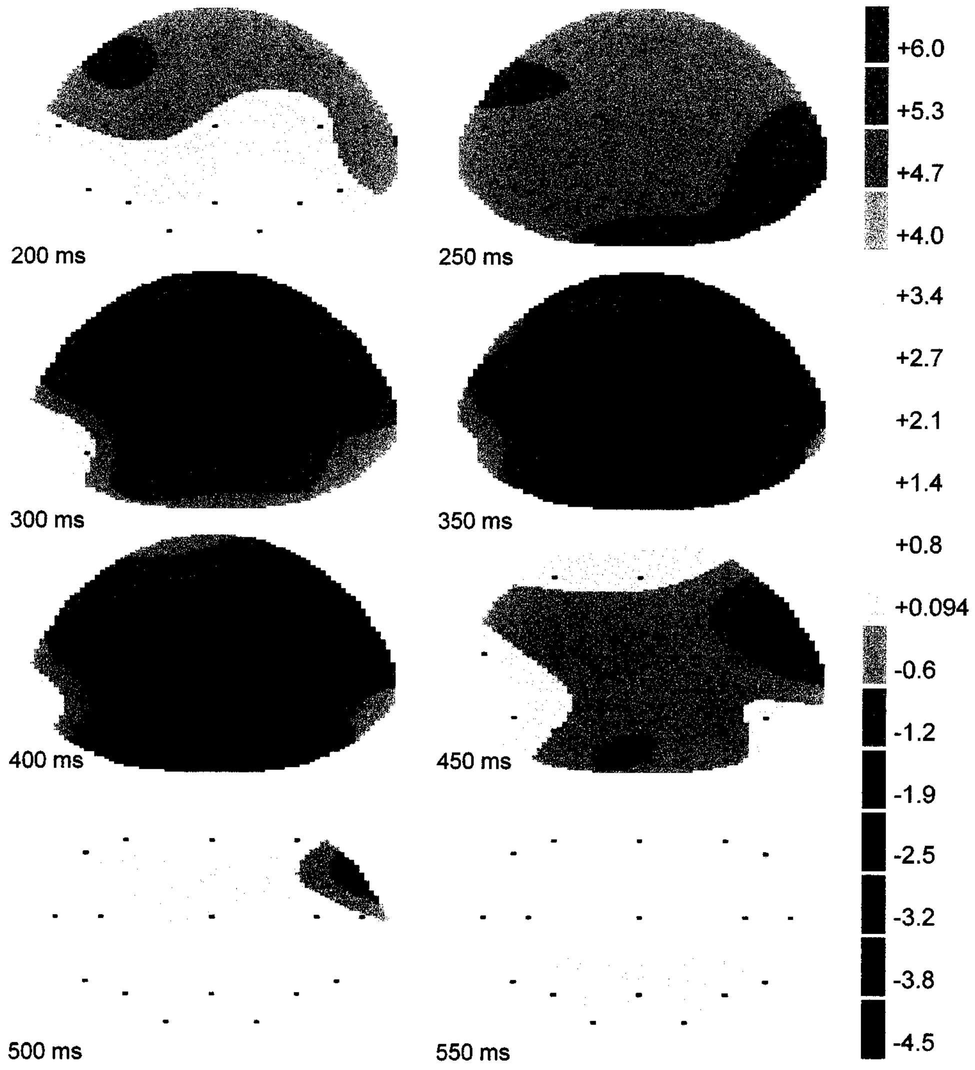
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.3
+4.7
+4.0
+3.4
+2.7
+2.1
+1.4
+0.8
+0.094
-0.6
-1.2
-1.9
-2.5
-3.2
-3.8
-4.5

Figure 21

Figure Caption

Figure 22. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Tone** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 2. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

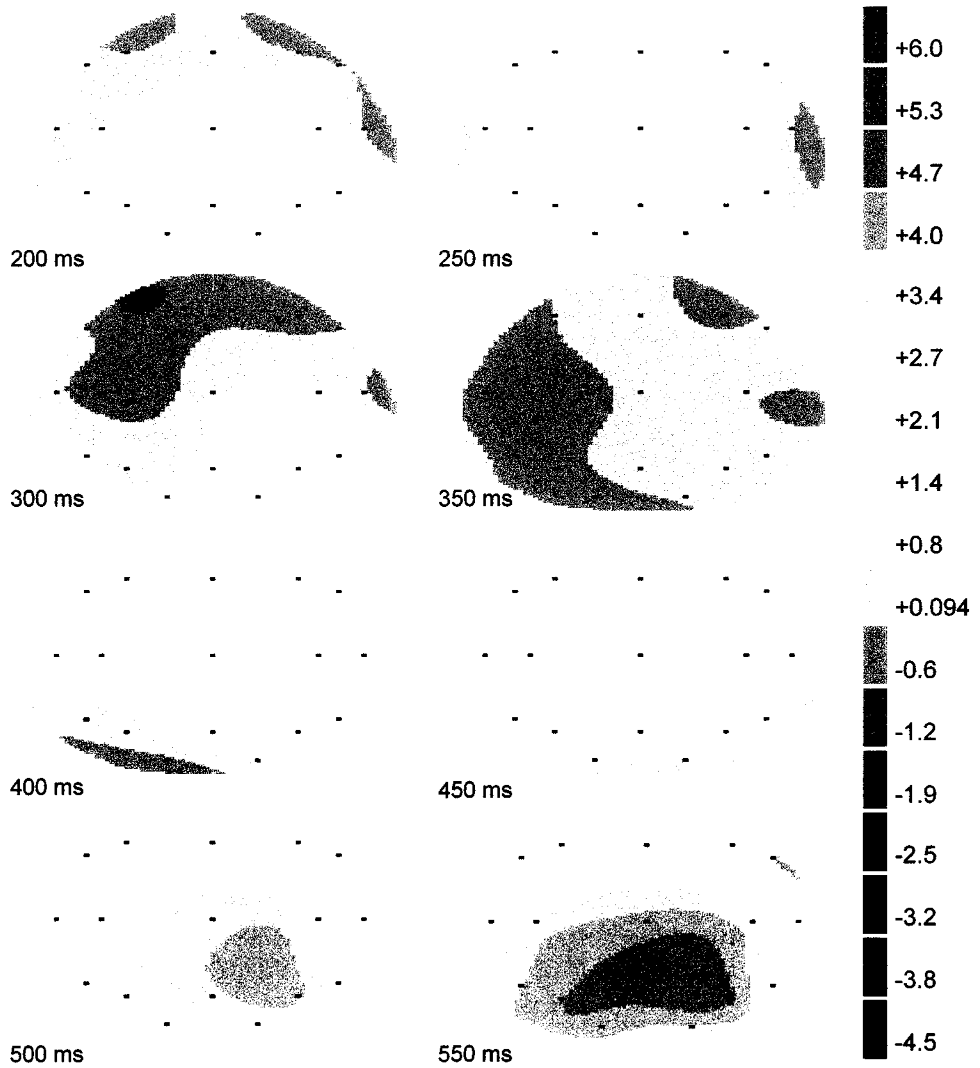
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.3
+4.7
+4.0
+3.4
+2.7
+2.1
+1.4
+0.8
+0.094
-0.6
-1.2
-1.9
-2.5
-3.2
-3.8
-4.5

Figure 22

Figure Caption

Figure 23. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Congruent** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 2. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

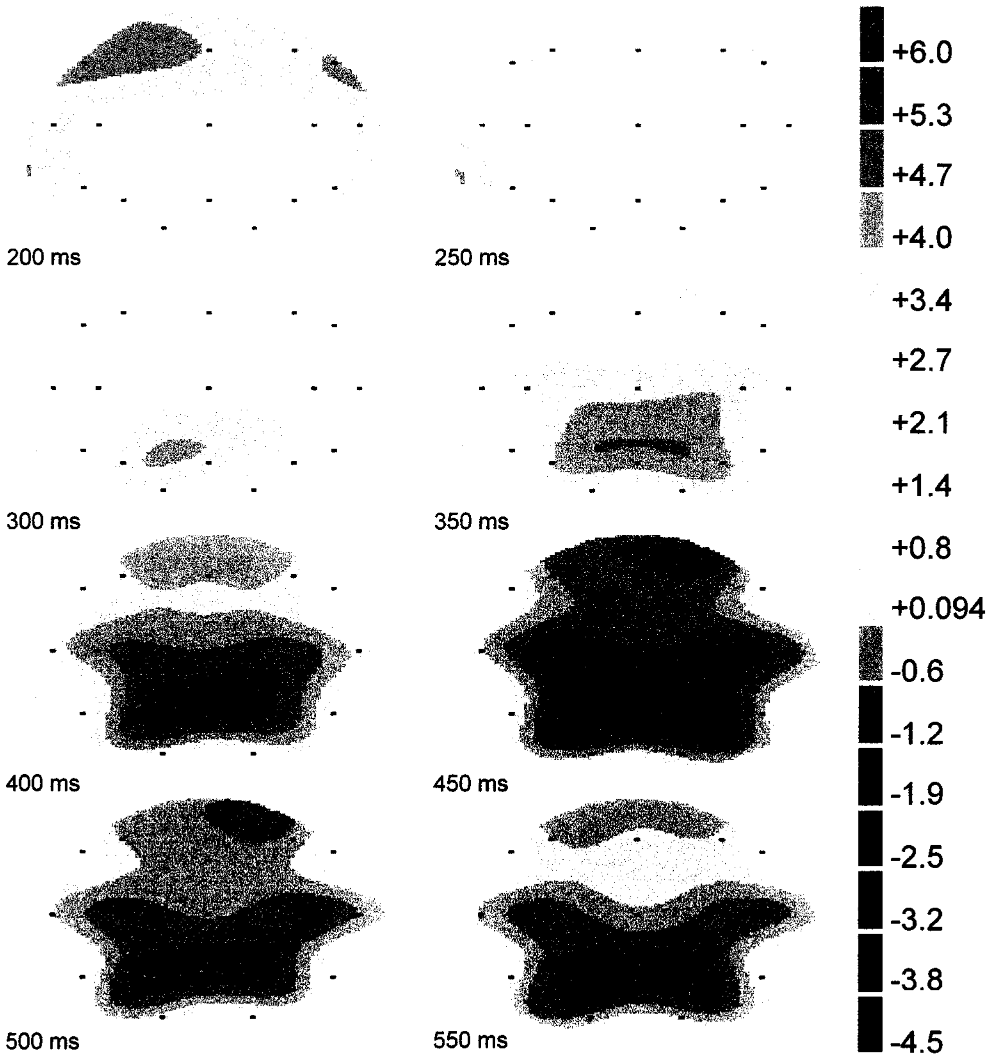
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.3
+4.7
+4.0
+3.4
+2.7
+2.1
+1.4
+0.8
+0.094
-0.6
-1.2
-1.9
-2.5
-3.2
-3.8
-4.5

Figure 23

Figure Caption

Figure 24. Individual waveforms of a representative participant at 17 recording sites for the **Congruent**, **Onset**, **Rime**, **Tone**, and **Incongruent** stimulus conditions presented in Experiment 2. For the display purpose, the low pass filter was set at 15 Hz for the ERP wave forms in this figure.

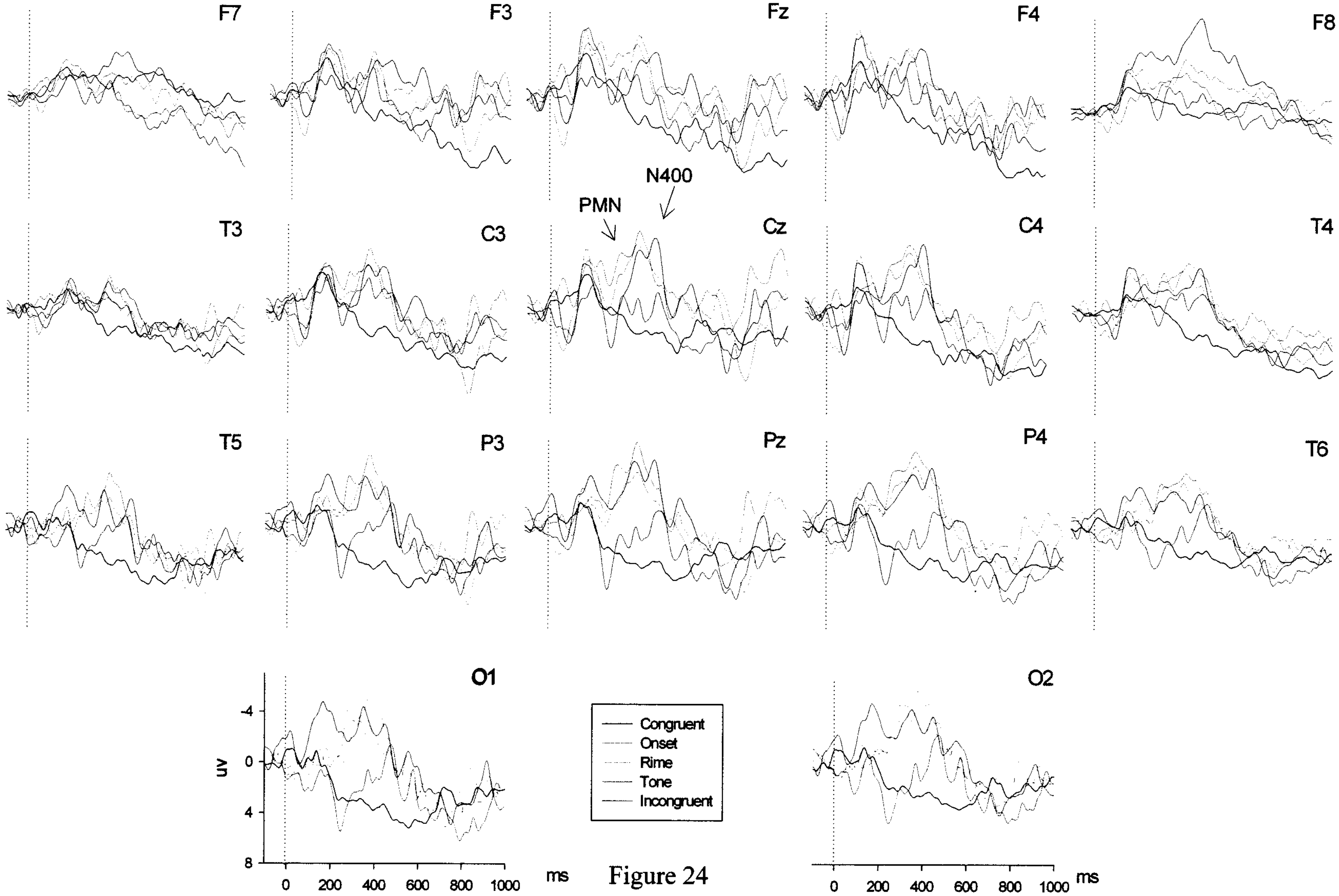
F7
F3
Fz
F4
F8
N400
PMN
T3
C3
Cz
C4
T4
T5
P3
Pz
P4
T6
O1
O2
uv
-4
0
4
8
0
200
400
600
800
1000
ms
Congruent
Onset
Rime
Tone
Incongruent

Figure 24

amplitudes was found) were larger than those in the 350-400 ms time interval in the **Rime** and **Onset** conditions (See Figure 18).

There was a significant Condition x Time x Site three-way interaction (F (192, 2688) = 2.75, p <0.02, ε =0.03). Subsequent analyses indicated that the amplitudes in different sites varied across the conditions in each time episode. In the time window of 350-400 ms, amplitudes in both the **Incongruent** and **Onset** conditions were larger at midline than at other sites (Figs. 19, 20). There was a left parieto-occipital to right centro-parietal characteristic in the **Rime** condition (Fig. 21) whereas a left temporo-occipital distribution in the **Tone** condition (Fig. 22). In the time window of 400-450 ms, there was a centro-parietal distribution in the **Incongruent** condition whereas a fronto-central distribution in the **Onset** condition. In the **Rime** condition, there was a left occipital and right temporo-central distribution. No significant scalp distribution effects were seen in the **Tone** condition although the amplitudes at the left occipital site appeared higher than that over other sites. In the time window of 450-500 ms, there was a fronto-central distribution in the **Onset** condition, while right frontal distribution in the **Incongruent** condition. There was a right fronto-temporal characteristic in the **Rime** condition while a right occipital to temporal characteristic that extended to the right frontal area in the **Tone** condition. In the time window of 500-550 ms, there was a frontal distribution across **Onset, Rime,** and **Tone** conditions, while more right frontal dominant in the **Incongruent** condition.

The Condition x Time x Region x Hemisphere analysis showed that there was no main effect of hemisphere and no interactive effect of Time x Condition x Hemisphere or Condition x Hemisphere, which indicates there was no clear specific hemisphere dominance across five conditions (See Appendix D, Table D7, D8).

3.5 Discussion

In this experiment five conditions were used to investigate the respective roles of word onset and rime, word prosody (tone in Chinese), semantic incongruence, and congruence on event-related brain potentials during spoken word processing. One of the main objectives of this study was to assess the relative importance of onset, rime and tone in the semantic comprehension of words in a tone-based language (Chinese Mandarin). The Chinese characters used presented as well known four-character proverbs. Within this type of context, the N400 that is related to semantic analysis (Bentin, Kutas, & Hillyard, 1993; Connolly & Phillips, 1994; Kutas & Hillyard, 1980) can be viewed as a reliable marker of the degree to which the proverb's context prepares the listeners for the terminal character recognition. One of the main findings of this experiment is that the N400 was observed in all of the conditions (**Incongruent, Tone, Rime,** and **Onset**) except the semantically congruous one. The N400 in response to the semantically **Incongruent** character appeared larger in amplitude than that to the segmental (**Onset** and **Rime**) change and **Tone** change characters. It seems that the N400 is responsive not only to semantic but also to phonological variables. The finding is generally in accordance with English ERP literature.

The N400 was first recognized as a late negative component elicited by the final words in a sentence/phrase (Kutas & Hillyard, 1980a, b) or the second word of a word pair (Kutas & Van Petten, 1988; Kutas & Hillyard, 1989) that did not semantically match the preceding sentence or word. It is thus widely believed that the N400 is an index of the degree of semantic association or priming between two words or a word and its context. Later on, it was found that in a rhyme discrimination task, the N400 component in response to non-rhyming words was greater in amplitude than to rhyming words. This phonological effect on the N400 has further been observed in a lexical decision task (Kramer & Donchin, 1987; Rugg, 1984; Rugg & Barrett, 1987; Sanquist, Rohrbaugh, Syndulko, & Lindsley, 1980; Praamstra & Stegeman, 1993). These findings demonstrate that the N400 component could also be modulated by the phonological variables between words in both visual and auditory modalities. Based on these data, it is argued that the N400 is not only sensitive to semantic processing but also phonological processing. Praamstra and Stegeman further argued (1993) that the N400 does not directly reflect phonological processing, per se. They state that the phonological effects on the N400 'are not related to temporary activation of the segmental phonological representations that establish contact between the input and corresponding lexical entry' during spoken language processing (Praamstra & Stegeman, 1993, p. 80). Rather, it is argued that the magnitude of the N400 reflects the ease with which phonological representations constrain semantic activation of a word in the mental lexicon. Following this account, the significant condition effects on the N400 in the present experiment can be seen as convincing evidence that onset, rime, and tone play a role in constraining lexical access to semantic meaning of a Chinese character.

In particular, the **Tone** condition effect on the N400 demonstrated that Chinese Mandarin made use of tone in constraining semantic activation. In fact, this is not a surprising result given that tone is a component of a Chinese character. Some researchers (Repp & Lin, 1990) believe that tone is of relevance at the word level in tone languages.

In addition to conveying affect, tone information 'serves the highly linguistic function of distinguishing lexical items' (Repp & Lin, 1990, p.483). It is a priori evident that the identification of a Chinese character (e.g., 母/mu3/ -- mother) merely provided by the onset and rime information "mu" would not be successful until the concrete tone is provided.

In line with this view, empirical evidence for the role of tone in lexical processing has been obtained from several lines of research. For instance, using a dichotic listening paradigm, most of the literature indicates that tones in tone languages evoke more of a right-ear advantage than affective intonation (Van Lancker & Fromkin, 1973; Shipley-Brown, Dingwall, & Berlin, 1988). An influential interpretation for this right ear advantage is that there is left hemisphere superiority in the processing of tonal information (e.g., Divenyi & Robinson, 1989). Given that the left hemisphere plays a key role in linguistic processing, the left hemisphere superiority in tone processing has been taken as compelling evidence for the lexical nature of tone. The failure to observe the typical right ear advantage in some tone perception studies was explained as due to atypical processes uncovered during the test and the variable nature of this technique itself since the dichotic listening paradigm is an indirect measure of neural function (Yang, 1991).

Considerable evidence supporting the lexical function of tone also comes from a large body of clinical literature. It has been found that left hemispheric lesions secondary to stroke inhibit tone perception and production while this pattern of deficit is not seen to right hemisphere patients (Basso, Casati & Vignolo, 1977; Blumstein, Baker & Goodglass, 1977; Square-Storer, Darley & Sommers, 1988; Gandour, 1981, 1988). It is worth noting that the scalp distribution of the **Tone** N400 in the present experiment presented a typical left posterior dominance as hypothesized. Although one may argue that because the N400 reflects the "outcome" of form-based encoding its topography might not provide direct insights on neural correlates of tone processing, the left posterior dominance in the **Tone** N400 nonetheless to some degree shows a close correspondence with the left hemisphere superiority proposed in both behavioral and clinical research.

A number of cross-linguistic behavioral studies have also been brought to bear on the argument that tone serves a linguistic function of discriminating lexical items in tone-based languages. The common finding is that tone processing is interfered with by the listener's specific language experience. For instance, in a seminal study, Fox and Unkefer (1985) carried out a categorization experiment in which a continuum was varied from one tone of Chinese Mandarin to another. The crossover point at which listeners switched from reporting one tone to reporting the other shifted as a function of whether the consonant-vowel (CV) syllable affiliated with a tone formed a real word when combined with one tone or the other. The control stimuli consisted of those in which both tones, or neither tone, formed a real word in combination with the CV. It was found that the lexical effect appeared only for native Mandarin listeners; English listeners showed no such shift, and on the control continua the two participant groups did not differ. Later cross-linguistic studies (Lee & Nusbaum, 1993; Lee, Vakoch, & Wurm, 1996; Repp & Lin, 1990) replicated and extended this finding. Different patterns in the interaction of segments (consonants and vowels) and tones were found in tone based and stress based languages. Based on these data, it is argued that tone is of use during Chinese character recognition. It seems likely that tone is part of the representation of a Chinese character in the mental lexicon.

However, Taft and Chen (1992) seemed reluctant to agree with the lexical view on tones. In a lexical decision task, they found that the behavioral responses to words varied in tones involved more errors and were much slower than those varied in segments. They propose that the process of lexical access operates without the use of tone information. It is clear that the disagreement in the literature comes from the interpretation of the above data. As stated previously, the behavioral measure itself is critically sensitive to the postlexical characteristics of a language process (e.g., decision-making and checking). Given this fact, behavioral results (e.g., accuracy and reaction time) may not provide direct insights on prelexical and lexical mechanisms. In this case, the view proposed by Taft and Chen may beg the question, for there is possibility that the 'more error-prone and slower responses' data were attributable to decision making postlexical subprocesses. The explanation for this possibility is considered more completed below.

It is noteworthy that the phonological effects on the N400 are clearly more evident in the **Rime** and **Onset** conditions than in **Tone** condition. A central issue with respect to tonal processing is the extent to which tone contributes to ongoing linguistic processing. It is worth mentioning that the different phonological effects observed between the **Tone** and segmental **(Onset & Rime)** conditions do not necessarily imply that tones are represented and processed in a different fashion from segments or even that tones are superimposed upon segments. In line with the Cohort model (Marslen-Wilson, 1984), a possible interpretation for the different phonological effects described above can be proposed using the concept of neighborhood density (e.g., Goldinger, Luce, & Pisoni, 1989; Marslen-Wilson, Moss, & Halen, 1996). In the mental lexicon, the nodes corresponding to Chinese homonyms that share the same initial consonants and final vowels regardless of tone features are closely related to each other as clusters (high neighborhood density). The activation of a node in such a cluster could be very easily transferred and in turn activate other nodes within the cluster. In this case, the target Chinese homonym can be easily recognized if it corresponds to one of the nodes in the cluster. However, speculating on the difficulty that post-lexical sub-processes (e.g., judgment-making) may meet in order to accurately identify a corresponding node within this cluster, the time for the lexical decision tends to be prolonged since it is more difficult to select an accurate candidate and much easier to make decision mistakes on homonyms than when the targets were not homonyms. This may be a possible account for the 'more error-prone and slower responses' data in the study by Taft and Chen (1992).

Another account of the difference in the phonological effects on the N400 may be derived from the dialect effects on Chinese Mandarin comprehension. It is stated that there is a wide variety of 'dialects' within the Chinese language, most of which are mutually unintelligible in the word prosodic dimension (Hua & Dodd, 2000). For example, the word 'food' will be pronounced 'fan 4' in Putonghua but 'fan 3' in the Shenyang dialect. Put differently, tones in Chinese contain a diacritic feature after being detached from a Chinese character. The tonal manipulation applied in the present experiment in fact followed the tonal rules in Chinese Mandarin. Chinese Mandarin, which is also called 'Putonghua', is one of the major 'dialectal groups' of Chinese spoken by more than 70% of the Chinese population. Since 1950, Chinese Mandarin has become an official language and been widely used in the mass media and taught in schools. In the present experiment, although all the participants were screened as Putonghua Chinese who grew up in the Putonghua-governed cities, they all had at some time been exposed to other dialect environments (e.g., Cantonese). It needs to be kept in mind that this dialect effect might undermine the specificity of the tonal representation in the mental lexicon, especially within such highly constrained contexts produced by Chinese proverbs. The dialect effect on tonal processing has so far rarely been tested especially on the basis of the neural correlates of linguistic processing. This is not a surprise, however, since finding a cohort of exclusive Mandarin speakers is highly challenging.

Taken together, the dialect interference coupled with the above mentioned theoretical interpretation may have contributed to the difference in the phonological effects on N400 elicited in the segment (i.e., **Onset** and **Rime**) conditions and **Tone** condition.

The other main finding in the present experiment is that there is a negativity occurring immediately before the N400, whose peak varies according to discrete phonological variations. In the **Onset** condition, this negativity peaks in 334 ms after the target onset. In the **Rime** condition, the peak of this negativity was situated in 309 ms post the target onset while the peak in the **Tone** condition occurs in 297 ms. If the peak of divergences in this negativity is taken as the temporal marker, then this negativity is much closer to the time window suggested for phonological processing rather than for acoustic perception. Meanwhile, the peaks of this negativity varied in these three conditions seem to reflect dynamic tone, onset, and rime processes.

In fact, an increasing group of English language studies have been brought to bear on this component. In the work of spoken English language processing on N400, it was found that there is a negativity that starts earlier than the classical N400 and then extends onto the classical N400 waveform. Connolly and his colleagues first proposed that this negativity is associated with different processes from the semantic processing manifested by the N400 and named this negativity the Phonological Mismatch Negativity (PMN). In order to confirm their hypothesis, Connolly and Phillips (1994) carried out a sentence comprehension experiment that permitted them to separately manipulate the processes respectively manifested by the PMN and the N400. Successfully, a clear PMN without the following N400 was elicited when the target ending word of the sentence had an initial phoneme different from that of the highest Cloze probability word but was semantically appropriate. In line with the above argument, it seems likely that the negativity that is sensitive to Chinese phonological variations is the PMN proposed by Connolly and Phillips (1994). Based upon the Cohort model (Marslen-Wilson, 1984), Connolly and Phillips (1994) further argued that the PMN might reflect phonological processing at the level of lexical selection in which phonological contextual information facilitates the word recognition processing. It is noteworthy that the disparity in Chinese PMN peaks between **Onset, Rime** and **Tone** conditions further corroborates this view.

According to the data in the present experiment, the PMN elicited in the **Tone** condition peaks earlier than that in the **Rime** condition. This finding indicates that tonal information is processed before vowel at the stage of phonological representation. This is unexpected in light of the current hypothesis that tones are primarily realized upon vowels, and they cannot be processed until the vowel information is available which gives rise to a delayed processing of tonal information (Gandour, Wong, Hsieh, et al, 2000; Lee & Nusbaum, 1993; Millers, 1978). The evidence of this hypothesis, as proponents of this hypothesis claim, is the asymmetric responses to Chinese vowel and tone perception in the speeded classification paradigm (Repp & Lin, 1990) and the slower response to tone compared with that to vowel in vowel and tone monitoring tasks with single characters as targets (Cutler & Chen, 1997; Miller, 1978; Lee & Nusbaum, 1993). It was believed that tones, as supra-segments, when brought up to the phonological stage, are translated into vowels and become part of vowels' phonological configuration.

Supposing that tones at the phonological level are represented by vowel configuration, the PMN in the **Tone** condition should show a similar time course and scalp distribution of that in the **Rime** condition. The data in the present experiment do not support this prediction, however. The dynamic pattern of the PMN scalp distribution in the **Rime** and **Tone** conditions appear independent of each other: the scalp distribution of the **Rime** PMN appears to have more midline dominance while the **Tone** PMN presents more local left centro-temporal distribution.

As a matter of fact, compatible with the present PMN findings, one behavioral study has also cast doubt on the hypothesis that tone perception is slower than vowel perception. In this study, Chinese proverbs were aurally presented and native Chinese listeners were required to decide whether the proverb-ending character contained a vowel /a/ in the vowel monitoring task or had tone 2 in the tone monitoring task (Ye & Connine, 1999). It was expected that if tone information were superimposed on vowel representation, tone information would not be available until vowel is completed, which should result in slower responses to tones than to vowels. However, the findings were inconsistent with what was expected. It was found that tone monitoring appeared faster than vowel identity in highly constrained proverb contexts. It is noteworthy that the tone advantage in this behavioral study is in congruence with the early peak of the PMN in **Tone** condition contrastive with that in the **Rime** condition. Although we do not yet have a full understanding of tones and the interaction of tones and segments, these data are sufficient to propose that tones are phonologically encoded independently of segmental processes. In this case, the hypothesis that vowels are the carriers of tones needs to be further modified and the neural correlates of processes through which tones are associated with segmental processes on the phonological level deserve further examination.

Viewed in this light, another issue of interest is to determine the representation of tone in Chinese Mandarin at a theoretical level. In fact, one of the most heavily researched questions concerning the nature of tones is whether they are supra-segments or auto-segments. Goldsmith (1976) believed that the tone is represented separately from segments in the mental lexicon and decoded independently of segments at the phonological stage during ongoing tonal language processing. The different time course and different brain area dominance of ERPs between the **Tone** condition and the two segment (**Onset** and **Rime**) conditions in the present study seem consistent with this argument. In addition, data from spontaneous tongue slips in Mandarin further confirm this argument (Chen, 1999). The results of this study demonstrated that mispronunciations of segments were common whereas mispronunciations of tones were rare. In addition, the error patterns were different between tone and segment: tone is mostly preserved (e.g., the mispronunciation of a character results from the perseverance of the tone of its preceding character) while segment is primarily anticipated (e.g., the segments of a character are mistakenly replaced with the segments of its following character). Most notable is the finding that tones were never affected by errors of segmental phonemes. The data suggest that tones may be represented and processed differently from segmental information. Put differently, tone can be seen as a distinct unit (like segment) in lexical memory and bears a position in the frame of phonological representation. Taken together, it might be more parsimonious to consider tones as auto-segments rather than supra-segments.

Given that tones and vowels are processed separately during phonological encoding, it is of interest to determine the reason for the literature's contradiction concerning different vowel and tone perception speeds using the same monitoring task. In view of the contradictory performance findings that vowels are recognized faster than tones when the target is presented in a neutral context while tones are identified earlier than vowels when the target is the ending character of highly constrained proverbs, Ye and Connine (1999) postulated that the tone advantage in Chinese proverbs may be due to the highly predictive contextual effect.

A line of supporting evidence for this view comes from a large group of studies on tone languages such as Mandarin (Shen, 1990; Xu, 1994), Taiwanese (Peng, 1997; Lin, 1988) and Thai (Gandour, Potisuk, Dechongkit, et al., 1992a, 1992b). The findings in these studies suggest that tone perception is largely sensitive to contextual information from the speech perceptual perspective. It is claimed that the interaction between neighboring tones is pervasive and universal. According to the location of the source of the effects, the influence of one tone on adjacent tones can be 'anticipatory' or ‘retrospective’ in nature. Using Chinese Mandarin as an example, anticipatory effects are present when the tone of a Chinese character is affected by that of the following character. Shen (1990) demonstrated that tone co-articulation could affect the overall tonal height of preceding tonal contours as well as onsets and offsets of the tones. In this view, it seems likely that tonal co-articulation provides perceptual cues that enhance the predictability of the tone of the following word while vowel perception is generally facilitated by the co-articulation on the basis of word level. This interpretation of available data could account for the tone advantage. At the same time, because tones are generally uttered over time and consequently persevere over time, it seems reasonable that the less rigid temporal constraints in generating tonal sequences contribute to the disparity of tone perception speed in different task contexts. However, the limited evidence from the above studies does not appear adequate enough to clarify the issues as to whether it is the highly predictive proverb contexts per se or co-articulation in speech which contributes to this tone advantage. If both contribute to the said advantage, what is the relationship between these two contexts in tonal processing? Future studies should explicitly manipulate these two contexts to better understand the mechanisms underlying the contextual effects on tone processing.

Chapter Four:

Experiment 3: ERP evidence for lexical tone, onset phonology, rime phonology, and semantic influences during silent Chinese reading

4.1 Objectives:

The results of experiments 1 and 2 demonstrate two important findings with regard to the processes of word prosody and sub-lexical phonology in both English and Chinese speech comprehension. First, both the PMN and N400 components demonstrate their sensitivity to lexical stress manipulation in spoken English word processing and lexical tone variation in spoken Chinese character processing. According to these findings, it is logical to argue that lexical prosody can be represented in the mental lexicon and plays a certain role in spoken word/character processing. Second, sub-lexical phonological cues exert facilitating effects on speech processing. The results of the first two experiments demonstrated that syllable-sized sub-lexical phonology (e.g. /lait/ in the target "po**lite**") facilitates whole spoken English word processing; and, both onset and rime phonology provide important contributions to spoken Chinese character understanding. The finding further suggests that phonological representations of spoken words/characters in the mental lexicon are not encoded in an exclusively all-or-none impoverished way that phonology can be of use in constraining semantic activation of a word only if it completely matches the phonological representation of the word – a "phonology-as a whole" process (e.g., phonology /wΛn/ can be only able to contact semantic representations of words such as "one" and "won" but not "wonder".). Rather, the phonological information of a spoken word/character is represented in a more detailed fashion, based on which the corresponding nodes to distinct words/characters in the mental lexicon are connected to each other by the means of shared sub-lexical phonological features.

The importance of word prosody and sub-lexical phonology in speech comprehension equally introduces another interesting issue into this thesis. That is, are lexical prosody and sub-lexical phonology of relevance in reading? If Lexical prosody is believed to contribute to ongoing language processing, it is reasonable to ask whether it is engaged during reading comprehension. As for the phonological processing in reading, a large body of performance data, as stated in the introduction, has shown that there is phonological activation during Chinese silent reading. However, given the limitations of behavioral research itself, the functional locus and time course of phonological processing in Chinese reading remain to be exploited.

For the investigation of word prosodic and segmental phonological processing in Chinese reading comprehension, an online measure that can capture detailed language processing traces as they unfold over time becomes of great value. With this need in mind, the ERP method was employed in this experiment to examine neural correlates of mental processes underlying Chinese reading comprehension. The major intention of this experiment was to define the functional locus of phonology in the reading processing stream and the time course of phonological representation and its relationship with semantic analysis during Chinese reading. In addition, as addressed in Experiment 2, an influential view on lexical tone in Chinese, which runs counter to the findings obtained in Experiment 2, claims that tonal cues as prosodic information exist only at the acoustic level. In line with this claim, the manipulation in tonal dimension should not influence a bit any aspect of reading processing given that the acoustic stage is the specific feature of speech. With this issue in mind, this experiment was further directed at examining the role word prosody (tones) may play in Chinese reading comprehension. This experiment employed the same phonologically mediated semantic priming paradigm, experimental design, and language stimuli as used in the second Chinese speech experiment and hereby examined the corresponding ERP manifestation. What is different, however, is that the stimuli used in this experiment were presented in the visual modality. See Table 2 for examples of sentences from the stimulus conditions.

4.2 Hypotheses:

If phonological processing takes place during Chinese character recognition, it was hypothesized that contextually primed segmental phonology would facilitate Chinese written character processing. On this view, it was expected that the N400 in response to either **Rime** targets or **Onset** targets should be smaller than that to the **Incongruent** targets. Otherwise, the N400s in response to segmental change targets should not appear any difference from that to the **Incongruent** targets. The logic behind is that the common features between targets and correct proverb ending characters in the **Onset** and **Rime** conditions appear only in the phonological domain given that they have completely different orthographic and semantic attributes (e.g., the target character in the **Onset** condition had the same tone and rime as, but different onset, word form, and meaning from, the correct proverb-ending character.).

If the present data showed that there was a phonological effect on the N400 in response to segmental change targets, the data would provide strong evidence in favor of the argument that there is phonological involvement in Chinese reading. However, given that the N400 is predominantly sensitive to semantic analysis, the phonological effect on the N400 itself cannot provide clear insights into the nature of the phonological activation in reading.

Before the account for the phonological effects on the N400 is proposed, two traditional terms are introduced here to account for the process of generating phonological representations from print (Patterson & Coltheart, 1987). The first 'assembled phonology' is a computational process that utilizes certain conversion rules by which orthographic units can be transformed into phonemic units (orthography-to-phonology transformation). The alternative to this computation procedure is named 'addressed phonology' as a phonological representation in this case is retrieved from lexical storage after orthographic cues of a printed word have addressed the lexicon. Assembled phonology is rule based as a detailed process regardless of lexical properties whereas addressed phonology is lexically driven and is thereby an abstract phonology-as-a-whole-unit process.

In line with this view, it was further hypothesized that if phonological transformation took place to constrain semantic meaning of a target character, the N270, which reflects form-based processing in reading, in response to both the **Rime** target and the **Onset** target, should appear smaller than that to the **Incongruent** target character. Furthermore, the morphology of the N270 should be sensitive to distinct segmental phonological manipulations given sophisticated features of assembled phonology. Otherwise data would indicate that phonological activation observed in the N400 component was not derived from orthography to phonology transformation. Rather, the phonological priming effects on the N400 may result from post-lexical contextual priming that pre-activated the corresponding lexical node in the mental lexicon.

Regarding the lexical prosodic issue, it was hypothesized that lexical tones are represented in the mental lexicon and play a certain role in semantic constraints of written Chinese characters. This hypothesis would give rise to the expectation that a smaller N400 should be observed in response to **Tone** targets in contrast with that to **Incongruent** targets.

4.3 Method:

4.3.1 Participants

Fifteen right-handed native speakers of Chinese Mandarin (7 male and 8 female) between 19 and 28 years of age (M = 23.73, SD = 2.46) were recruited at Dalhousie University and paid for their participation in this experiment. All the participants had normal or corrected-to-normal vision and were screened with a self-report measure for any history of neurological, psychiatric or reading disorders. They all provided informed consent and were debriefed fully after the experiment.

4.3.2 Stimuli and experimental conditions

260 four-character Chinese proverbs were used, which were identical to those used in the prior Chinese speech experiment (See Appendix C for sentence stimuli). The proverbs were presented one character at a time on a computer monitor and subtended a visual angle that averaged about 4 degrees. The characters were presented in yellow on a black background. All characters were exposed for 500 ms and the interval between

characters (ISI) was 500 ms. Prior to the presentation of the first character of a proverb a warning stimulus with duration of 300 ms was presented, which terminated 1s before the presentation of the first character. There was a 2 s pause between trials.

The proverb stimuli were divided across five experimental conditions in which the ending character of the proverbs varied with respect to onset, rime, tone and/or semantic appropriateness. In the **Congruent** condition, the ending character of a proverb was the accurate fourth character of that proverb (e.g., “入 乡 随 俗 /su2/” , in which the ending character (俗) is custom). In the **Onset** condition, the proverb-ending character had the same rime and tone as, but different onset, orthography, and meaning from, the correct fourth character of the proverb. For instance, 对 牛 弹 林/lin2/, in which the terminal character should have been ‘琴 /qin2/’ lute. However, the character ‘林’ means woods, which did not orthographically or semantically fit the context of the proverb. In the **Rime** condition, each proverb ended with a character that had the same onset and tone as, but different rime, orthography, and meaning from, the correct terminal character of the proverb. For example, 爱 屋 及 挖 /wa1/, in which the terminal character should have been ‘black’ (乌 /wu1/) but was, in fact, ‘dig’ (挖 /wa1/). In the **Tone** condition, the terminal character of the proverb had the same onset and rime as, but different tone, orthography, and meaning from, the correct fourth character of the proverb. Thus, ‘不 期 而 鱼 /yu2/’, is a proverb in which the terminal character should have been ‘meet’’遇 /yu4/’ but was in fact, ‘fish’ ‘鱼 /yu2/’. In the **Incongruent** condition, the onset, rime, tone, orthography, and meaning of the terminal character were completely different from those of the correct fourth character of the proverb. For example, “循 序

渐 盘 /pan2/” should have ended with the character meaning ‘improve’’进 /jin4/’ but in fact ended with the character meaning ‘plate’ ‘盘’. 160 fillers were proverbs with accurate ending characters. The conditions were presented pseudo-randomly with the restriction that any of the five conditions could not be presented successively more than three times.

4.3.3 Procedure

Each participant was tested in one experimental session that lasted around 35 minutes. During the experiment, participants were seated in a comfortable, padded chair in a sound attenuated room adjoining the room with the recording equipment. Stimuli were presented on a computer monitor located about 100 cm in front of the participant. Participants were told not to blink from the start of the warning signal until a short pause after the fourth character presentation and instructed to read silently the proverbs that were presented one character at a time. When the second or third character appeared they were told to think, as fast as they could, of a proper Chinese character to end the proverb. As soon as the fourth character was presented they were asked to <u>think of</u> its meaning as quickly as possible. Participants were told that they might be asked questions about the proverbs at the end of the experiment. Participants were given time to practice until they fully understood the task demands for this experiment.

4.3.4 EEG Recording

The electroencephalogram (EEG) was recorded by means of Tin electrodes from 17 scalp-recording sites with linked ears as the reference (with a ground on the mid-forehead). Recording locations were derived from standard international 10-20 system scalp locations (Jasper, 1958) at Fz, F3, F4, F7, F8, Cz, C3, C4, Pz, P3, P4, T3, T4, T5, T6, O1, O2. The electro-oculogram (EOG) recorded vertical eye movements with electrodes placed above and below the right eye. Horizontal eye movements were recorded with electrodes placed over the outer canthi of the right and left eye. Electrode impedance was kept below 5 kΩ. The analogue EEG recordings were obtained with a half-amplitude bandpass of 0.01-100Hz, digitally sampled at 500 Hz, low pass filtered off line at 30 Hz, and high pass filtered offline at 0.1 Hz. The sample duration began 100ms before the proverb ending character onset and continued until 1000ms after the target character onset. Trials contaminated by EOG greater than ±75μv or any other artifacts (e.g., muscle activity) were excluded from the analysis. The remaining EEG trials were then averaged by experimental condition. The individual average ERPs were also averaged to create grand average waveforms for the various conditions.

4.3.5 Data analysis

Data were analyzed several ways in order to obtain the best possible view of the negative components observed in this experiment in the 200-550 ms time range. The main assessment approach used the interval scoring method by taking the mean amplitudes of 50 ms epochs successively from 200-350 ms post stimulus onset for the N270 component and from 350-550 ms after the stimulus onset for the N400 component. Meanwhile, in order to better isolate the variations in the N270 component across varied conditions, the peak identification method was applied to the data. The N270 was scored as the most negative point in the 200-350 ms time range. Peak latency was defined as the time from stimulus onset to the point scored as the most negative within the latency range.

The statistical analyses of the ERP data were conducted with a repeated measures analysis of variance (ANOVA) with conservative corrections to the degrees of freedom (Greenhouse & Geisser, 1959). For the time interval approach, the primary analyses of the N270 and the N400 components included the Condition factor (5 levels: **Incongruent, Onset, Rime, Tone**, and **Congruent**), a Time factor (for the N270, the time intervals were 200-250ms, 250-300ms, and 300-350ms; and for the N400, the time intervals were 350-400ms, 400-450ms, 450-500ms, and 500-550ms), and a Site factor (17 levels). In addition, a secondary analysis was conducted to examine further the spatial distribution of ERPs. This analysis was conducted with Condition (5 levels) as one factor and the recording sites divided into two factors: hemisphere (left, right) and region (5 levels: frontal, central, parietal, temporal and occipital). The frontal level combined F3 and F7 for a left hemisphere frontal value and F4 and F8 for the corresponding right hemisphere value. A similar approach was conducted with temporal sites combining T3 and T5 for the left and T4 and T6 for the right. The central level consisted of C3 and C4 values while the parietal level consisted of the P3 and P4 values. The Condition x Time x Region x Hemisphere analysis was done as a way of simplifying the assessment of scalp topography differences. Results from the Condition x Time x site analyses were presented first and those from the Condition x Time x Region x Hemisphere analyses were presented only if they provide new information or clarification of previous results.

As for the peak scoring approach, the first analysis of the data included Condition (5 levels) and Site (17 levels) as factors in the ANOVA. The second analysis was conducted to facilitate the assessment of topographical differences by investigating the data associated with specific brain areas and hemispheres. This analysis included Condition (5 levels) as one factor and the recording sites were divided into two factors: Hemisphere (2 levels) and Region (5 levels). These factors were also analyzed in an ANOVA with repeated measures. Results from the primary analyses are presented in full. For the sake of brevity, results from the secondary analyses are presented only when new information is provided.

Significant main effects and interactions were subjected to further post hoc analyses by means of the Tukey Honestly Significant difference (HSD) test. Furthermore, data involved in interactions between conditions and sites or regions and hemispheres were reassessed with repeated measures ANOVAs of normalized amplitudes, as underscored by McCarthy and Wood (1985). Only those interactions that were significant on normalized amplitude measures were reported. An alpha level of $p < .05$ was required for statistical significance.

4.4 Results

4.4.1 Waveforms

Observation of the grand average waveforms (Fig.25A) reveals clear differentiation among the five experimental conditions. There is a late negative component peaking in the 350-400 ms range (N400) associated with the **Incongruent** condition. Although the N400 is also observed in the **Onset**, **Rime**, and **Tone** conditions, the general shapes of these N400s are apparently smaller than that in the **Incongruent** condition. A closer inspection demonstrates that the N400s for the **Onset**, **Rime**, and

Figure Caption

Figure 25 A. Grand average waveforms at 17 recording sites from 15 individuals for the **Congruent**, **Onset**, **Rime**, **Tone**, and **Incongruent** stimulus conditions presented in Experiment 3.

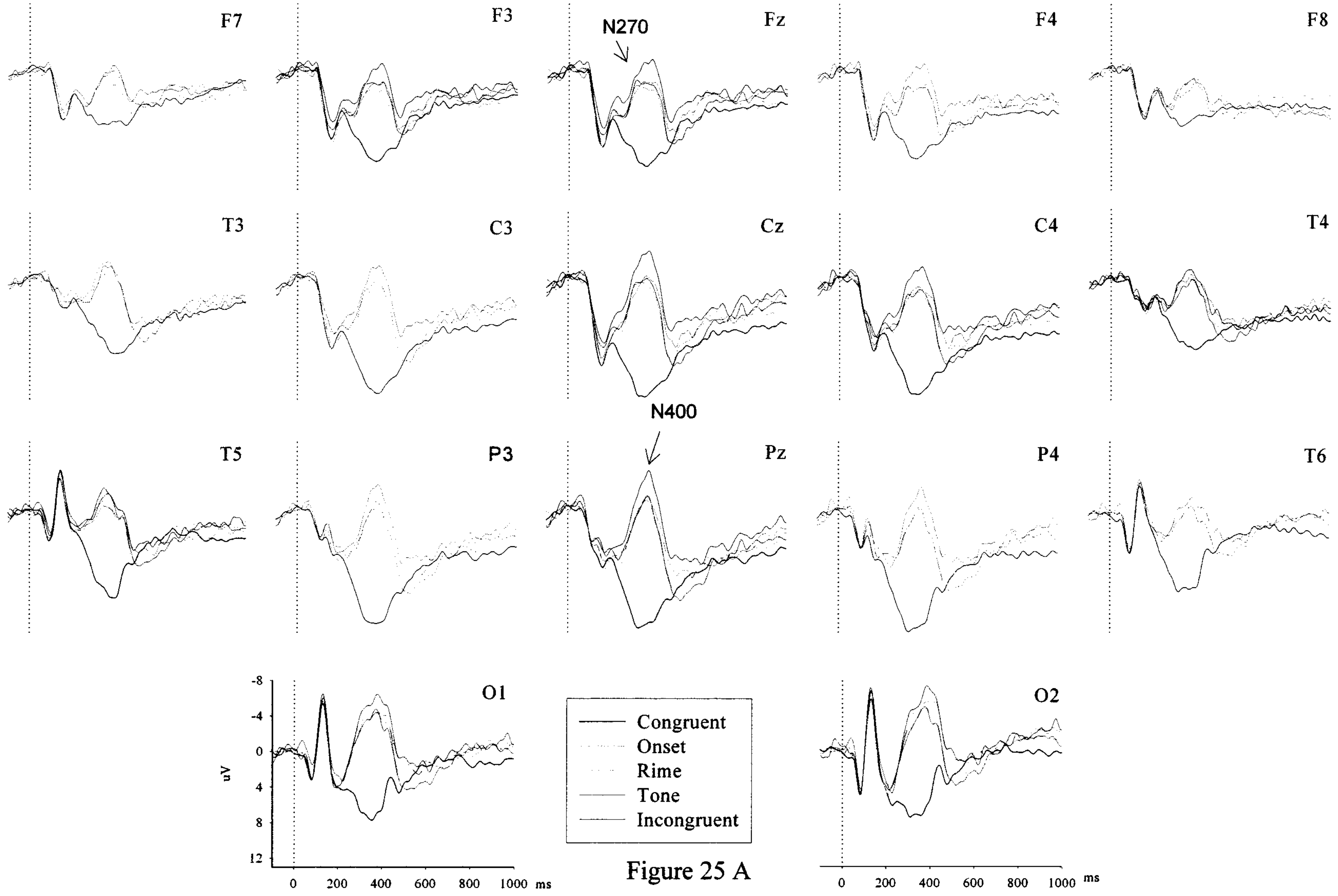
F7
F3
N270
Fz
F4
F8
T3
C3
Cz
C4
T4
N400
T5
P3
Pz
P4
T6
O1
O2
-8
-4
0
4
8
12
uV
0
200
400
600
800
1000
ms
Congruent
Onset
Rime
Tone
Incongruent

Figure 25 A

Figure Caption

Figure 25 B. Grand average waveforms at Fz from 15 individuals for the **Congruent**, **Onset**, **Rime**, **Tone**, and **Incongruent** stimulus conditions presented in Experiment 3.

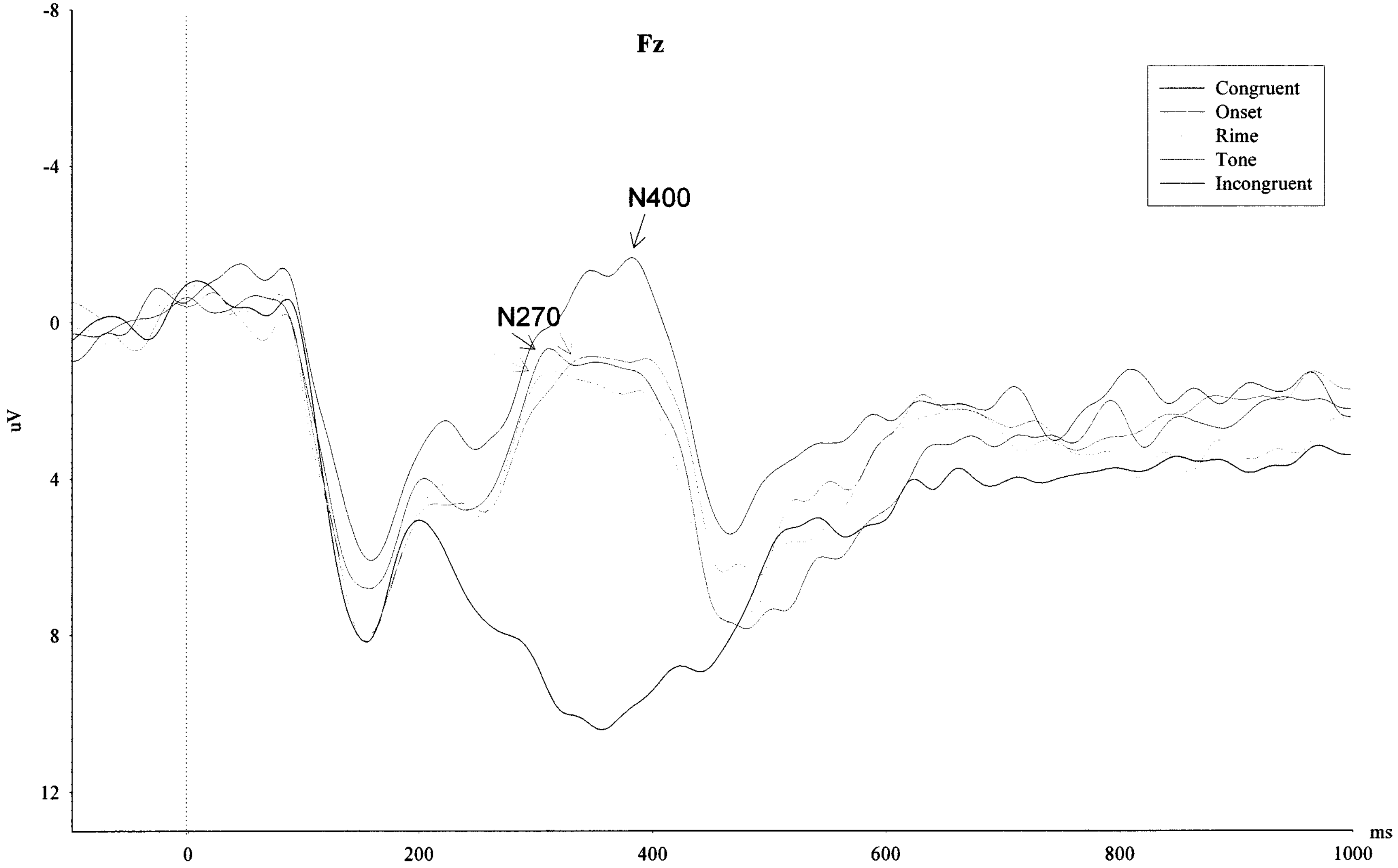
Fz
N400
N270
Congruent
Onset
Rime
Tone
Incongruent
-8
-4
0
4
8
12
uV
0
200
400
600
800
1000
ms

Figure 25B

Figure Caption

Figure 25 C. Grand average waveforms at Pz from 15 individuals for the **Congruent**, **Onset**, **Rime**, **Tone**, and **Incongruent** stimulus conditions presented in Experiment 3.

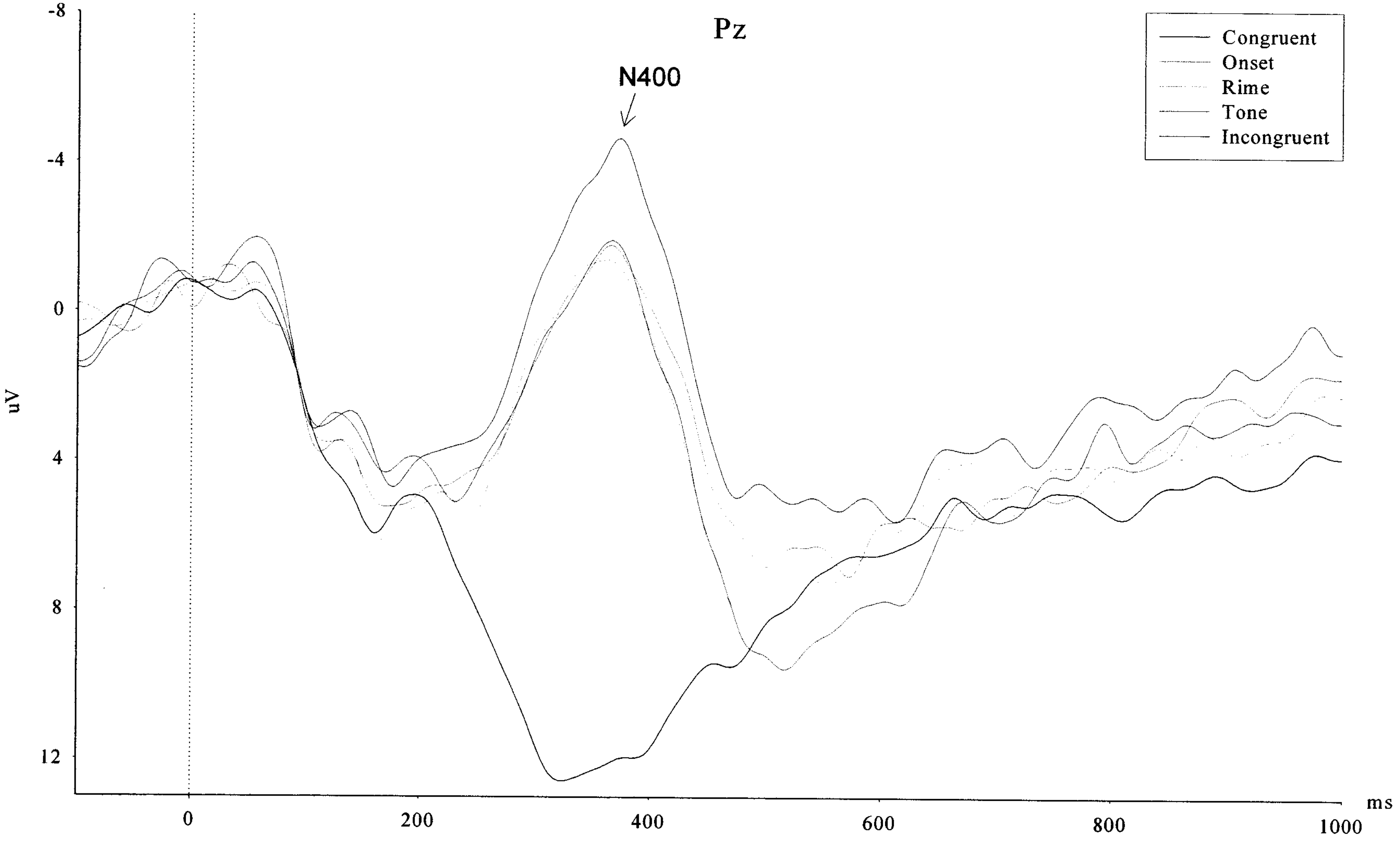
Pz
N400
Congruent
Onset
Rime
Tone
Incongruent
-8
-4
0
4
8
12
uV
0
200
400
600
800
1000
ms

Figure 25C

Tone conditions show almost identical wave shapes with parietal sites having the largest amplitudes (in particular at Pz) (See Fig. 25C). In the time range of N400, there is a clear deep positive deflection in the **Congruent** condition. Prior to the N400 component, there is another negative component (N270) that is extending over and overlapping with the N400. The N270 can be observed in the **Incongruent** condition most apparently but is also clear in the **Onset**, **Rime**, and **Tone** conditions. The N270 appears to peak differently across the **Onset**, **Rime**, and **Tone** conditions. It seems that the N270 to the **Rime** and **Tone** targets peaks earlier than that to the **Onset** target. This phenomenon can be seen most clearly over fronto-temporal recording sites (See Figs. 25B). The condition effects on the N400 and the N270 observed in the grand average waveforms are identifiable in each individual participant ERP waveforms. Figure 35 demonstrates the waveforms of a representative participant.

4.4.2 Statistical analyses

The N270

The interval scoring analyses:

Analysis indicated the sensitivity of the N270 to the different manipulations of the phonological units (onset, rime and tone) ($F(4, 56) = 45.64$, $p < 0.001$, $\varepsilon = 0.54$) (See Appendix E, Table E1): the mean amplitudes in the **Incongruent** condition ($M = 0.52$ µV, SE = 0.73) were significantly larger than those in the **Onset** ($M = 1.75$ µV, SE = 0.69), **Rime** ($M = 1.85$ µV, SE = 0.77), and **Tone** ($M = 1.73$ µV, SE = 0.83) conditions which were further larger than those in the **Congruent** ($M = 6.72$ µV, SE = 0.98) condition. No differences in amplitude were observed among any of the **Onset**, **Rime**, and **Tone**

Figure Caption

Figure 26. Graph depicting the significant interaction between the Condition and Time factors for N270 amplitude data in Experiment 3.

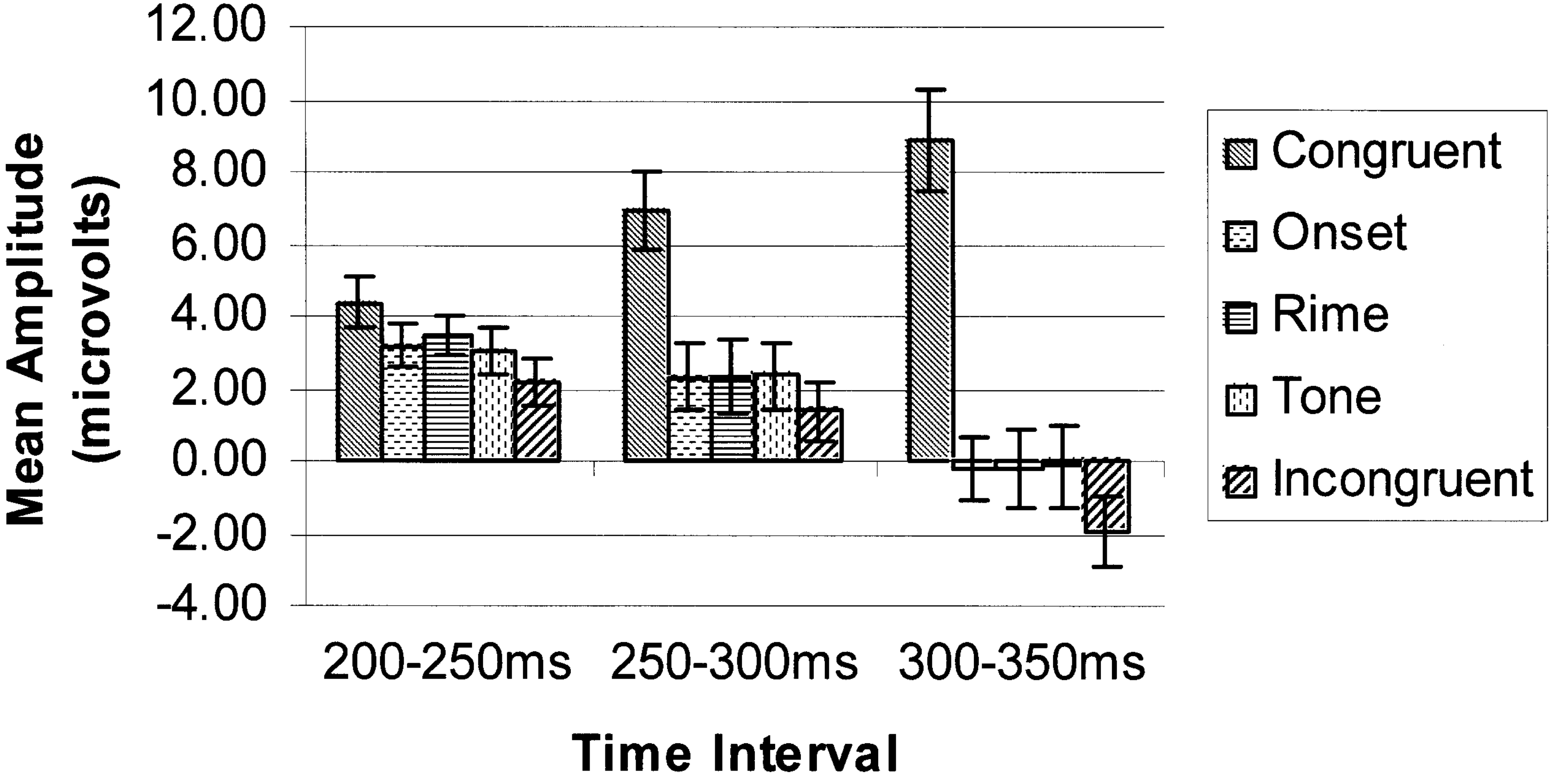
Mean Amplitude (microvolts)
12.00
10.00
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
200-250ms
250-300ms
300-350ms
Time Interval
Congruent
Onset
Rime
Tone
Incongruent

Figure 26

conditions in the three time intervals (See Figure 26). A main effect of Time ($F (2, 28) = 5.17$, $p< 0.05$, $\varepsilon = 0.64$) was found. The Condition x Time interaction ($F (8, 112) = 59.01$, $p< 0.001$, $\varepsilon = 0.38$) and subsequent post-hoc analyses revealed that the largest mean amplitudes in the **Incongruent**, **Onset**, **Rime**, and **Tone** conditions appeared in the 300-350 ms time window while those in the semantically **Congruent** condition appeared in the 200-300ms time window.

There was a main effect of Site ($F (16, 224) = 5.04$, $p < 0.05$, $\varepsilon = 0.13$) and a significant Condition x Time x Site three-way interaction ($F (128, 1792) = 10.31$, $p< 0.001$, $\varepsilon = 0.04$) (See Appendix E, Table E2). Subsequent analyses indicated that the amplitudes at different sites varied across the conditions in each time interval. In the 200-300 ms time range, there was a temporo-occipital distribution in the **Onset**, **Rime**, and **Incongruent** conditions but a left temporo-occipital distribution in the **Tone** condition. In the 300-350 ms time range, there was a left temporo-parietal distribution extending to the occipital area in the **Incongruent** condition. There was a tempro-occipital distribution that extended to the parietal area in the **Onset** and **Rime** conditions while a left temporo-occipital distribution in the **Tone** condition.

There was a main effect of region ($F (4, 56) = 6.69$, $p<0.05$, $\varepsilon = 0.34$) and a significant Condition x Time x Region interaction ($F (32,448) = 11.90$, $p<0.001$, $\varepsilon = 0.11$) (See Appendix E, Tables E3, E4), which further confirmed the results from the Condition x Time x Site analyses. There was also a Condition x Hemisphere interaction ($F (4, 56) = 2.75$, $p<0.05$, $\varepsilon = 0.83$) (See Figure 27). Subsequent analyses indicated that there was a left hemisphere distribution in the **Tone** condition while no hemisphere lateralization in the other four conditions was found.

Figure Caption

Figure 27. Graph depicting the significant interaction between the Condition and Hemisphere factors for N270 amplitude data in Experiment 3.

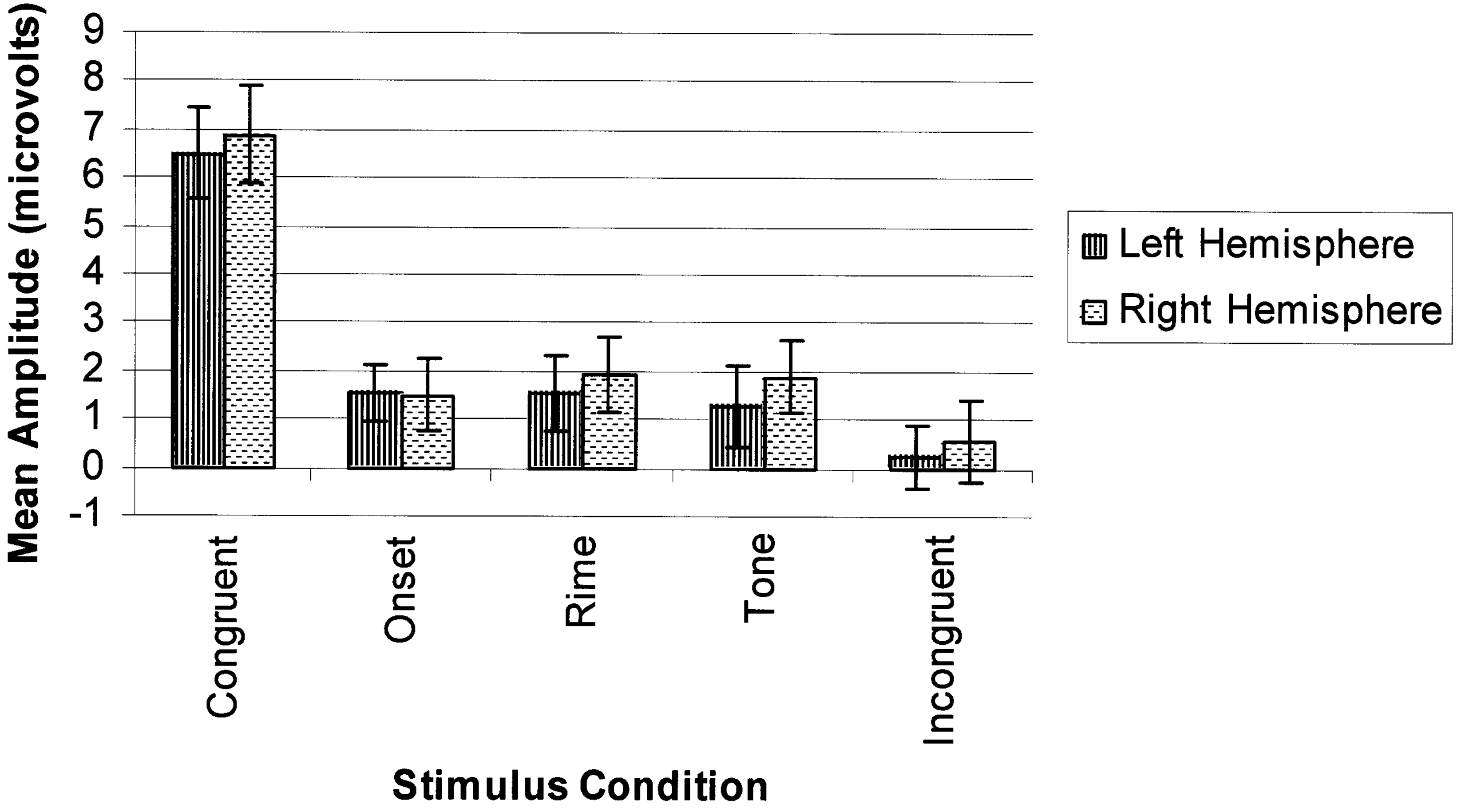
Mean Amplitude (microvolts)
9
8
7
6
5
4
3
2
1
0
-1
Congruent
Onset
Rime
Tone
Incongruent
Stimulus Condition
▥ Left Hemisphere
▤ Right Hemisphere

Figure 27

The peak scoring analyses:

The results of the peak scoring analyses showed that the peak latency of the N270 component varied as a function of the stimulus condition (F (4, 56) = 21.38, p<0.001, ε = 0.53) (See Appendix E, Table E5). Subsequent post-hoc analyses revealed that the peak of the N270 component to the **Rime** targets (307 ms, SE = 4.36) and the **Tone** targets (307 ms, SE = 5.07) appeared earlier than that to the **Onset** characters (330 ms, SE = 4.78) and the **Incongruent** characters (323 ms, SE = 5.26). No differences in the peak latency of the N270 were observed between the **Onset** and **Incongruent** conditions (See Figure 28).

N400

Analyses confirmed the sensitivity of the N400 amplitude to the experimental conditions (F (4, 56) = 61.65, P<0.0001, ε = 0.70) (See Appendix E, Table E6): the mean amplitudes of the N400 in the **Incongruent** condition (M = 0.90 μV, SE= 0.70) were greater than those in the **Onset** (M = 2.22 μV, SE = 0.67), **Rime** (M = 3.04 μV, SE = 0.56), and **Tone** (M = 3.37 μV, SE = 0.73) conditions which were further greater than those in the **Congruent** condition (M = 7.31 μV, SE = 0.66). Main effects of time (F (3, 42) = 48.50, p< 0.001, ε = 0.53) and site (F (16, 224) = 10.07, p<0.001, ε = 0.13) were also found. There was a significant Condition x Time interaction (F (12, 168)= 45.70, p<0.001, ε = 0.22), and subsequent post-hoc analyses using the Tukey HSD test indicated that the amplitudes in the **Incongruent** condition were greater than those in the **Onset**, **Rime**, and **Tone** conditions in the 350-400 ms time range. However, in the 400-550 ms time ranges, the mean amplitudes in the **Incongruent** and **Onset** conditions did not differ

Figure Caption

Figure 28. Graph depicting the peak latency of the N270 response varied as a function of the stimulus condition in Experiment 3.

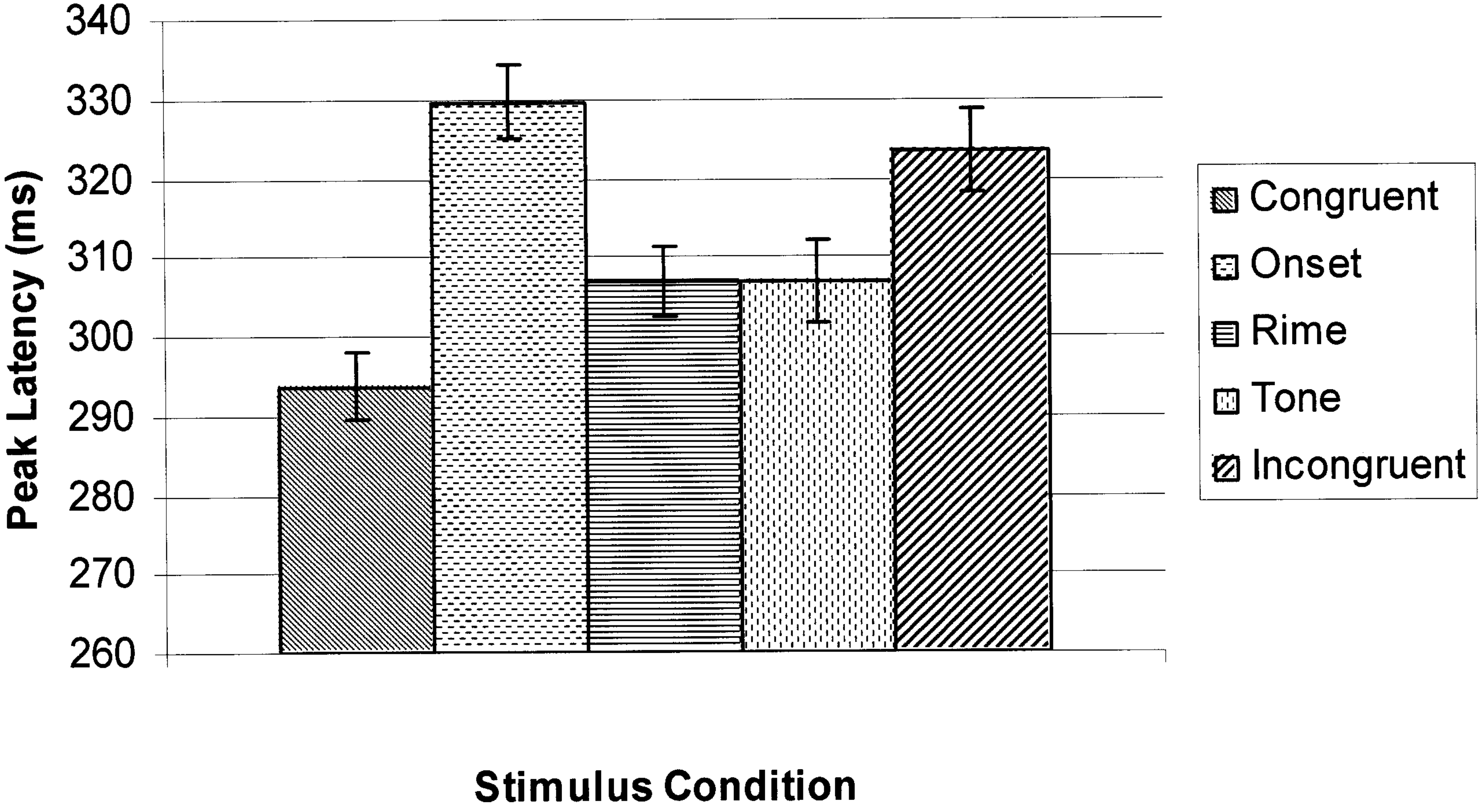
Peak Latency (ms)
340
330
320
310
300
290
280
270
260
Stimulus Condition
▧ Congruent
▤ Onset
▤ Rime
▥ Tone
▨ Incongruent

Figure 28

Figure Caption

<u>Figure 29.</u> Graph depicting the significant interaction between the Condition and Time factors for N400 amplitude data in Experiment 3.

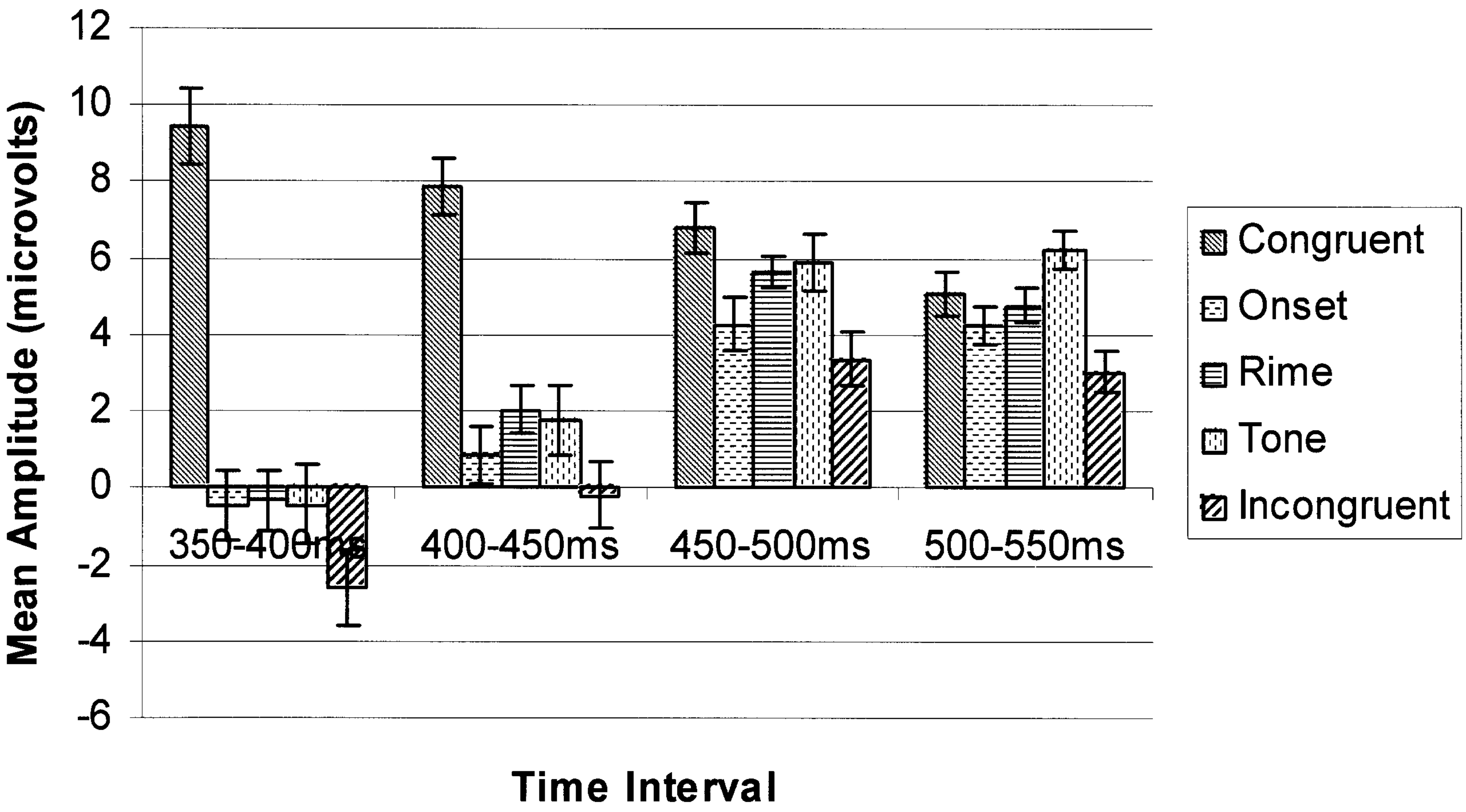
12
10
8
6
4
2
0
-2
-4
-6
Mean Amplitude (microvolts)
350-400ms
400-450ms
450-500ms
500-550ms
Time Interval
▧ Congruent
▤ Onset
▤ Rime
▥ Tone
▨ Incongruent

Figure 29

from each other although both were greater than those in the **Rime** and **Tone** conditions (see Figure 29).

There was a significant Condition x Time x Site interaction (F (192, 2688) = 8.52, $p<0.001$, $\varepsilon = 0.03$) (See Appendix E, Table E7). Subsequent post-hoc analyses indicated that in the 350-400 ms time range, there was a parietal to occipital characteristic in the **Onset**, **Rime**, and **Incongruent** conditions (Figs 30, 31, 32), but a left temporal distribution extending to the parieto-occipital area in the **Tone** condition (Fig 33). Moreover, the amplitudes over left-sided sites tended to be greater than those over right-sided sites only in the **Tone** condition. In the 400-450 ms time range, there was a parietal distribution that extended to the occipital region in the **Incongruent** condition while a temporo-parietal distribution in the **Tone** condition but a parietal distribution extending to the right temporal area in the **Onset** and **Rime** conditions. In the 450-500 ms time range, there was a right temporal extending to the occipital area in all four conditions other than **Congruent** condition (Fig 34). In the 500-550 ms time range, there was a left frontal extending to the occipital area in the **Incongruent** condition and the **Tone** condition. However, there was a right temporal to occipital characteristic in the **Onset** and **Rime** conditions.

The time/condition/region/hemisphere analysis of the amplitude found a significant Condition x Time x Region interaction (F (48, 672) = 9.28, $p< 0.0001$, $\varepsilon =0.07$), a marginally significant Condition X Region X Hemisphere interaction (F (16,224) = 2.60, $p< 0.05$, $\varepsilon =0.32$), and the four-way Condition x Time x Region x Hemisphere interaction (F (48, 672) = 2.38, $p< 0.05$, $\varepsilon =0.13$), which further confirmed the results from the Time/Condition/Site analyses (See Appendix E, Table E8).

Figure Caption

Figure 30. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Incongruent** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 3. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

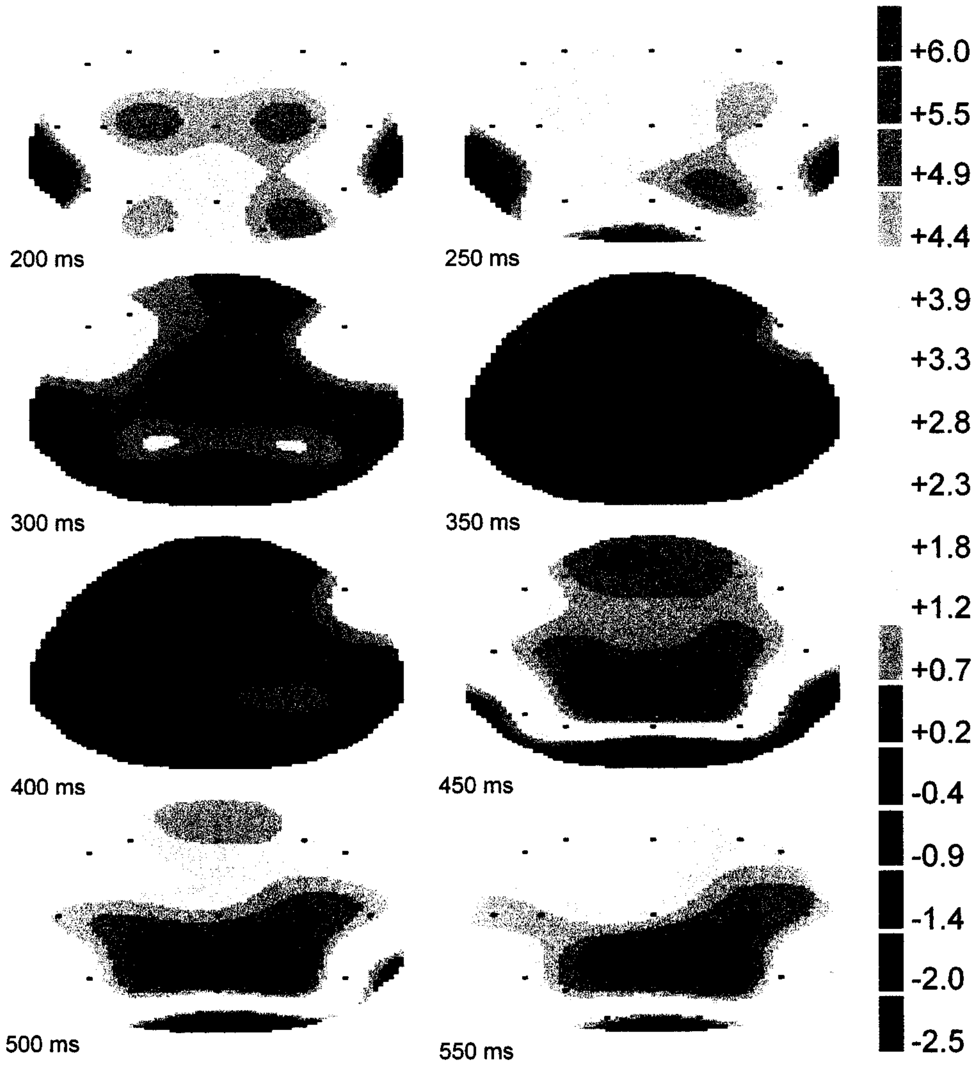
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.5
+4.9
+4.4
+3.9
+3.3
+2.8
+2.3
+1.8
+1.2
+0.7
+0.2
-0.4
-0.9
-1.4
-2.0
-2.5

Figure 30

Figure Caption

Figure 31. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Onset** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 3. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

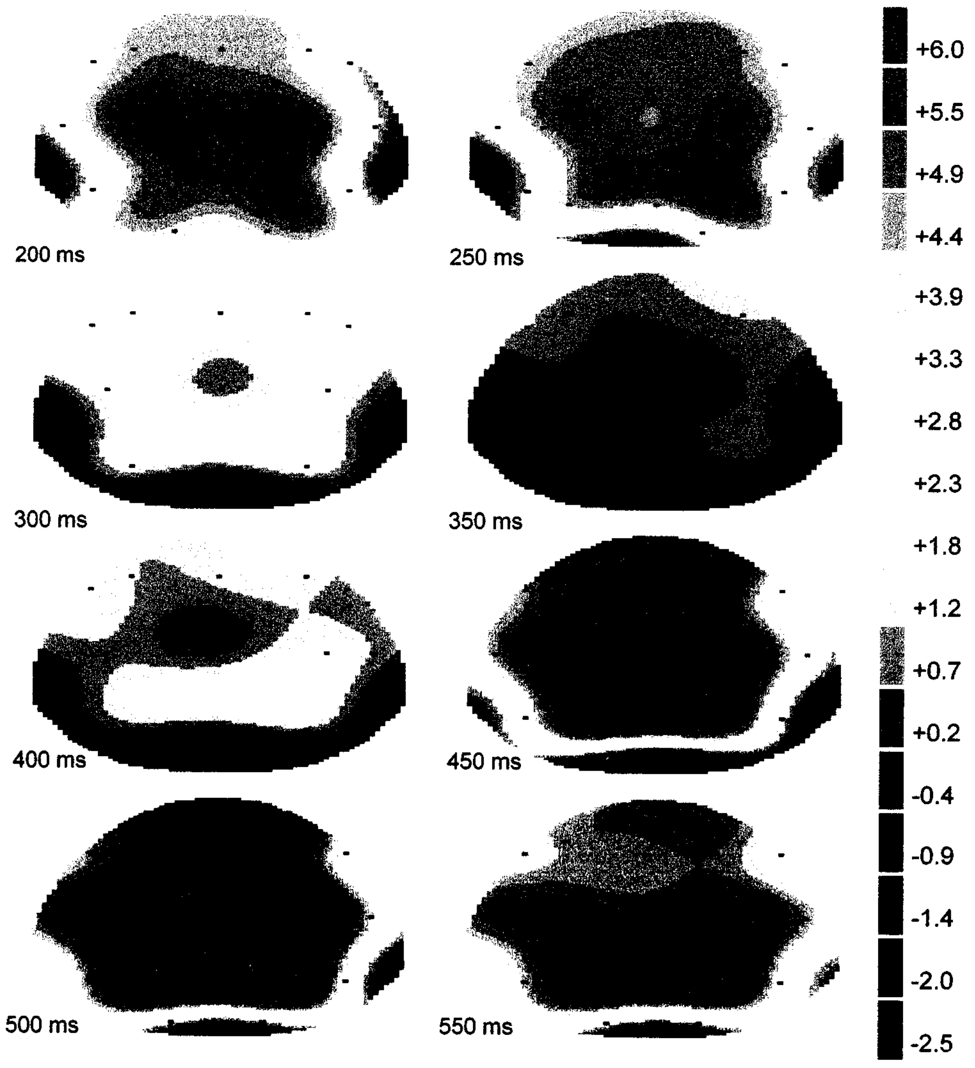
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.5
+4.9
+4.4
+3.9
+3.3
+2.8
+2.3
+1.8
+1.2
+0.7
+0.2
-0.4
-0.9
-1.4
-2.0
-2.5

Figure 31

Figure Caption

Figure 32. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Rime** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 3. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

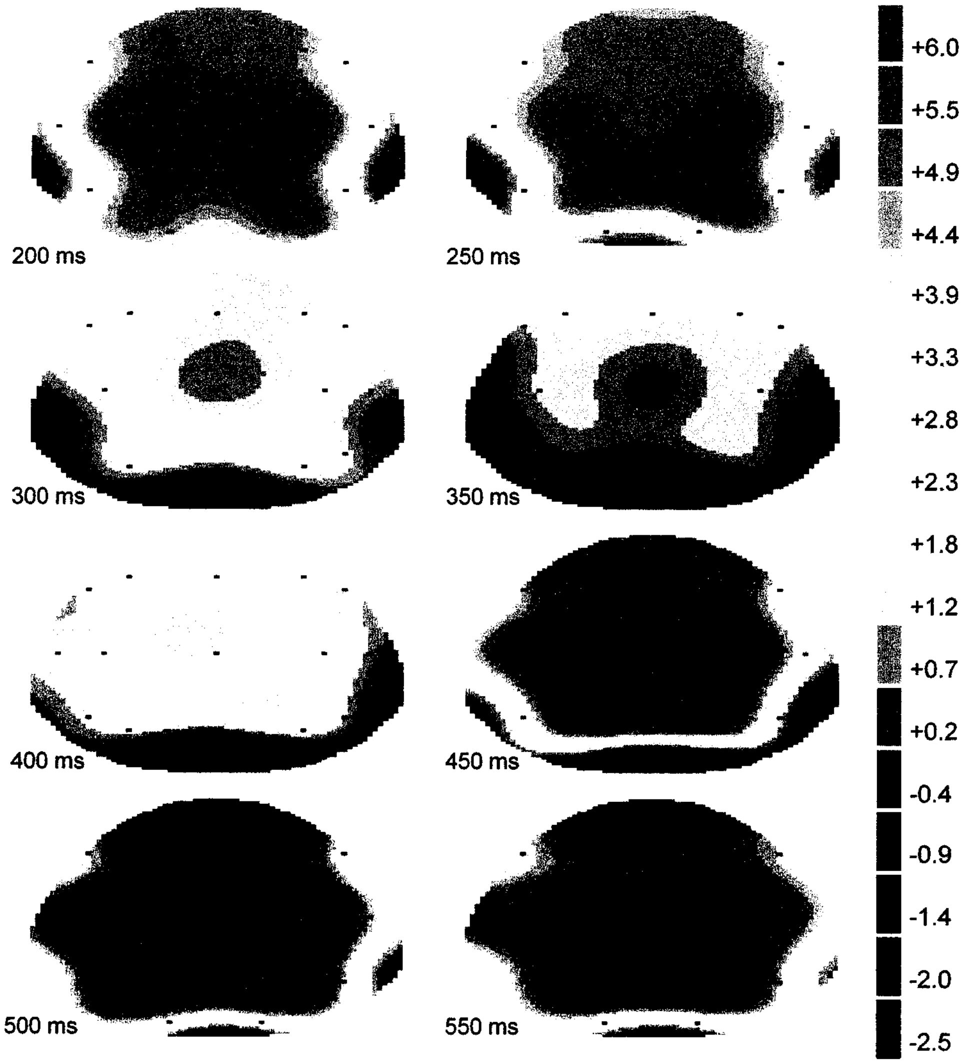
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.5
+4.9
+4.4
+3.9
+3.3
+2.8
+2.3
+1.8
+1.2
+0.7
+0.2
-0.4
-0.9
-1.4
-2.0
-2.5

Figure 32

Figure Caption

Figure 33. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Tone** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 3. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

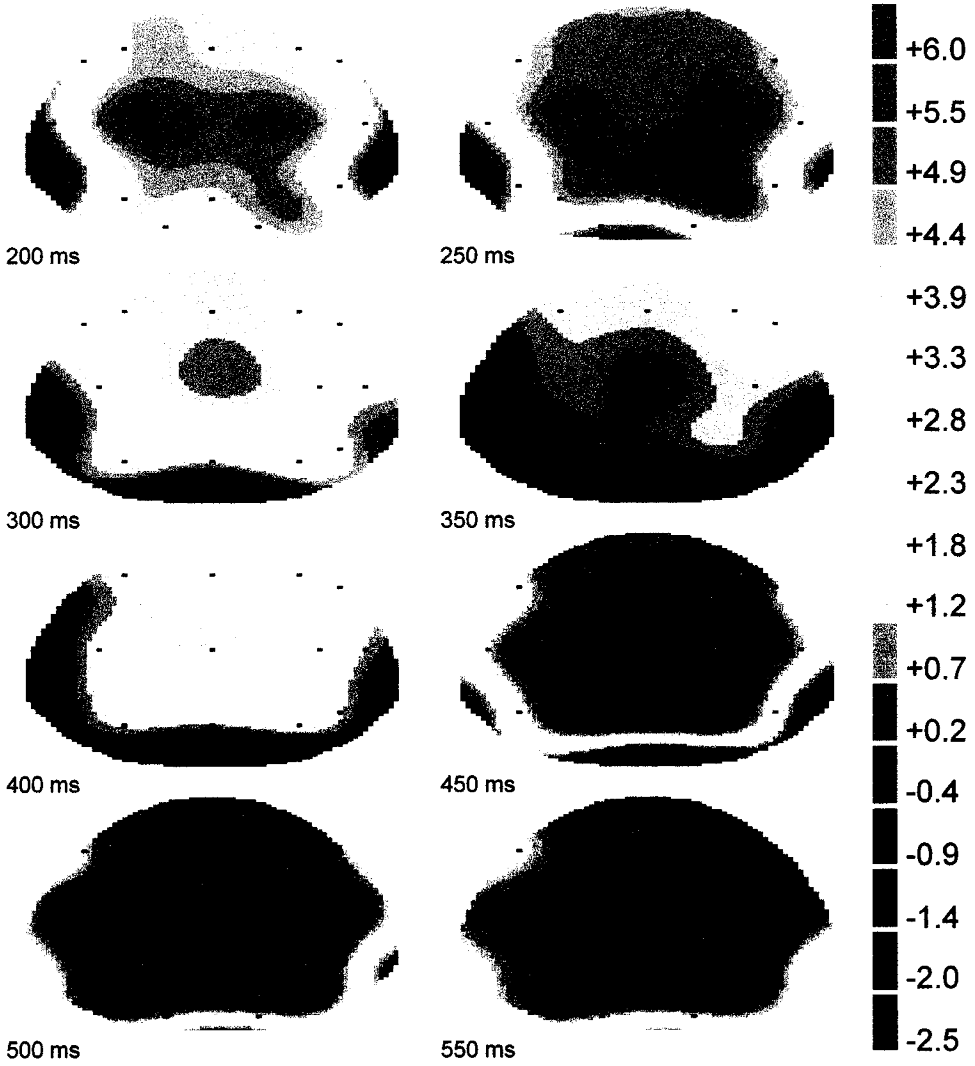
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.5
+4.9
+4.4
+3.9
+3.3
+2.8
+2.3
+1.8
+1.2
+0.7
+0.2
-0.4
-0.9
-1.4
-2.0
-2.5

Figure 33

Figure Caption

Figure 34. Graph depicting bird-view 2D scalp potential distributions of grand average ERPs (from 15 individuals) in the **Congruent** condition at 8 different latencies between 200 and 550 ms post stimulus in Experiment 3. Red hue represents positive voltages, and blue hue, negative voltages. The voltage scale bar is presented on the right side of the figure.

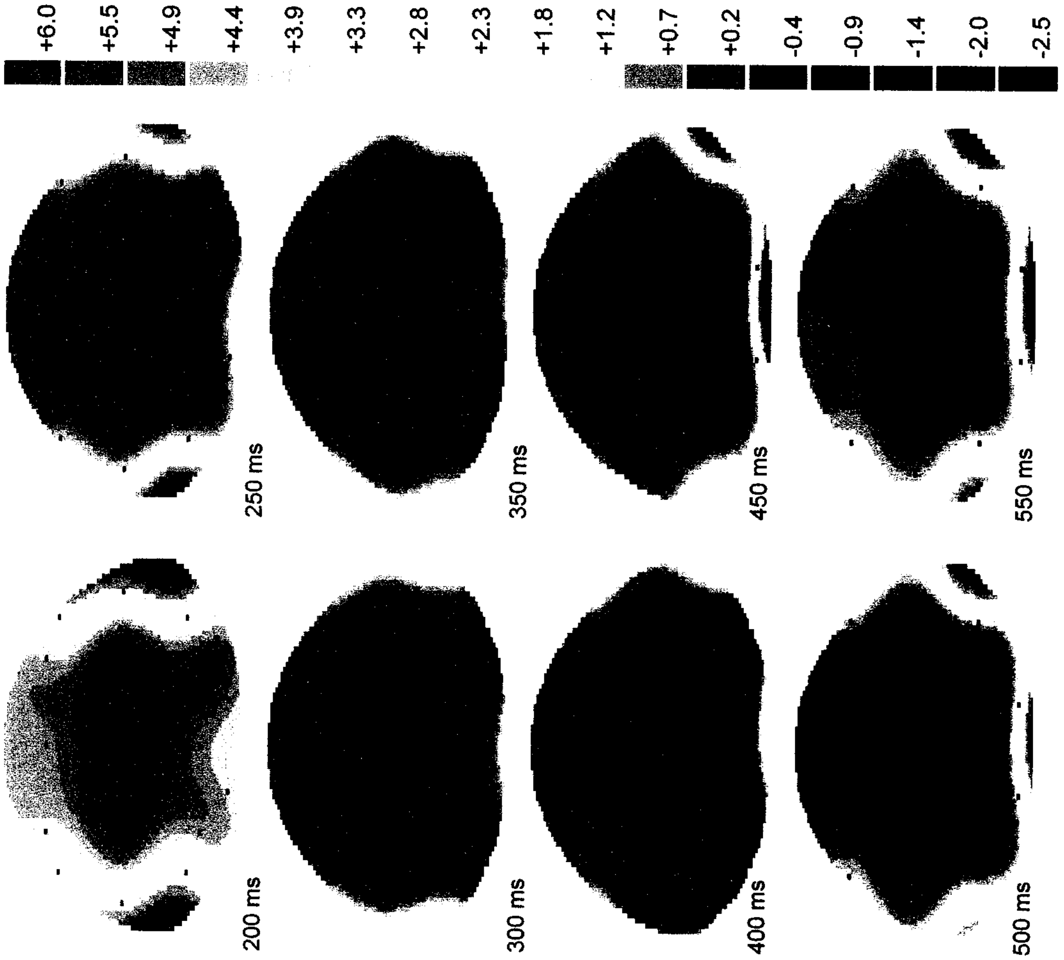
200 ms
250 ms
300 ms
350 ms
400 ms
450 ms
500 ms
550 ms
+6.0
+5.5
+4.9
+4.4
+3.9
+3.3
+2.8
+2.3
+1.8
+1.2
+0.7
+0.2
-0.4
-0.9
-1.4
-2.0
-2.5

Figure 34

Figure Caption

Figure 35. Individual waveforms of a representative participant at 17 recording sites for the **Congruent**, **Onset**, **Rime**, **Tone**, and **Incongruent** stimulus conditions presented in Experiment 3. For the display purpose, the low pass filter was set at 15 Hz for the ERP wave forms in this figure.

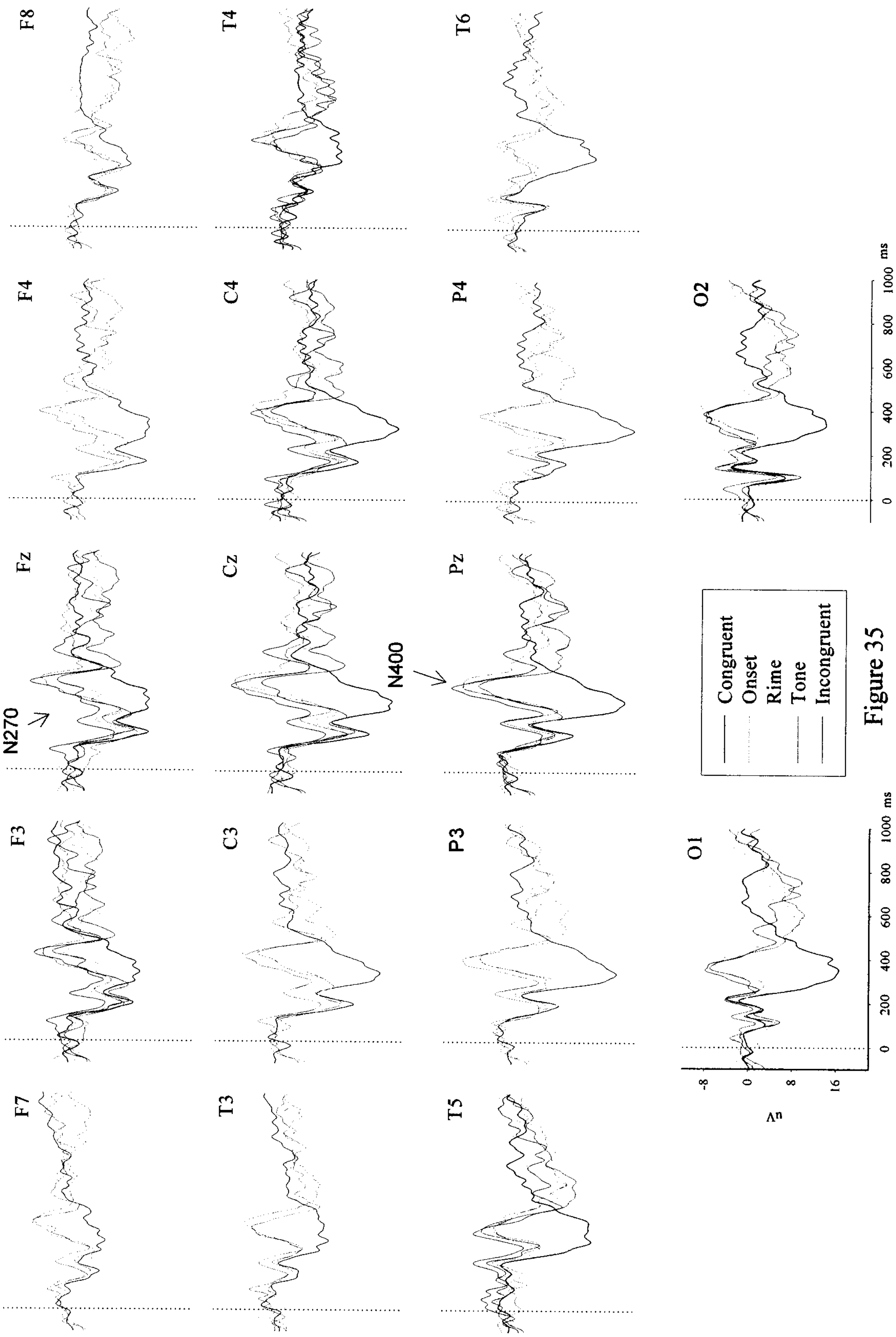
F7
F3
Fz
F4
F8
T3
C3
Cz
C4
T4
T5
P3
Pz
P4
T6
O1
O2
N270
N400
uV
-8
0
8
16
0
200
400
600
800
1000
ms
Congruent
Onset
Rime
Tone
Incongruent

Figure 35

Discussion

The purpose of this experiment was to define the roles that word prosodic and segmental phonology may play in reading Chinese characters. The present experiment enabled a detailed look at the influences of onset, rime, and tone information on Chinese reading comprehension using ERP. The ERP data revealed that segmental and tonal variations modulated ERP waveforms amongst the five conditions. Two features of the present study are particularly important and will be discussed below.

The first is that a negative ERP component whose peak occurred between 350-400ms after target character onset and which had a posterior scalp distribution was observed in the four non-congruent conditions. This response, the N400, was largest in the **Incongruent** condition and appears indistinguishable from the classic N400 elicited by semantic incongruity in the English literature (e.g., Kutas & Hillyard, 1980a,b). Smaller but still strong N400 responses were observed in the **Rime**, **Tone**, and **Onset** conditions, suggesting its sensitivity to segmental and tonal phonology manipulations – either directly or indirectly. For example, the N400 showed almost identical morphology at parietal sites for both segmental and tonal changes. This finding indicated the N400's insensitivity to the nature of the phonological incongruity (segmental and tonal) encountered during reading.

Kutas and Hillyard (1980a) first demonstrated that the N400 is associated with semantic processing. Subsequent studies further argued that the N400 amplitude can be constrained not only by word expectancy engendered by context but also by automatic activation that is restricted by the degree of association between individual words in the mental lexicon (Kutas & Hillyard, 1984). The influence of contextual constraints on

N400 has also been supported by a number of other English ERP studies (e.g., Connolly, Stewart, & Phillips, 1990; Van Petten & Kutas, 1990). Taken together, these findings provide convincing evidence that the N400 amplitude reflects the ease with which the semantic code of a word is accessed. Put differently, any cognitive processes that facilitate or impede the activation of semantic codes of a word would in turn influence N400 amplitudes. In line with this view, the existence of segmental and tonal influences on the N400 in this study suggests that phonological information comes into play in Chinese reading comprehension. Furthermore, the similarity of the N400 to discrete tonal and segmental phonological incongruity provides further evidence that the N400 does not reflect form-based representations per se (Praamstra & Stegeman, 1993).

The magnitude of the effects of phonological priming on the N400 in this study is remarkable because there is considerable debate in the English literature as to the reliability of the phonological effects on lexical access in reading, let alone how it should be interpreted. In fact, a few English studies have demonstrated that the visual N400 can be modulated by phonological similarities between written words. In a rhyme-judgment task, non-rhyming words elicit larger N400s than rhyming ones (e.g., Rugg, 1984, 1985). Rugg argues that most of the rhyming pairs used in their studies were orthographically dissimilar (e.g., moose-juice) so that the reduced N400 amplitude results from rhyme priming effects. There is a caveat, however: phonological effects on the N400 can be influenced by the experimental tasks that place explicit emphasis on phonological activation. In this regard, conscious awareness of the phonological relationship between the prime and the target may change the way that words are naturally processed.

Polich and colleagues (1983) conducted a study using both rhyme and word form similarity judgment tasks and found that phonological priming of the N400 was observed only for rhyme similarity judgments but not for word form similarity judgments. Based on these findings, it was concluded that the phonological priming of the English N400 is clearly task dependent. However, the argument that the sensitivity of the visual N400 to phonological variables occurs only in explicitly rhyme-orientated tasks is made less likely by a study that made use of pseudohomophones (Newman & Connolly, 2003). Using a silent reading sentence comprehension task, Newman and Connolly (2003) visually presented English sentences ending with pseudo-words whose pronunciation was the same as the highly expected sentence ending words (e.g., 'The ship disappeared into the thick *phog*'). It was found that the N400 elicited by such pseudo-homophones (e.g., 'phog') was absent relative to the amplitude of the N400 to traditionally incongruent sentence-ending words (e.g., 'The piano was out of *nose*'). Unfortunately, given the possibility that different mechanisms may be applied during pseudo-word processing such that unnatural language processing is highlighted, the mechanisms underlying the phonological effects on N400 in reading tasks remain less than clear-cut.

This study represents the first ERP study that presents compelling evidence on the onset, rime and tone effects on the N400 during reading comprehension. The clear effects that segmental and tonal variations have on the visual N400 suggest that phonological cues are actively involved in Chinese reading. The target character in the **Tone** condition, for example, shares the same rime and onset attributes as the accurate proverb-ending character that is highly expected. On the basis of the Cohort model (Marslen-Wilson, 1984), the node corresponding to the highly expected Chinese character can be activated when preceding contextual information provides adequate cues. The activated node then sends activity (via the link) to any other nodes that share related features in both semantic and form-based (phonology or orthography) domains. In this regard, the pre-activated node can be very easily activated and pass over its threshold level when the corresponding character is presented. According to the spreading-activation theory proposed in the Cohort model, the influence of the preceding contextual information on word processing occurs long before the word can be recognized.

Of particular importance is that the N400 in response to the **Onset** targets presents almost identical morphology to the N400 to the **Rime** and **Tone** targets. This insensitivity of the N400 to distinct phonological incongruity provides converging evidence that the apparent phonological effects on the N400 reflect the consequence of phonological processing which facilitates subsequent semantic activation in the lexical processing sequence. In line with this view, it is worth emphasizing that the neural mechanisms manifested by the N400 do not represent phonological recoding per se.

Given the sensitivity of the N400 amplitude to phonological variables, two possible explanations of how the preceding contextual information might facilitate the recognition of upcoming target characters in this experiment can be offered. The first possibility is that phonological contextual facilitation derived from the pre-activated node corresponding to the accurate proverb ending character acts on the processing through which orthographic codes are transformed to phonological codes. In turn this process facilitates subsequent semantic activation (sub-lexical top-down phonological priming). The second possibility is that phonological contextual facilitation directly activates the node corresponding to the target so that the node that shares tonal or segmental phonological attributes with the accurate ending character is raised from its normal resting level, the consequence of which facilitates subsequent semantic activation (lexical/post-lexical top-down phonological priming). Given that the N400 is more tied to the functional locus of semantic activation, however, one cannot identify which of these two possibilities sub-serves the phonological priming reflected in the N400.

Fortunately, the second striking feature of the ERP data in this study provides helpful information as to the nature of the phonological priming effects on the N400. The second feature is that an earlier occurring temporo-occipitally distributed negative component peaking about 300ms after the character onset showed distinct amplitude and peak latency differences amongst the **Tone**, **Rime**, **Onset**, and **Incongruent** conditions. This negativity, identified by virtue of its latency and distribution as the N270, and its conditional variations proved important in interpreting this experiment (Forbes, 1993, 1998; Newman, 2000). The data revealed that the N270 to the **Incongruent** targets was much larger in amplitude than that to the other three phonological change targets. The most notable finding was that the negativity to the **Rime** and **Tone** targets peaked earlier than that to the **Onset** and **Incongruent** targets.

The temporal separation and the topographic distinctions between the N270 and the N400 suggest that these two components reflect distinct neural processes. The sensitivity of the N270 peak's latency to individual segmental and tonal variations demonstrates that the N270 is sensitive to distinct aspects of phonological (segmental and tonal) analysis. This responsiveness to these factors is in stark contrast to the lack of differential sensitivity of the N400 to discrete phonological manipulations (as reflected across the four conditions having semantic incongruities). It suggests that the segmental and tonal effects on the N270 are probably specific to neural substrates of segmental and tonal phonological encoding. These results can be further construed as compelling evidence that the parietally distributed N400 and the temporo-occipitally distributed N270 are likely to be interactive but temporally sequential components.

The sensitivity of the N270 to phonological manipulation in this experiment seems at a first glance inconsistent with previous reading studies (Forbes, 1993, 1998; Newman, 2000) demonstrating that the N270 is associated with orthographic expectation that is independent of phonological recoding. However, taking a closer examination of the data presented in the contemporary studies, one cannot deny that the N270 appeared associated with phonological variables as well. In the first experiment of Forbes' (1998), participants were required to judge whether visually presented sentence-ending words were semantically correct. Terminal words of highly constrained sentences were manipulated in three conditions. Sentence endings consisted of: 1) semantically expected words (Congruent); 2) words that were homophonic to highly expected sentence-ending words but semantically incongruent to the sentence context (FOIL); and, 3) words that were semantically incongruent to the sentence context and which were homophonic to semantically unexpected sentence-ending words (Incongruent). The result of this experiment revealed that the N270 was reduced in the homophonic Foil condition. In the second experiment, Forbes took into consideration the magnitude of similarity in orthography between non-congruent targets and highly expected sentence endings. Terminal words were divided into: 1) highly expected sentence endings (Congruent); 2) words that were semantically incongruent to sentence contexts but orthographically similar to highly expected sentence endings (OS); 3) words that were both semantically and orthographically incongruent to sentence contexts (OD); and, 4) homophone foils (FOIL). The results of the second experiment demonstrated that the N270 elicited in the OS and FOIL conditions was smaller in amplitude than that in the OD condition. Although this finding indicated that the N270 is sensitive to the magnitude of orthographic variation, there is no convincing evidence that the N270 is not also sensitive to the orthography to phonology transformation. Parallel to these studies, Newman (2000) argued that the N270 is sensitive to orthographic incongruity. However, she also reported that the N270 is reduced in response to the phonologically congruent targets. Although she leaves open the issue of whether the N270 reflects the orthography to phonology transformation, she does not restrict the neural processes sub-serving the N270 to orthographic encoding only. It is worth emphasizing that if phonological activation is involved in reading, it may occur at the same moment as the activation of orthographic encoding (Tan & Perfetti, 1998). In the present experiment, given the opaque relationship between phonology and orthography of a Chinese character, the target characters in the **Onset**, **Rime**, and **Tone** conditions were similar to the accurate proverb ending characters in terms of their phonological features only (e.g., the target and the correct proverb ending in the **Rime** condition had the same onset and tone attributes but differed in both word form and meaning). In this view, it is reasonable to argue that the amplitude of the N270 reflects the ease with which phonology is recoded according to orthographic information as well as contextual cues. Meanwhile, although the peak latency of an ERP response does not necessarily equal the maximum number of synchronous neural clusters engaged in a certain process, the two are probably connected (e.g., Kutas & Van Petten, 1994). In this view, the different peak latencies of the N270 in response to discrete phonological incongruity suggest distinct onset, rime, and tone phonological recoding.

In fact, previous reading research has suggested that the N270 may be associated with the orthography-to-phonology transformation. It has been recognized that reading difficulties found in dyslexia populations is not exclusively derived from deficits in visual processing. In particular, a subtype of dyslexia named phonological dyslexia is believed to result from a core deficit in phonological awareness (Shaywitz, 1996; Wagner & Torgesen, 1987). It is argued that if phonological recoding in reading is not mastered, reading and writing cannot achieve a proficient level. By this account, the deficit in phonological processing in phonological dyslexia might be expected to alter the normal morphology of the N270 or even eliminate its appearance if the neural processes manifested by the N270 reflect phonological recoding in reading. Indeed, Helenius, Salmelin, Service, and Connolly (1999) conducted a reading comprehension study using magneto-encephalography (MEG) and ERP and confirmed this proposition. In this study, the authors compared the activated brain areas engaged in reading between normal and phonological dyslexic adults. Stimuli used were based on previous work (Connolly & Phillips, 1994) and comprised sentences ending with: 1) highly probable words; 2) semantically congruent but not highly probable words; 3) semantically incongruent words but which shared the initial letters with the highly probable words; and finally, 4) incongruent words in both semantic and orthographic domains. They found that the N400 was reliably elicited in the dyslexic group although the strength of the N400 was apparently decreased in contrast with the normal group. However, with the ERP recordings it was clear that the healthy control participants showed a three-phased P200-

N270-P300 complex in the reading task while the dyslexic group failed to show the N270 element of this complex response. The absence of the N270 in dyslexic group was interpreted as compelling evidence that the N270 is sensitive to phonological recoding.

There is evidence that this N270 seen in these studies may be related or identical to another component that has been linked to phonological processing. Bentin, Mouchetant-Rostaing, Giard, et al. (1999) conducted two experiments to examine the neural correlates of phonetic/phonological processing. In a phonological/phonetic decision task, participants were required to look at two types of stimuli: one consisted of words and pseudo-words rhyming with the French word *vitrail*, the other type of words were pseudo-words and strings of consonants that did not rhyme with *vitrail.* A negative component peaking about 320ms after the stimulus onset was clearly elicited by pronounceable stimuli but not by unpronounceable stimuli. This component was distributed over temporal areas. Given that the rhyme task required activation of the phonetic codes of stimuli, this negativity was associated with phonetic recoding. In a following phonological/lexical processing task, a negative component peaking about 350 ms after the stimulus onset was conspicuously elicited by the phonologically legal but not by phonologically illegal stimuli. The distribution of this negativity appeared similar to that of the N320. What is different is that the N350 appeared broader and had a more temporo-parietal distribution than the N320. The authors concluded that the difference between the N320 and the N350 in the timing and scalp distribution was attributable to different weightings of phonetic and phonological recoding due to different task demands (the N320 being seen in the rhyme task whereas the N350 was observed in the lexical decision task). Overall, given the similar latency as well as scalp distribution of the

N320/N350 to the N270 observed in the present experiment, the components may come from the same family of ERP potential that is associated with phonological recoding in reading.

Returning to the major issue regarding the nature of phonology in reading, one may first wonder about the nature of segmental and tonal processing sub-serving the temporo-occipitally distributed N270, which seems to occur prior to, but highly overlap with, the semantic activation reflected by the N400. As introduced in the hypothesis section of this chapter, there are two ways of recoding phonology from print. The first "assembled phonology" is referred to as the process by which phonology can be recoded according to orthographic information (the orthography-to-phonology transformation). In contrast, the second "addressed phonology" mainly emphasizes the process through which phonological representations of a word are retrieved, in an abstract impoverished fashion, from activated 'corresponding' node in the mental lexicon that is directly addressed by orthographic representations (Patterson & Coltheart, 1987). In line with this theory, the sensitivity of the N270 to detailed segmental and tonal variables can be construed as assembled phonological processing that is transformed from orthographic cues, which confirms the interpretation of the N270 as being associated with phonological recoding. Given the relationship between the N270 and the N400 in timing, it seems reasonable to interpret the phonological effects on the N400 as the consequence of contextual facilitation on the orthography-to-phonology transformation.

Taken together, the impact of phonology on both the N270 and the N400 observed in this reading experiment provides strong evidence against the "direct access" theory that phonology does not come into play in reading. The findings in the present experiment seem at first sight to lead to the "phonological recoding" theory underscoring that phonological recoding is the default process used to access semantic meaning. The hazards of regarding phonology as a default are clear when the logic is taken to its extreme. Of critical relevance is that the overlapping nature of the N270 with the N400 can not be taken as explicit evidence that orthography to phonology transformation and semantic activation are processed in a sequence. There is possibility that semantic processing does not need to wait for the completion of phonological recoding; semantic meaning of a word may be accessed not only by phonological recoding but also by activated orthographic codes as well as contextual information (such as syntactic). Such a pattern may be more congruent with a cascade (Seidenberg & McClelland, 1989) than with a serial-processing model of word recognition (Coltheart, Patterson, & Marshall, 1980). As a matter of fact, given the view that linguistic working memory is phonetically organized (Baddeley, 1982), it seems reasonable to argue that one of the functions of phonological recoding in reading may be to enable the character to be held in working memory and its context facilitates further Chinese sentence comprehension.

Parsimony requirements would appear to force an interpretation of the present ERP findings as providing compelling evidence to swing the pendulum toward an interest in the parallel-distributed model (Seidenberg & McClelland, 1989). According to this model, phonological transformation occurs in reading and serves to constrain semantic activation along with orthographic information. Most importantly, the model emphasizes that two processing routes are not independent of, but actively interact with, each other.

Chapter Five

General Discussion

The goals of the three experiments in the thesis were to determine 1) whether word prosody in both English and Chinese speech plays a role in constraining semantic activation of a spoken word/character; 2) whether the processing of continuous acoustic input of a word/character proceeds in a strictly sequential manner or in a relatively 'time-independent' fashion; 3) whether lexical activation proceeds in an all-or-none fashion - that is, to explore what function sub-lexical phonology (discrete syllables in multi-syllabic words or segments in monosyllabic Chinese characters) may play during whole word recognition; 4) whether segmental phonology is involved in accessing semantic meaning in silent reading; and, 5) whether word prosodic information is recoded as part of phonological specification during Chinese reading comprehension. The three experiments combine to provide converging evidence in favor of the argument that word prosody exerts important effects in both English and Chinese languages, which implies a language universal mechanism underlying prosodic phonological processing. The findings of the three experiments have important implications for both speech and reading models of word recognition. In addition, the present research implies that ERP are of great use when the timing and sequencing of cognitive processes are relevant to the issues of prime concern.

5.1 Word prosodic influences in lexical processing

The present research investigated whether word prosody has any influence on lexical processing in both speech and reading. The findings of the first English speech experiment demonstrated that both the PMN and the N400 that respectively reflect phonological encoding and semantic analysis were evidently modulated by lexical stress manipulation. The amplitude of the PMN was determined predominantly by stressed syllables in contrast with unstressed syllables. Although the amplitude of the N400 was responsive to both primarily stressed and unstressed information, it was heavily influenced by primarily stressed syllable encoding. That is, the weighting of the influence of sub-lexical phonological cues on semantic constraint varies according to their stress status. Equally, the results of the second Chinese speech experiment indicated that both the PMN and the N400 were sensitive to lexical tone manipulation. As well as the sensitivity of the PMN's amplitude to segmental and tonal variables, its peak latency was also modulated by discrete segmental and tonal phonological information. The N400's amplitude appeared more sensitive to segmental than tonal phonology, although the sensitivity of the N400 to tonal information was still observed. Based on these findings, it seems logical to argue that word prosody is represented in the mental lexicon. The coding of word prosodic information is a crucial part of lexical processing. Given the importance of word prosody in both English and Chinese spoken word processing, one tends to conclude that the influence of word prosody on spoken word recognition serves as a language universal mechanism.

The results of the first two experiments in this thesis seem in accord with several lines of behavioral research indicating that word prosodic information contributes to auditory word processing. Quite early in the field of psycholinguistics, there has been compelling evidence that stress pattern can facilitate auditory word recognition. In a recognition task, Robinson (1977) presented listeners with nonsense bisyllablic stimuli that varied with stress pattern (e.g., BIsev, juBIM, JUlem). The results of this study demonstrated that foil stimuli that preserved the stress of individual syllables from the original presentation (e.g., EIlem, when BIsev and JUlem had been presented) induced a large number of false alarms in contrast with the foil stimuli that changed a syllable's stress (e.g., biLEM). In a speech perception experiment, Darwin (1975) required listeners to repeat one of two simultaneous messages that were presented one to each ear. It was found that the cognitive system was sensitive to the organizing function of the prosodic pattern: listeners tended to switch their shadowing to the wrong channel to follow prosodic continuity even if this switch made the whole message that they repeated nonsensical. It is concluded that when prosodic continuity and semantic continuity conflict, listeners tend to give more attention to the former. A contemporary study conducted by Speer, Crowder, and Thomas (1993) further replicated the findings of Darwin (1975).

Parallel to this evidence, another line of behavioral research directly taps on the processing of intact prosody to examine its role in speech perception (Aull, 1984; Huttenlocher & Zue, 1983; Slowiaczek, 1990; Waible, 1988) and retrieval (Levelt, 1989; 1992). Results from this research underscore the influence of word prosody on word recognition. For example, Slowiaczek (1990) conducted several behavioral tasks to determine the influence of lexical stress on the processing of auditory single words. In three shadowing tasks, it was reliably observed that correctly stressed items facilitated the time listeners took to name them in contrast with incorrectly stressed items. Similarly, in a lexical decision task, correctly stressed words were recognized much faster than incorrectly stressed words. Based on these findings, it was concluded that the code of lexical stress reflects an important aspect of word representation in the mental lexicon. In a seminal behavioral study, participants were asked to categorize an ambiguous initial stop consonant (/d/ vs. /t/) in either DIgress-TIgress (tigress being the real word by virtue of correct stress) or diGRESS-tiGRESS (digress the real word) (Connine, Clifton, & Cutler, 1987). The results of this study demonstrated a significant effect of stress-determined lexical status in that listeners were more likely to report an ambiguous initial segment that resulted in a real word (diGRESS or TIgress) than one that resulted in a non-word (DIgress or tiGRESS). It seems evident that word prosodic information plays a role in resolving phonetically ambiguous words.

Likewise, a group of studies in the stress-based language of Dutch demonstrated the use of stress information in word recognition. Using minimal stress pairs such as SERvisch vs. serVIES in a sentence context, listeners were asked to discriminate two successively presented Dutch stimuli with partially presented phonetic transcription (Jongenburger & van Heuven, 1995). It was found that listeners could accurately differentiate two minimal stress Dutch word pairs as soon as the whole of the initial syllable and part of the following vowel quality were provided. A follow-up study (Jongenburger, 1996) further demonstrated that listeners could accurately discriminate between two Dutch words that shared a segment-identical initial syllable but different stress pattern, such as ORgel and orKEST, or a minimal pair such as SERvisch and serVIES, when presented with only the first syllable. In all, it seems likely that stress based languages make use of stress information during word recognition.

Further evidence of the importance of stress information in spoken word processing comes from observation of the effects of ill-formed prosody (such as, misplacement of stress or accent-induced stress distortion). For instance, it was expected that incorrect stress, derived from foreign accents in natural settings, would violate word recognition if stress information were of relevance in segmenting the speech and initiating lexical search. Bansal (1966) presented native English listeners with English words spoken by an Indian speaker. The results of this experiment were indeed consistent with what was expected. In other research using incorrectly stressed English words, it was also found that listeners tended to recognize incorrectly stressed English words much more slowly than correctly stressed words (Bond & Small, 1983; Cutler & Clifton, 1984). Similar results have been found for Dutch by van Heuven (1985).

As stated in the introduction, tone based languages that boast a highly restricted range of syllable structures contain a relatively rich tone repertoire. Although lexical tone research is not extensive in contrast with the literature on lexical stress, there are a number of studies, the majority of which are spoken word recognition studies, attempting to get to grips with the nature of lexical tone processing. It has been argued that lexical tone constrains lexical processing in tone based languages. As stated previously, a perceptual right ear advantage (REA) in dichotic listening experiments with lexical tone manipulation demonstrates left hemisphere dominance in lexical tone processing, which further indicates the linguistic nature of lexical tone. In this regard, it is worth noting that the N400 in response to **Tone** targets in the Chinese speech and reading experiments of this thesis has a left hemisphere distribution. This finding provides further evidence on the linguistic nature of tone information.

Further evidence of the importance of tone as a linguistic element comes from behavioral studies of word identification and lexical decision. For example, Ching (1985,

1988) asked Cantonese listeners to identify lip-read Cantonese words and found that the identification rate became higher when the fundamental frequency (F_0) was presented in the form of high-pass filtered pulses synchronized with the speaker's larynx frequency. In an auditory lexical decision task, lexical priming effects of tonal overlap between a prime word and its target were observed reliably. More to the point, the magnitude of the tonal effect seems similar to that of the segmental effect on lexical decision (Cutler & Chen, 1995).

In conclusion, it is reasonable to argue that word prosodic cues are made use of during auditory word processing, which are independent of segmental phonological information. However, this argument has not gone without controversy. As indeed in most of the world's stress based languages, lexical stress is presented as a minimally distinctive attribute between words. Taking English language as an example, pairs of semantically unrelated but homo-segmental words that differ merely in stress pattern are fewer than a dozen; they include FORbear—forBEAR, FOREarm-foreARM, REtail-reTAIL, INcite—inSIGHT. Based on the rarity of such minimal stress pairs, it is argued that lexical stress is not a very valuable source in constraining lexical search. However, given the aforementioned evidence that a multi-syllabic word with the presence of its initial syllable was much more easily identified if the stress status of that syllable was provided (Connine, Clifton, & Cutler, 1987; Connine, Blasko, & Titone, 1993), the view based on the rarity of minimal stress pairs is not self-evident.

Nonetheless, given the fact that, although rare, such minimal stress word pairs do exist in English, Cutler (1986) manipulated those word pairs in a cross modal priming experiment in an attempt to clarify the issue of whether prosody is part of the lexical access code. In the cross-modal priming task, the participant listened to a sentence; at some point during the sentence, a visually presented string of letters (target) was displayed on the screen. The task demand for the listener was to judge whether or not the visually presented target was a word. A putative homophone effect observed in such a task is that if the visually presented target appears right after a homophone in the sentence, the speed of the listener's response is faster when it is related to any meaning of the homophone. For example, both targets "ant" and "spy" are detected faster than the control word "desk" if the sentence bears the homophone "bug" (Swinney, 1979). In line with this view, Cutler (1986) hypothesized that if prosody did constrain lexical search as segments do, the target "ancestor" should not be primed when the sentence contained the word "forBEAR" rather than "FORbear". The findings of her experiment, however, demonstrated that word pairs like *forbear* behaved functionally as homophones: both FORbear and forBEAR improve the lexical decision on words related to either of them (e.g., ancestor, tolerate). On the basis of these findings, she argued that the lexical codes for these two word forms are not distinguishable between each other. In other words, it was proposed that the lexical stress pattern does not function in constraining lexical search to exclude one of the two possibilities.

In this regard, it seems likely that lexical stress, although available at the acoustic level, does not play a role in constraining lexical search. Consistent with this argument are some data from experiments on the processing of tone based Chinese language indicating that the tonal information has no function in word recognition. For example, in a categorization task, non-word consonant-vowel syllables were manipulated in terms of segments (consonant or vowel) or tone dimension (Repp & Lin, 1990). It was found that the tonal categorizations were responded to much more slowly than the segmental decisions. In an auditory lexical decision task, Cutler and Chen (1997) further reported that a non-word that differs from a real word only in tonal dimension was easily mistaken as a real word. Using a same-different judgment task, Cutler and Chen note that it is much slower and more error-prone for Cantonese listeners to make a judgment when two syllables differ only in tone domain whereas this is not the case when two syllables differ in segmental dimensions. Based on these findings, it seems that word prosodic information is only available at the acoustic level, which fails to function in constraining lexical search.

Caution needs to be exercised before embracing these conclusions, however. It is worth mentioning that the logic behind the argument addressed by Cutler (1987) is that lexical access is processed in an all-or-none fashion. FOREarm and foreARM correspond to either exactly the same lexical access code (to indicate that word stress has no function) or completely unrelated lexical access codes even though there is segmental phonology overlap between them. This logic is not tenable, however. As stated in the first and second experiments of this thesis, segmental phonology of the target word/character, which has overlapping features (syllables in the first English experiment and onsets and rimes in the second Chinese speech experiment) with a contextually primed word/character, provides significant facilitation effects on semantic analysis. These sub-lexical phonological effects on the semantic constraints observed in these two experiments suggest that lexical coding may not be all or none but rather indicates a probabilistic feature of lexical processing: any partial phonological information can to some degree facilitate the whole word processing. In this regard, it is reasonable to argue that the different coding of lexical stress which comes into play during lexical processing may be overshadowed by the common segmental features of the minimal stress pairs, the result of which gives rise to the homophone effect. In line with this view, the observation of homophonic-like effects on minimal stress pairs has not undermined the reliability of the view that word prosody constrains lexical processing. Rather, it has deepened our understanding of the nature of lexical processing.

Equally, the "more error-prone and slower responses" observed in the Chinese performance studies are not evidence enough by themselves to undermine the view that lexical tone contributes to lexical processing. As stated previously, the behavioral measure itself is critically influenced by the postlexical characteristics of a process. In this regard, the safe interpretation of findings employing behavioral measures is limited by an inability to disentangle lexical processing and post-lexical operations. In line with the findings observed in this thesis demonstrating the influence of lexical tone during Chinese character processing, it is reasonable to argue that the "more error-prone and slower responses" to tonal variables may primarily reflect the influence of lexical stress information on certain post-lexical sub-processes. For instance, lexical stress information is used to check an item after access has been completed. In particular, it has been shown that the neighborhood density for homophones that have a common segmental phonology but vary in the tonal dimension is very intensive in the Chinese language. On the basis of the Modern Chinese Frequency Dictionary compiled in 1986, 4574 Chinese characters make up full coverage of a 1,800,000-character corpus in daily use. The most notable is that when tone variables are disregarded, there are only about 420 distinct segmental-defined syllables in Chinese. That is, there are almost 11 Chinese characters that share the common segmental phonology (Tan & Perfetti, 1997). Taking this high homophonic density into consideration, the "more error-prone and slow response" to tonal manipulation can be convincingly accounted for by the relative difficulty in discriminating between homophonic characters identical in segments but different in tone during post-lexically sub-processes (e.g., decision making). On this view, the "more error-prone and slow response" may not be a counter example against the view that lexical stress is represented in the mental lexicon. Instead, such data may, from a post-lexical position, highlight the impact of lexical tone on language processing.

In short, the present findings in this thesis along with the behavioral data in the literature provide compelling evidence that word prosody is decoded during lexical processing. These data also indicate that word prosody is involved in lexical processing in a way that is not specific to only a single language but indeed reflects the common aspects of language processing between two very different languages such as English and Mandarin. However, it is worth emphasizing that, although the involvement of word prosody in constraining lexical search implies a language universal, the means and procedures by which word prosody influences lexical processing may vary with different languages insofar as the prosodic attributes of the languages themselves differ. For example, lexical stress in the first experiment functions at the level of monitoring the initiation of lexical attempts and manipulating the weighting of syllable-sized sub-lexical phonology in constraining semantic meaning. In contrast, lexical tone in the second experiment does not show exclusively the same function given the phonological features of the Chinese language itself (the strictly monosyllabic structure for each Chinese character).

Given the fact that lexical tone plays a role in spoken word processing, one may wonder whether word prosody exerts some effects on reading comprehension. The findings of the third Chinese reading experiment imply that lexical tone is encoded during visual Chinese character processing and along with segmental phonology plays a crucial role in visual character recognition. Although requiring replication, the evidence in the third experiment that lexical tone is exploited in reading processing provides an opportunity for further exploring the mechanisms underlying word prosody in reading comprehension. Meanwhile, these findings in turn further corroborate the view that word prosody is part of a lexical representation in the mental lexicon.

5.2 Implications for Speech Models of Word Recognition

5.2.1 The directionality of lexical processing:

The present research questions the position held by the TRACE model that the mapping of the acoustic input onto mental representations of word forms is non-directional (Elman & McClelland, 1986; McClelland & Elman, 1986). The proponents of the TRACE model argue that the directionality of the mapping is less relevant. What is more important, on the basis of the TRACE model, is the overall goodness of fit between the acoustic input and a selected lexical representation, in contrast with the input's total amount of overlap with other potential candidates. However, the predominant role of primary stressed syllables in lexical processing in the first experiment implies that this is not the case. Rather, the salience that primary stressed information has for lexical processing suggests the intrinsically directional properties of the mental representation of lexical forms.

The present research findings are not in line with the claim at the heart of the Cohort model that lexical activation progresses in an "exclusively" sequential fashion (Marslen-Wilson & Zwitserlood, 1989). It is evident that speech input has intrinsic directionality in time. The speech signal is unfolded over time, moving successively from beginning to end, with limited overlap between phonemes (co-articulation). Following this account, proponents of the Cohort Model stress that lexical processing must follow the deployment of speech in a time-shadowing fashion. The acoustic information arriving first in time is always processed first. The mental representation of a lexical form is seen as "a sequence of independent representational units (such as segments), strung together like a string of beads" (Marslen-Wilson & Zwitserlood, 1989, p.584). Although this description sounds tenable on intuitive grounds, the strictly sequential claim has been repeatedly questioned by an increasing number of behavioral studies. As stated in the introduction, there exists considerable evidence that word initial information does not play a predominant role in lexical processing, relative to later portions of the word (e.g., Connine, Blasko, & Titone, 1993; Slowiaczek, Nusbaum, & Pisoni, 1987; Radeau, Morais, & Segui, 1995). Also, there is ample evidence that word onsets are not always accessed before later portions of a word (e.g., Cutler, 1992). Yet, given the critical sensitivity of the behavioral performance itself to post-lexical operations, despite the abundance of contradictory data, the influence that the sequential processing claim has particularly on the contemporary neuro-psycholinguistic research, seems to be as lively as ever.

In contrast with behavioral measures, ERPs are of great use to provide reliable information pertaining to the timing and sequencing of cognitive processing engaged during language. With this in mind, the present research provides direct evidence that there is an intrinsic directionality in lexical processing; nonetheless, the directionality is not determined by the temporal sequential nature of speech but appears dependent upon the position of primary stressed syllables in a word. Simply put, lexical processing gives its priority to primary stressed syllables (irrespective of their positions in words) instead of temporally early input. With initial-stress words (e.g., "**won**der"), lexical activation may proceed in a sequential fashion at the syllabic level. However, with non-initial-stress words (e.g., "i**dea**", "po**lite**"), lexical activation begins on the primary stressed syllable in a non-sequential manner. The findings of the first experiment further suggest that a complementary process on initial-unstressed syllables concurs with the process dominated by primary stressed information (the parallel hypothesis proposed in the Chapter two). As a matter of fact, the directionality view that takes word prosody into account also sounds convincing from the speech segmentation stand point: if word onset information is always highlighted in speech stream it may well be optimally efficient for lexical activation to progress in a sequential fashion as the acoustic input unfolds over time. However, as stated in the introduction, word boundary is not specified in continuous speech. In this account, the stress guide strategy can be effectively seen as an optimal approach to conduct lexical search (Cutler, 1992).

In an attempt to get a better understanding of the nature of the directionality in lexical processing, the complexity provided by the above view for multi-syllabic English word processing brings another issue into focus. That is, does lexical processing of monosyllabic words proceed in a strictly sequential way or does it progress in a 'time-independent' fashion? It is noteworthy that the peak latency of the PMN in the **Rime** condition in the Chinese speech experiment appears earlier than that in the **Onset** condition. On the basis of the assumption that the different peak latencies of the PMN corresponding to onset and rime manipulations may reflect temporally distinct aspects of the processing sequence, one may speculate that rime encoding in monosyllabic Chinese characters proceeds earlier than onset encoding. If this speculation holds a grain of truth, the strictly sequential strategy even fails to be exploited in monosyllabic word recognition.

Compatible with the aforementioned Chinese findings, there is a considerable body of evidence that the rime and the onset for monosyllabic English words present distinct patterns of phonological overlapping effects. It has been repeatedly reported that onset phonological overlap between spoken monosyllabic probe words and preceding primes (e.g., Black—Bless) in the shadowing task produces little facilitation effect at short lags, for instance, with a 20 ms interval between prime and probe (ISI). However, with the increase of the ISI, a facilitation effect is observed initially and then an increasing inhibition effect of onset phonological overlap surfaces (Goldinger, 1999; Hamburger & Slowiaczek, 1996; Monsell & Hirsh, 1998; Radeau & Colin, 1996; Slowiaczek & Hamburger, 1992; Radeau, Morais, & Segui, 1995). Given that these results cannot be replicated when preceding primes are non-words (Radeau, Morais, Dewier, 1989; Slowiaczek & Hamburger, 1992), the inhibition effect is thought to be influenced by sub-processes at the lexical and post-lexical stages. It is surmised that the activation of onset phonological representation may occur close to lexical-level processes.

Different from the onset overlap effect, the rime phonological overlap between auditory target and auditory prime produces a reliable facilitation effect at the short lag (e.g., 20-ms ISI) in both shadowing task and lexical decision task (e.g., Radeau, Morais, & Segui, 1995; Slowiaczek, McQueen, Soltano, & Lynch, 2000). Of particular interest is that the facilitation effect is equally observed with both word and non-word primes. Given the assumption that non-word primes would not activate monosyllabic lexical representations to any significant degree (Marslen-Wilson, Moss, & van Halen, 1996), the equivalent effect with word and non-word primes implies that the facilitation is more susceptible to perceptual processing (pre-lexical). Put differently, one may argue that activation of rime phonology in auditory English monosyllabic words may occur temporally earlier than that of onset phonological representation.

Taken together, the present research along with the behavioral findings cited above suggests that there is a certain directionality of mapping processes that, however, does not follow the sequential continuous nature of the speech stream. Furthermore, the directionality of lexical processing seems language-universal. Unfortunately, given the limited nature of the information presented in the literature, the exact nature of the processing directionality needs to be further explored in future research.

5.2.2 Sub-lexical influences on whole word processing

The present research provides compelling evidence that sub-lexical phonology (e.g., /wΛn/ in the word "**won**der" or the segmental information /m/ in the Chinese character /ma3/ 马) exerts critical effects on spoken word processing. In this view, the mechanism underlying spoken word processing involves a less binding commitment instead of an all-or-none decision. In particular, the present research challenges the inhibition position held by the TRACE model. Instead, it supports the facilitation claim advocated by the Cohort model that sub-lexical phonology including syllables in multi-syllabic words and segments (onsets and rimes) in monosyllabic characters produces a facilitation effect on word recognition.

As stated in Chapter Two, compelling evidence for this facilitation view has been obtained in a number of behavioral studies. For example, using a cross-modal priming paradigm, Shillcock (1990) asked subjects to make lexical decisions on the visually presented target item that was preceded by aurally presented primes. The target was either related to the homophone of one syllable of the spoken prime (e.g., "rib" for "trombone") or was unrelated (e.g., "bee"). A clear facilitation instead of inhibition effect was found: lexical decisions to the targets such as "rib" were much faster and more accurate than those to the targets such as "bee". Consistent with the Cohort model, it is reasoned that the most suitable candidate for the auditory prime does not deactivate or inhibit those alternatives that are semantically or phonologically associated with part of the acoustic input (syllables in this case). Instead, those alternatives are active for a period of time, which makes it possible that the sub-lexical syllable "bone" in "trombone" primes the target "rib".

According to the findings in the first experiment, it is further emphasized that the weighting of the facilitation roles of syllables in multi-syllabic English word processing varies according to their stress status: stressed syllables may give more salient contributions to semantic activation than unstressed syllables. In this account, the **modified Cohort model** that takes word prosody into account needs to be considered in speech research.

Some behavioral data provide converging, albeit implicit, evidence in favor of the modified Cohort model. For instance, in a cross-modal priming study (Swinney, 1981), the sub-lexical facilitation effect was more apparent in initial embedded words (e.g., "boy" in "boycott") but less discernible in final embedded words (e.g., "cot"). These results seem at odds with the sub-lexical phonological facilitation claim. However, close examination of the stimuli used in this study reveals that the syllables of interest in most final embedded words were unstressed. In this regard, the results do not in fact challenge but truly support the viability of the **modified Cohort model**. On the basis of the **modified Cohort model**, the weak facilitation effect on final embedded words can be appropriately accounted for with the stress status of priming syllables. As for the Shillcock's (1990) study mentioned above, the reliable facilitation effect observed can be interpreted in line with the modified Cohort model in that sub-lexical cues of primes that produced robust facilitation effects were all stressed syllables.

5.3 Implications for Reading Models of Word Recognition

The present findings in the third experiment of the thesis provide explicit converging evidence in favor of the "phonological mediation" theory (e.g., parallel distributed model proposed by Seidenberg & McClelland, 1989). According to this theory, phonology serves an important function in reading comprehension (Xu, 1991; Zhang & Perfetti, 1993). Of particular importance is that phonology plays a role in accessing the meaning of a word. Nonetheless, contrary to the "phonological recoding" theory, phonological recoding on the basis of the mediation theory is not the exclusive solution in access to semantic meaning. Rather, orthography and phonology interact with each other and both exert their influence on semantic activation of a word.

The present research stands in stark contrast with the "direct access" theory. To date, a widespread viewpoint of Chinese reading has been the direct access theory. According to this theory, readers of Chinese access semantic meaning of a character on the basis of the orthographic or visual representation of the character without the involvement of phonology (e.g., Barron, 1978; Smith, 1985; Shen & Forster, 1999; Wong & Chen, 1999). The logic behind this theory has relied mainly on analyses of the Chinese writing system. Chinese is an ideographic language in which each Chinese character has its own symbol. The characteristics of Chinese morpho-syllabic writing system make viable the assumption that the Chinese graphemes map onto meaningful morphemes instead of phonemes in the spoken language as alphabetic languages do. In line with this assumption, advocates of the direct access theory claim that readers of ideographic languages comprehend written characters by means of a visual-orthographic pathway; they do not necessarily use the strategy of decoding orthography to phonology in order to constrain semantic meaning.

Although the direct access explanation sounds convincing on intuitive grounds, the present research nonetheless does not support it. The major fallacy of the direct access claim is that it only takes into consideration analyses of writing systems rather than of human language and cognitive processing. It is argued that the form of a writing system such as a sign script of Chinese may influence the means and processes through which orthographic representations are transformed to phonological codes. However, a writing system does not necessarily function in restraining the occurrence of phonological activation in an all-or-none manner (Perfetti & Zhang, 1995). That is, if phonological recoding reflects aspects of human language reading, it may be optimally efficient for the reader to transform orthographic representations into phonological codes regardless of what type of orthographic symbols he or she is reading (e.g., Perfetti & Zhang, 1995; Tzeng, Hung, & Wang, 1977). Undoubtedly, the present research strongly supports this view.

According to the findings in the third experiment of the thesis, the sensitivity of the N270 to the distinct segmental and tonal information, along with the reliable phonological priming effects on the N400, reinforces the argument that phonological processing occurs as a component of Chinese language reading. In support of the arguments for the involvement of phonology in Chinese reading, there is considerable behavioral evidence that phonology helps constrain semantic meaning of words and ample evidence that phonology is recoded automatically (e.g., Chua, 1999; Perfetti & Zhang, 1995; Tan & Perfetti, 1997; Xu, Pollatsek, & Potter, 1999).

In particular, the phonological priming effects on the N270 and the N400 have been reliably observed in a good number of English reading studies (e.g., Forbes, 1999; Newman, 2000). Although it remains an open question concerning the extent of the confounding of orthography with phonological function in English reading, the reliable phonological effects observed in both English and Chinese ERP studies nonetheless lead to a reasonable argument that phonological recoding may reflect aspects of language-universal mechanisms underlying reading comprehension.

The present research further demonstrates that lexical tonal information, along with segmental cues, plays an important role in access to semantic meaning in Chinese reading. This gives rise to the argument that word prosody is part of phonological codes transformed from print. Compatible with the present research, several behavioral studies have also provided converging evidence in favor of this argument (e.g., Spinks, Liu, Perfetti, et al., 2000; Xu, Pollatsek, & Potter, 1999). For example, in an attempt to explore the function of phonology in semantic activation of Chinese characters, Spinks and colleagues (2000) carried out a Stroop study in which subjects named the ink color of viewed Chinese characters or color patches. In this study, color characters and their homophones without orthographic similarity to color characters were used as target stimuli. The homophone stimuli included homophones with the same tone and homophones with different tones. It was hypothesized that if meaning activation is constrained by both orthographic and phonological representations, the homophones of color characters should produce interference with color naming although the magnitude of this interference would be smaller than that induced by color characters. The findings of this study confirmed their hypothesis and therefore corroborate the argument that phonology is indeed activated and involved in the access of semantic meaning of a character. The most notable finding in this study is that the interference effect was more robust for homophones with the same tone than for those with a different tone. Based on this finding, it seems likely that a Chinese character's phonological code activated in reading includes both segmental information (onset and rime) and tonal information (word prosody).

Taking all the above into account, it is reasonable to argue that phonology is involved in the access of semantic meaning in reading. More relevantly, phonological representations are not recoded from print in an abstract impoverished manner. Instead, they contain, in effect, a complete phonological specification of a Chinese character, which includes not only segmental but also word prosodic information.

5.4 Directions for Future Research

To date, there has been growing evidence for the different presentation forms of the function of word prosodic information in lexical processing between English and Chinese. In the future, it will be important to address the issue of whether the effects of word prosody in word recognition are brought into play simply by the presence of a particular language prosodic structure itself in the speech input. A good number of speech perception studies have suggested that the utility of word prosodic information differs as a function of the listener's language experience. For instance, native speakers of tone-based languages (e.g., Chinese) have difficulty in discriminating stress contrasts in English language (e.g., Dupoux, Pallier, Sebastian, et al., 1997). In turn, speakers of stress-based languages (e.g., English) have trouble in determining Cantonese and Mandarin tones. It is worth mentioning that speakers of these two tone-based languages were found to be better at discriminating tone contrasts of their mother language than of the other tone language (Lee, Vakoch, & Wurm, 1996).

Taking these performance data into consideration, it is hypothesized that lexical stress during spoken English word recognition may not exert its optimal effect on semantic constraints if subjects are Chinese speakers who have English as their second language, in contrast with native English speakers. In an attempt to determine the viability of this proposal, a future ERP experiment could be carried out, in which the same stimuli and experimental designs as the first experiment would be used and native Chinese speakers who can speak English fluently would be tested.

Summary

Event related brain potentials were used in the three experiments of the thesis to examine the roles of word prosody and sub-lexical phonology in speech and reading. The findings obtained in the present research suggest that word prosodic information, along with segmental phonology, plays an important role in both Chinese and English spoken word processing. The involvement of word prosody in both English and Chinese speech implies a language universal mechanism underlying prosodic phonological processing. The findings also suggest that lexical processing does not proceed in an all-or-none fashion. Rather, sub-lexical information plays a facilitation role in whole word processing. The results of the third experiment provide strong support for the "phonological mediation" theory. It is concluded that word prosody is recoded as part of complete phonological specification of a Chinese character during silent reading.

APPENDIX A

Sentence Stimuli used in each of the five experimental conditions in Experiment 1.

Stimuli in the CONGRUENT Condition

1. The daisy is a very pretty **flow**er.
2. The maple leaf is a well-known national **sym**bol.
3. The townspeople voted to elect a **may**or.
4. Julie took her car over to PEI on the **fe**rry.
5. Salt and pepper are **con**diments.
6. At the end of a drum roll the drummer usually hits a **cym**bal.
7. The apartment was described in the newspaper's rental **sec**tion.
8. The bank robber escaped with all the **mo**ney.
9. Every night Jenny's mother told her a bedtime **sto**ry.
10. The police officer gave jack a speeding **ti**cket.
11. Tommy brought a present to the birthday **par**ty.
12. Tony wants to go to the rock **con**cert.
13. The author's novel was made into a **mo**vie.
14. The comedian's joke was really **fu**nny.
15. Jason lost his driver's **li**cense.
16. Michael borrowed a novel from the public **li**brary.
17. She tied her hair up in a blue **ri**bbon.
18. Three people were killed in a major highway **ac**cident.
19. Plants will not grow in dry **wea**ther.
20. Sharon dried the bowls with a **tow**el.
21. Paddy cut the fabric with a pair of **sci**ssors.
22. Mike pointed to the picture with his index **fin**ger.

23. The lost motorist pulled over to ask for **di**rections.

24. The lawyer feared his client was **guil**ty.

25. New York is a very busy **ci**ty.

26. The better students thought the test was too **ea**sy.

27. Jill looked back through the open **win**dow.

28. The surgeon tried in vain to save his **pa**tient.

29. No one wants to hire a man who just got out of **pri**son.

30. The side of the road was covered with bottles and other **li**tter.

31. The almond orchard had blossomed but a disease made the flowers **wi**ther.

32. They use X-ray machines to search everyone's luggage for **wea**pons.

33. The dog is a **bea**gle.

34. The crocodile is a strange-looking **crea**ture.

35. Every morning she cooks omelets for **break**fast.

36. Until he died did the dog remain faithful to his **mas**ter.

37. That actor attended the last Cannes Film **Fes**tival.

38. The cat family includes lions and **ti**gers.

39. The baker was covered in white **flour**.

40. Billy put his tooth under his pillow for the tooth **fai**ry.

Stimuli in the LIGHT/PO**LITE** (Stress noninitial – Prime noninitial) condition

1. That large brown animal is a grizzly for**bear** (bear).

2. Karen went on the ferris wheel at the country a**ffair** (fair).

3. When he gets to the end of the walk he must open the ne**gate** (gate).

4. Eight minus six equals ta**ttoo** (two).

5. Bill's trust was something Jane had to work hard to a**gain** (gain).

6. She told the lost tourist to turn right at the traffic po**lite** (light).

7. In polka game, every one has a chance to for**bid** (bid).

8. It is easier to view the stars at ig**nite** (night).

9. Tomorrow the ship will set as**sail** (sail).

10. The dog wagged its de**tail** (tail).

11. The shoes were too small for Mike's de**feat** (feet).

12. The baker baked a dozen loaves of in**bred** (bread).

13. The spaghetti was served with an Italian tomato re**source** (sauce).

14. After surgery the blind man regained his ex**cite** (sight).

15. Colleen bought an outfit to a**ware** (wear).

16. Cindy took out enough money for bus un**fair** (fare).

17. The king's death ended his long te**rrain** (reign).

18. He put the toast on the breakfast be**tray** (tray).

19. The laundry soap did not remove the coffee ab**stain** (stain).

20. Donna built a sandcastle with her shovel and im**pale** (pail).

21. The prisoner had a sink in his ex**cel** (cell).

22. The number of legs a spider has is cre**ate** (eight).

23. The campers listened to their leader tell an old wives' en**tail** (tale).

24. Table salt is fine grained but rock salt is re**course** (coarse).

25. Before the director was satisfied, we had to do a dozen of mis**takes** (takes).

26. Lori remembered she had to meet her friend but she forgot when and be**ware** (where).

27. Frank buttered his dinner en**rol** (roll).

28. The dog whistle is too high for humans to ad**here** (hear).

29. Many people refuse to wear animal pre**fer** (fur).

30. This type of tree is called a Douglas in**fer** (fir).

31. Bob bought an apple, an orange and a com**pare** (pear).

32. Mary raked the autumn leaves into a large com**pile** (pile).

33. Janice knocked on her neighbor's front a**dore** (door).

34. The losing gambler insisted placing another a**bet** (bet).

35. The bird on the birdfeeder ate every last ex**ceed** (seed).

36. Shuffle the cards before you or**deal** (deal).

37. The number of students was 120 at the last dis**count** (count).

38. At last the time for action had be**come** (come).

39. She took the baby out of the theatre when he started to de**cry** (cry).

40. A single sheet of glass in a window frame is commonly termed cam**paign** (pane).

Stimuli in the ONE/**WON**DER (Stress initial – Prime initial) condition

1. The boxing match started at the sound of the **Bel**ted (bell).
2. The worker climbed up the telephone **poul**try (pole).
3. Eight minus seven equals **won**der (one).
4. To keep the boat from sinking the man had to **bale**ful (bail).
5. To hide the treasure the pirates dug a deep **hol**der (hole).
6. The patient was given medicine to ease her **pain**ting (pain).
7. The students learned how to multiply, divide, subtract and **ad**vertise (add).
8. Eight thousand people were surveyed in a Gallup **pol**ka (poll).
9. Make sure everything is ready when the committee members **com**fort (come).
10. Ned has a daughter and a **Sun**day (son).
11. Nov 11 is Canadian Remembrance **da**ta (day).
12. The class baked a cake and everyone had a **peace**ful (piece).
13. The sum of all the parts equals a **hol**ster (whole).
14. When the sun set, the prisoner went back to his **sel**dom (cell).
15. When Gerald bent over his pants ripped along the **seem**ly (seam).
16. Sam adjusted the sound of the stereo using the treble and the **base**ment (bass).
17. Such a shameless request deserves the answer **no**ble (no).
18. The church steeple has a weather **vain**ly (vane).
19. Belts are worn around your **waste**ful (waist).
20. There are seven days in a **weak**ness (week).
21. Water and sunshine help plants **gro**cery (grow).
22. After finishing their homework the children went out to **Pla**to (play).

23. Before a law is passed, it is called a **buil**ding (bill).

24. Ryan writes with his left **hand**some (hand).

25. The children climbed up the apple **trea**ty (tree).

26. An honest person could never tell a **li**on (lie).

27. Little Joe took off one shoe but kept wearing the other **shoo**ting (shoe).

28. The author dedicated the book to those that had encouraged him to **right**ful (write).

29. She was too impatient to stand around and **weight**less (wait).

30. The antique bed looks beautiful but very uncomfortable to sleep **in**cident (in).

31. After the talk everyone had cheese and crackers with **win**ding (wine).

32. He hung the clothes on the line to air **dri**ver (dry).

33. When the moon is full it is hard to see many stars or the milky **weigh**ted (way).

34. We became impatient because the service was very **slo**gan (slow).

35. April eleventh is when Tina's math assignment is **du**ty (due).

36. Some people like to drink coffee and some like to drink **te**dious (tea).

37. Their favorite dessert is Apple **pi**rate (pie).

38. With the microscope you can see a special red blood **sel**fish (cell).

39. Nancy likes the apple pie but I prefer the cherry **pi**lot (pie).

40. He tried to put the pieces of the broken plate back together with **gloo**my (glue).

Stimuli in the EYE/**IDEA** (Stress noninitial – Prime initial) condition

1. A female deer can be called a do**mes**tic (doe).
2. A Volvo is a popular car**toon** (car).
3. The pirate wore a patch over his **iden**tical (eye).
4. The fact that the Canadian flag is red and white is something that most people no**bel** (know).
5. Under the sun, the diamond shone with every hu**mane** (hue).
6. Is this the book you are looking for**ma**tion (for).
7. Your mother told me to ask **uni**ted (you).
8. The pupil is the black opening in your **iro**nic (eye).
9. Kevin's car was brand neu**tra**lity (new).
10. In class the students sat in the first ro**bust** (row).
11. Fred stubbed his big to**na**lity (toe).
12. George wants to go to the movie and so do **Idea** (I).
13. The congregation sang a him**self** (hymn).
14. The answer to the question was either yes or no**ma**dic (no).
15. The Bible says stealing food is a sin**cere** (sin).
16. Your life ends when you dy**na**mic (die).
17. A female sheep is called a u**ni**que (ewe).
18. When Jim's car went into the ditch he called a garage to get a to**ma**to (tow).
19. The grass was wet from the morning du**et** (dew).
20. Joan handed her essay in when it was du**pli**city (due).
21. She unlocked the door with a ki**ne**tic (key).

22. The sad film made the woman cri**te**rion (cry).

23. The squirrel stored some nuts in the tre**men**dous (tree).

24. He liked lemon and sugar in his tee**to**taller (tea).

25. Eight plus two equals ten**den**tious (ten).

26. On the last Thanksgiving Day Grandma baked a pumpkin pio**neer** (pie).

27. Her parents call her Susan but her friends just call her su**pine** (sue).

28. She did not believe this story and replied sarcastically: "Is that so**bri**ety" (so).

29. His parents were killed in a plane rec**tan**gular (wreck).

30. I had to wear Velcro sneakers until I learned how to ti**ta**nic (tie).

31. The basketball star had to sit out the season because of his sprained mne**mo**nic (knee).

32. A stream of radiation from outer space is called a cosmic re**gime** (ray).

33. He did not worry about burglars because he kept that fierce dog**ma**tic (dog).

34. The antique chair was beautiful but very uncomfortable to sit in**a**dequate (in).

35. My car has broken down; will you give me a to**ta**lity (tow).

36. A hat can also be called a cap**ti**vity (cap).

37. The indefinite article modifying the word 'elephant' is an**ces**tral (an).

38. The room was hot so he turned on the electric fan**tas**tic (fan).

39. Joe was not accustomed to refusing people by saying No**vem**ber (no).

40. This chemical that gives the wall a shining blue is a special type of di**ver**sity (dye).

Stimuli in the INCONGRUENT condition

1. The manager phoned the applicant that he decided to **plas**ter (hire).
2. The waiter recommended a sparkling white **di**fference (wine).
3. Susan joined a convent and became a **pla**net (nun).
4. The article discussed the inequality between the rich and the **lan**guage (poor).
5. Janice went outside to get some fresh **em**bassy (air).
6. The highest part of the mountain is the **ra**bbit (peak).
7. The soldier's wounds took a long time to **mon**key (heal).
8. Oak and mahogany are both expensive types of **dia**logue (wood).
9. The dove is a sign of **bi**cycle (peace).
10. The yellow part of an egg is called the **e**ffort (yoke).
11. When the princess kissed the frog it turned into a **ro**bbery (prince).
12. Fortunetellers talk about the present, future and **sli**ppers (past).
13. The teacher told her class a story about a tortoise and a **di**nner (hare).
14. The piano is out of **di**ary (tune).
15. Eight minus seven equals **A**frica (one).
16. The teacher wrote the problem on the **fo**rest (blackboard).
17. The child was born with a rare **hea**ter (disease).
18. The movie was so packed that they couldn't find a single **beau**tiful (seat).
19. Fred realized the old house was up for **el**bows (sale).
20. If the crowd quiets down the band will **di**amond (start).
21. Steve is a student who hangs on the professor's every **ma**rathon (word).
22. The company has gradually bought out its smaller **ye**llow (rivals).

23. The parents pleaded with their daughter to come **va**lley (home).
24. We sometimes forget that golf is just a **pen**cil (game).
25. His leaving home amazed all his **pos**ters (family).
26. My aunt liked to read the daily **moun**tain (newspaper).
27. Doug is strong but Joe is really **swim**ming (weak).
28. After dinner they washed the **au**tumn (dishes).
29. The campers cooked dinner over an open **shif**ty (fire).
30. By way of apology he sent her a dozen long stemmed **lob**sters (roses).
31. There were advantages to living in a city but Marsha moved to a small town for the peace and **wed**ding (quietness).
32. The old milk tasted very **ma**chine (sour).
33. He hadn't paid any of the bills so the power company cut off the **pi**llow (electricity).
34. A person who inherits a family fortune is called an **ab**sent (heir).
35. A wild pig is also called a **li**ar (boar).
36. The puppy was so expensive because it was pure **neigh**bor (bred).
37. Yoghurt is sold either with fruit or it is just **hea**ven (plain).
38. The power went out last night and all the food went **dic**tionary (bad).
39. The stagnant green pond was a breeding ground for **li**steners (mosquitoes).
40. The spider sat in its web awaiting a **clo**set (fly).

APPENDIX B

Table B1

ANOVA for PMN Amplitude Data for Condition X Time X Site Analysis for Experiment 1

Source	df	MSE	F
Condition	4, 56	96.65	30.57*
Time	2, 28	10.69	0.63
Site	16, 224	10.66	11.04*
Condition X Time	8,112	11.29	7.02*
Condition X Site	64,896	6.16	4.19*
Time X Site	32,448	0.57	2.89*
Condition X Time X Site	128,1792	0.67	2.60*

*$p < .05$

Table B2

Mean Amplitudes (in μV) (and Standard Errors) for the Condition X Time X Region analysis for the PMN component in Experiment 1.

Mean amplitudes in the 200-250 ms time interval

	Frontal	Central	Parietal	Temporal	Occipital
Congruent	0.50(0.53)	1.81(0.65)	2.71(0.67)	0.73(0.38)	1.87(0.67)
Count-Discount	-0.80(0.58)	-0.27(0.54)	1.28(0.54)	0.20(0.30)	1.39(0.51)
One-Wonder	-0.26(0.45)	1.30(0.69)	2.32(0.82)	0.68 (0.48)	1.74(0.78)
Car-Cartoon	-1.65(0.67)	-1.56 (0.58)	-0.70(0.52)	-0.46 (0.27)	-0.00 (0.56)
Incongruent	-2.16 (0.60)	-2.95 (0.59)	-2.63 (0.70)	-1.35 (0.37)	-1.93 (0.61)

Mean amplitudes in the 250-300ms time interval

	Frontal	Central	Parietal	Temporal	Occipital
Congruent	0.58(0.47)	2.09(0.63)	3.25(0.63)	1.14(0.35)	2.50(0.59)
Count-Discount	0.06(0.56)	0.79(0.49)	2.03(0.57)	0.40(0.31)	1.73(0.61)
One-Wonder	0.43(0.43)	1.92(0.57)	2.74(0.76)	0.94(0.44)	1.91(0.80)
Car-Cartoon	-2.18(0.59)	-2.31(0.53)	-1.22(0.44)	-0.98(0.28)	-0.65(0.52)
Incongruent	-2.77(0.60)	-3.41(0.56)	-2.93(0.58)	-1.71(0.34)	-1.85(0.50)

Table B2 con't

Mean amplitudes in the 300-350ms time interval

	Frontal	Central	Parietal	Temporal	Occipital
Congruent	0.21(0.49)	2.45(0.67)	4.17(0.64)	1.31(0.33)	3.60(0.63)
Count-Discount	0.79(0.53)	1.70(0.56)	2.85(0.59)	0.64(0.36)	2.32(0.66)
One-Wonder	0.26(0.52)	1.45(0.70)	2.06(0.71)	0.70(0.44)	1.26(0.63)
Car-Cartoon	-3.11(0.57)	-3.07(0.58)	-1.54(0.60)	-1.42(0.30)	-0.50(0.66)
Incongruent	-3.42(0.52)	-4.17(0.55)	-3.52(0.66)	-2.30(0.37)	-2.49(0.63)

Table B3

ANOVA for PMN Amplitude Data for Condition X Time X Region X Hemisphere Analysis for Experiment 1.

Source	df	MSE	F
Condition	4,56	58.62	29.56*
Time	2,28	6.80	0.37
Region	4,56	12.88	17.03*
Hemisphere	1,14	25.55	1.54
Condition X Time	8,112	6.68	6.55*
Condition X Region	16,224	8.02	4.86*
Time X Region	8,112	0.66	2.60
Condition X Time X Region	32,448	0.84	3.03*
Condition X Hemisphere	4,56	8.33	0.51
Time X Hemisphere	2,28	1.47	3.13
Condition X Time X Hemisphere	8,112	1.38	1.81
Region X Hemisphere	4,56	1.78	2.14
Condition X Region X Hemisphere	16,224	1.40	1.79
Time X Region X Hemisphere	8,112	0.14	2.23
Condition X Time X Region X Hemisphere	32,448	0.17	1.62

*$p < .05$

Table B4

ANOVA for N400 Amplitude Data for Condition X Time X Site Analysis for Experiment 1.

Source	df	MSE	F
Condition	4, 56	226.85	27.30*
Time	5, 70	87.18	34.44*
Site	16, 224	53.45	31.00*
Condition X Time	20, 280	22.98	6.28*
Condition X Site	64,896	15.93	3.93*
Time X Site	80, 1120	2.18	22.93*
Condition X Time X Site	320, 4480	0.98	2.63*

*p < .05

Table B5

Mean amplitudes (in μV) (and Standard Deviations) for the PMN and N400 components in Experiment 1

Stimulus Conditions	PMN	N400
Congruent	1.93(0.44)	3.77(0.50)
Count- Discount	1.01(0.37)	3.10(0.55)
One- Wonder	1.30(0.54)	0.74(0.62)
Car- Cartoon	-1.42(0.38)	0.03(0.44)
Incongruent	-2.64(0.44)	-0.93(0.51)

Table B4

ANOVA for N400 Amplitude Data for Condition X Time X Region X Hemisphere Analysis for Experiment 1.

Source	df	MSE	F
Condition	4, 56	136.22	27.35*
Time	5, 70	55.79	36.65*
Region	4, 56	64.41	45.86*
Hemisphere	1, 14	58.23	5.03*
Condition X Time	20, 280	13.61	6.79*
Condition X Region	16,224	20.72	4.87*
Time X Region	20, 280	2.99	29.22*
Condition X Time X Region	80, 1120	1.13	3.50*
Condition X Hemisphere	4, 56	20.39	0.78
Time X Hemisphere	5, 70	2.26	0.68
Condition X Time X Hemisphere	20, 280	2.39	1.31
Region X Hemisphere	4, 56	10.30	3.34*
Condition X Region X Hemisphere	16, 224	4.09	2.64*
Time X Region X Hemisphere	20, 280	0.41	3.09*
Condition X Time X Region X Hemisphere	80, 1120	0.19	2.04*

*$p < .05$

Table B5

Mean N400 amplitudes (in µV) (and Standard Deviations) for main effect for Hemisphere factor for Experiment 1.

Hemisphere	
Left hemisphere	1.37(0.42)
Right hemisphere	1.88(0.41)

APPENDIX C

Chinese four-character proverb stimuli used in each of the five experimental conditions in Experiments 2 & 3.

Stimuli in the **Congruent** Condition

1. 爱憎分明 /ming3/
2. 八仙过海 /hai3/
3. 白日作梦 /meng4/
4. 包罗万象 /xiang4/
5. 笨鸟先飞 /fei1/
6. 闭关自守 /shou3/
7. 表里如一 /yi1/
8. 杞人忧天 /tian1/
9. 不加思索 /suo3/
10. 海枯石烂 /lan4/
11. 后患无穷 /qiong2/
12. 平淡无奇 /qi2/
13. 焕然一新 /xin1/
14. 家喻户晓 /xiao3/
15. 见缝插针 /zhen1/
16. 金碧辉煌 /huang2/
17. 举世闻名 /ming2/
18. 量体裁衣 /yi1/
19. 宽宏大量 /liang4/
20. 老谋深算 /suan4/

21. 立竿见影 /ying3/

22. 良药苦口 /kou3/

23. 伶牙俐齿 /chi3/

24. 藏龙卧虎 /hu3/

25. 奇耻大辱 /ru3/

26. 承上启下 /xia4/

27. 持之以恒 /heng2/

28. 唇亡齿寒 /han2/

29. 粗制滥造 /zao4/

30. 大惊失色 /se4/

31. 胆战心惊 /jing1/

32. 地大物博 /bo2/

33. 粉身碎骨 /gu3/

34. 风土人情 /qing2/

35. 奉公守法 /fa3/

36. 梦寐以求 /qiu2/

37. 肝胆相照 /zhao4/

38. 归心似箭 /jian4/

39. 逢凶化吉 /ji2/

40. 怒发冲冠 /guan4/

Stimuli in the **Onset** Condition

1. 唉声叹立(**气**) /li4/←/**qi4**/
2. 安分守起(**己**) /qi2/←/**ji2**/
3. 按部就帆(**班**) /fan1/←/**ban1**/
4. 饱经风荒(**霜**) /huang1/←/**shuang1**/
5. 逼上梁班(**山**) /ban1/← /**shan1**/
6. 闭门造科(**车**) /ke1/←/**che1**/
7. 别具一河(**格**) /he2/←/**ge2**/
8. 波澜壮落(**阔**) /luo4/←/**kuo4**/
9. 不学无肚(**术**) /du4/←/**shu4**/
10. 和风细女(**雨**) /nu3/←/**yu3**/
11. 饥寒交墨(**迫**) /mo4/←/**po4**/
12. 物美价前(**廉**) /qian2/←/**lian2**/
13. 见义勇陪(**为**) /pei2/←/**wei2**/
14. 尽善尽给(**美**) /gei3/←/**mei3**/
15. 刻舟求变(**剑**) /bian4/←/**jian4**/
16. 狂妄自差(**大**) /cha4/←/**da4**/
17. 乐极生飞(**悲**) /fei1/← /**bei1**/
18. 炉火纯听(**青**) /ting1/←/**qing1**/
19. 才华横碧(**溢**) /bi4/←/**yi4**/
20. 沉默寡拦(**言**) /lan2/←/**yan2**/

21. 敲诈勒朵**(索)** /duo3/←**/suo3/**

22. 叱咤风浑**(云)** /hun2/←**/yun2/**

23. 出神入挂**(画)** /gua4/←**/hua4/**

24. 触目惊音**(心)** /yin1/←**/xin1/**

25. 绰绰有驴**(余)** /lu2/←**/yu2/**

26. 寸步难鸣**(行)** /ming2/←**/xing2/**

27. 大公无妻**(私)** /qi1/←**/si1/**

28. 大显神钟**(通)** /zhong1/←**/tong1/**

29. 胆小如补**(鼠)** /bu3/←**/shu3/**

30. 得寸进笔**(尺)** /bi3/←**/chi3/**

31. 颠倒黑抬**(白)** /tai2/←**/bai2/**

32. 东施效勤**(颦)** /qin2/← **/pin2/**

33. 短小精蛋**(悍)** /dan4/← **/han4/**

34. 发号施硬**(令)** /ying4/← **/ling4/**

35. 飞黄腾茶**(达)** /cha2/← **/da2/**

36. 赴汤蹈果**(火)** /guo3/← **/huo3/**

37. 甘拜下撑**(风)** /cheng1/← **/feng1/**

38. 孤注一赤**(掷)** /chi4/← **/zhi4/**

39. 寡不敌洞**(众)** /dong4/← **/zhong4/**

40. 诡计多砖**(端)** /zhuan1/← **/duan1/**

Stimuli in the **Rime** Condition

1. 爱不释少(**手**) /shao3/←**/shou3/**
2. 暗箭伤荣(**人**) /rong2/←**/ren2/**
3. 跋山涉闪(**水**) /shan3/←**/shui3/**
4. 百花齐愤(**放**) /fen4/←**/fang4/**
5. 班门弄访(**斧**) /fang3/← **/fu3/**
6. 报仇雪唤(**恨**) /huan4/←**/hen4/**
7. 本末倒震(**置**) /zhen4/←**/zhi4/**
8. 变化无迟(**常**) /chi2/←**/chang2/**
9. 彬彬有冷(**礼**) /leng3/←**/li3/**
10. 不谋而寒(**合**) /han2/←**/he2/**
11. 豪言壮影(**语**) /ying3/←**/yu3/**
12. 哗众取喘(**宠**) /chuan3/←**/chong3/**
13. 回光返正(**照**) /zheng4/←**/zhao4/**
14. 豁然开里(**朗**) /li3/←**/lang3/**
15. 惊慌失菜(**措**) /cai4/←**/cuo4/**
16. 开门见湿(**山**) /shi1/←**/shan1/**
17. 来龙去梦(**脉**) /meng4/← **/mai4/**
18. 老当益站(**壮**) /zhan4/←**/zhuang4/**
19. 了如指枕(**掌**) /zhen3/←**/zhang3/**
20. 绿林好画(**汉**) /hua4/←**/han4/**

21. 草木皆奔(**兵**) /ben1/←/**bing1**/

22. 长篇大泪(**论**) /lei4/←/**lun4**/

23. 趁热打塔(**铁**) /ta3/←/**tie3**/

24. 乘虚而绕(**入**) /rao4/←/**ru4**/

25. 重蹈覆宅(**辙**) /zhai2/←/**zhe2**/

26. 出生入伞(**死**) /san3/←/**si3**/

27. 穿针引穴(**线**) /xue4/←/**xian4**/

28. 错综复族(**杂**) /zu2/←/**za2**/

29. 大刀阔粉(**斧**) /fen3/←/**fu3**/

30. 大器晚查(**成**) /cha2/←/**cheng2**/

31. 大张旗阁(**鼓**) /ge3/←/**gu3**/

32. 得意忘旋(**形**) /xuan2/← /**xing2**/

33. 独出心从(**裁**) /cong2/← /**cai2**/

34. 对牛弹权(**琴**) /quan2/← /**qin2**/

35. 发扬光凳(**大**) /deng4/← /**da4**/

36. 奋不顾杀(**身**) /sha1/← /**shen1**/

37. 风流人胃(**物**) /wei4/← /**wu4**/

38. 富丽堂横(**皇**) /heng2/← /**huang2**/

39. 高谈阔乐(**论**) /le4/← /**lun4**/

40. 孤芳自省(**赏**) /sheng3/← /**shang3**/

Stimuli in the **Tone** Condition

1. 爱财如明(**命**) /ming2/←/**ming4**/
2. 安家落壶(**户**) /hu2/←/**hu4**/
3. 白头偕牢(**老**) /lao2/←/**lao3**/
4. 川流不喜(**息**) /xi3/←/**xi1**/
5. 暴跳如累(**雷**) /lei4/← /**lei2**/
6. 不攻自婆(**破**) /po2/←/**po4**/
7. 不毛之滴(**地**) /di1/←/**di4**/
8. 鞭长莫鸡(**及**) /ji1/←/**ji2**/
9. 调虎离扇(**山**) /shan4/←/**shan1**/
10. 独断专醒(**行**) /xing3/←/**xing2**/
11. 多愁善干(**感**) /gan1/←/**gan4**/
12. 恩将仇包(**报**) /bao1/←/**bao4**/
13. 翻箱倒龟(**柜**) /gui1/←/**gui4**/
14. 繁荣昌绳(**盛**) /sheng2/←/**sheng4**/
15. 风平浪惊(**静**) /jing1/←/**jing4**/
16. 风云变欢(**幻**) /huan1/←/**huan4**/
17. 画蛇添阻(**足**) /zu3/← /**zu2**/
18. 固执己肩(**见**) /jian1/←/**jian4**/
19. 光明磊锣(**落**) /luo2/←/**luo4**/
20. 好事多墨(**磨**) /mo4/←/**mo2**/

21. 和睦相初(**处**) /chu1/←/**chu4**/

22. 高枕无友(**忧**) /you3/←/**you1**/

23. 皆大欢溪(**喜**) /xi1/←/**xi3**/

24. 精益求净(**精**) /jing4/←/**jing1**/

25. 扣人心县(**弦**) /xian4/←/**xian2**/

26. 来日方唱(**长**) /chang4/←/**chang2**/

27. 冷若冰爽(**霜**) /shuang3/←/**shuang1**/

28. 废寝忘师(**食**) /shi1/←/**shi2**/

29. 临危不菊(**惧**) /ju2/←/**ju4**/

30. 临阵磨墙(**枪**) /qiang2/←/**qiang1**/

31. 成家立爷(**业**) /ye3/←/**ye4**/

32. 重温旧蒙(**梦**) /meng2/← /**meng4**/

33. 出口成丈(**章**) /zhang4/← /**zhang1**/

34. 吹毛求刺(**疵**) /ci4/← /**ci1**/

35. 打草惊射(**蛇**) /she4/← /**she2**/

36. 大海捞枕(**针**) /zhen3/← /**zhen1**/

37. 刀光剑鹰(**影**) /ying1/← /**ying3**/

38. 不堪设香(**想**) /xiang1/← /**xiang3**/

39. 暴风骤遇(**雨**) /yu4/← /**yu3**/

40. 耐人寻围(**味**) /wei2/← /**wei4**/

Stimuli in the **Incongruent** Condition

1. 稳如泰闷(**山**) /men4/←/**shan1**/
2. 昂首阔草(**步**) /cao3/←/**bu4**/
3. 百家争碗(**鸣**) /wan3/←/**ming2**/
4. 暗箭伤浪(**人**) /lang4/←/**ren2**/
5. 卑躬屈蜡(**膝**) /la4/← /**xi1**/
6. 背信弃坡(**义**) /po1/←/**yi4**/
7. 遍体鳞贺(**伤**) /he4/←/**shang1**/
8. 兵慌马鞋(**乱**) /xie2/←/**luan4**/
9. 不拘小态(**节**) /tai4/←/**jie2**/
10. 不务正脚(**业**) /jiao3/←/**ye4**/
11. 好高鹜截(**远**) /jie2/←/**yuan3**/
12. 鹤立鸡麦(**群**) /mai4/←/**qun2**/
13. 欢天喜团(**地**) /tuan2/←/**di4**/
14. 祸不单花(**行**) /hua1/←/**xing2**/
15. 家破人树(**亡**) /shu4/←/**wang2**/
16. 艰苦奋脸(**斗**) /lian3/←/**dou4**/
17. 节外生卡(**枝**) /ka3/← /**zhi1**/
18. 惊天动蒙(**地**) /meng2/←/**di4**/
19. 雕虫小手(**技**) /shou3/←/**ji4**/
20. 苦尽甘卧(**来**) /wo4/←/**lai2**/

21. 滥竽充爬(**数**) /pa2/←/**shu4**/

22. 背井离踝(**乡**) /hui2/←/**xiang1**/

23. 良师益醉(**友**) /zui4/←/**you3**/

24. 落井下蛋(**石**) /dan4/←/**shi2**/

25. 插翅难汗(**飞**) /han4/←/**fei1**/

26. 陈词滥犬(**调**) /quan3/←/**diao4**/

27. 痴心妄全(**想**) /quan2/←/**xiang3**/

28. 愁眉苦社(**脸**) /she4/←/**lian3**/

29. 出类拔盘(**萃**) /pan2/←/**cui4**/

30. 初露锋嫁(**芒**) /jia4/←/**mang2**/

31. 垂头丧晚(**气**) /wan3/←/**qi4**/

32. 粗茶淡萌(**饭**) /meng2/← /**fan4**/

33. 打抱不穿(**平**) /chuan1/← /**ping2**/

34. 大失所头(**望**) /tou2/← /**wang4**/

35. 单枪匹叶(**马**) /ye4/← /**ma3**/

36. 倒背如氧(**流**) /yang3/← /**liu2**/

37. 低声下勇(**气**) /yong3/← /**qi4**/

38. 锋芒毕响(**露**) /xiang3/← /**lu4**/

39. 孤陋寡恐(**闻**) /kong3/← /**wen2**/

40. 海底捞险(**月**) /xian3/← /**yue4**/

APPENDIX D

Table D1

ANOVA for PMN Amplitude Data for Condition X Time X Site Analysis for Experiment 2

Source	df	MSE	F
COND	4, 56	104.26	18.38*
TIME	2, 28	18.78	0.46
SITE	16, 224	9.01	0.88
COND * TIME	8, 112	18.70	4.04*
COND * SITE	64, 896	5.30	3.32*
TIME * SITE	32, 448	0.75	4.01*
COND * TIME * SITE	128, 1792	0.55	1.92

*p < .05

Table D2

Mean Amplitudes (in μV) (and Standard Deviations) for the PMN component for the Condition x Site Analysis for Experiment 2.

	Stimulus Conditions		
	Congruent	Onset Change	Rime Change
Fz	1.05(0.60)	-2.29(0.62)	-1.49(0.55)
F3	1.04(0.53)	-2.08(0.43)	-1.23(0.51)
F4	1.04(0.53)	-2.56(0.46)	-1.47(0.53)
F7	0.49(0.49)	-1.42(0.35)	-1.17(0.56)
F8	0.46(0.42)	-1.28(0.49)	-1.08(0.48)
Cz	2.10(0.57)	-2.56(0.50)	-1.59(0.58)
C3	1.80(0.51)	-2.22(0.44)	-1.52(0.51)
C4	1.84(0.48)	-2.67(0.44)	-1.50(0.53)
Pz	2.86(0.50)	-2.44(0.40)	-1.54(0.54)
P3	2.57(0.48)	-1.64(0.36)	-1.37(0.51)
P4	2.37(0.42)	-2.32(0.41)	-1.37(0.57)
T3	0.94(0.42)	-0.31(0.52)	-0.90(0.36)
T4	1.25(0.59)	-1.16(0.57)	-1.57(0.56)
T5	1.09(0.36)	-1.00(0.26)	-0.76(0.35)
T6	1.29(0.31)	-1.36(0.30)	-1.21(0.52)
O1	2.01(0.49)	-1.53(0.34)	-1.51(0.45)
O2	2.03(0.45)	-1.74(0.29)	-1.29(0.55)

Table D2 con't

	Stimulus Conditions	
	Tonal Change	Incongruent
Fz	0.03(0.83)	-1.84(0.53)
F3	-0.25(0.70)	-1.18(0.57)
F4	-0.36(0.67)	-1.94(0.50)
F7	-0.27(0.52)	-1.19(0.84)
F8	-0.26(0.72)	-2.22(0.57)
Cz	0.66(0.94)	-2.88(0.61)
C3	-0.18(0.74)	-2.41(0.60)
C4	0.52(0.78)	-2.79(0.64)
Pz	0.98(0.89)	-3.70(0.92)
P3	0.41(0.66)	-2.65(0.76)
P4	0.79(0.75)	-3.47(1.16)
T3	-0.07(0.52)	-1.38(0.54)
T4	-0.26(0.76)	-2.14(0.39)
T5	0.19(0.38)	-1.89(0.71)
T6	0.47(0.44)	-1.92(0.65)
O1	0.56(0.66)	-2.68(0.87)
O2	0.75(0.67)	-2.79(0.93)

Table D3

Mean Amplitudes (in μV) (and Standard Deviations) for the PMN component for the Time x Site Analysis for Experiment 2.

	Time Interval		
	200-250ms	250-300ms	300-350ms
Fz	-0.69(0.39)	-1.08(0.40)	-0.94(0.41)
F3	-0.63(0.34)	-0.78(0.34)	-0.80(0.35)
F4	-0.91(0.31)	-1.22(0.33)	-1.05(0.36)
F7	-0.82(0.29)	-0.70(0.27)	-0.62(0.31)
F8	-0.86(0.35)	-0.98(0.44)	-0.80(0.46)
Cz	-0.37(0.45)	-0.95(0.41)	-1.24(0.44)
C3	-0.64(0.43)	-0.89(0.39)	-1.19(0.40)
C4	-0.62(0.38)	-1.12(0.39)	-1.02(0.43)
Pz	-0.49(0.52)	-0.71(0.46)	-1.10(0.49)
P3	-0.45(0.39)	-0.44(0.37)	-0.72(0.38)
P4	-0.66(0.52)	-0.89(0.52)	-0.87(0.56)
T3	-0.52(0.32)	-0.30(0.34)	-0.22(0.35)
T4	-0.85(0.39)	-0.98(0.48)	-0.49(0.50)
T5	-0.67(0.32)	-0.37(0.27)	-0.39(0.29)
T6	-0.64(0.31)	-0.66(0.32)	-0.34(0.37)
O1	-0.56(0.38)	-0.52(0.36)	-0.82(0.36)
O2	-0.48(0.42)	-0.67(0.41)	-0.68(0.45)

Table D3

ANOVA for PMN Amplitude Data for Condition X Time X Region X Hemisphere Analysis for Experiment 2

Source	df	MSE	F
COND	4, 56	58.62	29.56*
TIME	2, 28	6.80	0.37
REGION	4, 56	12.88	17.03*
HEMIS	1, 14	25.55	1.54
COND X TIME	8, 112	6.68	6.55*
COND X REGION	16, 224	8.02	4.86*
TIME X REGION	8, 112	0.66	2.61
COND X TIME X REGION	32, 448	0.84	3.03*
COND X HEMIS	4, 56	8.33	0.51
TIME X HEMIS	2, 28	1.47	3.13
COND X TIME X HEMIS	8, 112	1.38	1.81
REGION X HEMIS	4, 56	1.78	2.14
COND X REGION X HEMIS	16, 224	1.40	1.79
TIME X REGION X HEMIS	8, 112	0.14	2.23
COND X TIME X REGION X HEMIS	32, 448	0.17	1.62

*$p < .05$

Table D4

Mean Amplitudes (in µV) (and Standard Deviations) for the PMN component for the Condition x Time x Region Analysis for Experiment 2.

200-250ms Time Interval

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Frontal	0.50(0.53)	-0.80(0.58)	-0.26(0.45)	-1.65(0.67)	-2.16(0.60)
Central	1.81(0.65)	-0.27(0.54)	1.30(0.69)	-1.56(0.58)	-2.95(0.59)
Parietal	2.71(0.67)	1.28(0.54)	2.32(0.82)	-0.70(0.52)	-2.63(0.70)
Temporal	0.73(0.38)	0.20(0.30)	0.68(0.48)	-0.46(0.27)	-1.35(0.37)
Occipital	1.87(0.67)	1.39(0.51)	1.74(0.78)	0.00(0.56)	-1.93(0.61)

250- 300ms Time Interval

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Frontal	0.58(0.47)	0.06(0.56)	0.43(0.43)	-2.18(0.59)	-2.77(0.60)
Central	2.09(0.63)	0.79(0.49)	1.92(0.57)	-2.31(0.53)	-3.41(0.56)
Parietal	3.25(0.63)	2.03(0.57)	2.74(0.76)	-1.22(0.44)	-2.93(0.58)
Temporal	1.14(0.35)	0.40(0.31)	0.94(0.44)	-0.98(0.28)	-1.71(0.34)
Occipital	2.50(0.59)	1.73(0.61)	1.91(0.80)	-0.65(0.52)	-1.85(0.50)

Table D4 con't

300- 350ms Time Interval

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Frontal	0.21(0.49)	0.79(0.53)	0.26(0.52)	-3.11(0.57)	-3.42(0.52)
Central	2.45(0.67)	1.70(0.56)	1.45(0.70)	-3.07(0.58)	-4.17(0.55)
Parietal	4.17(0.64)	2.85(0.59)	2.06(0.71)	-1.54(0.60)	-3.52(0.66)
Temporal	1.31(0.33)	0.64(0.36)	0.70(0.44)	-1.42(0.30)	-2.30(0.37)
Occipital	3.60(0.63)	2.32(0.66)	1.26(0.63)	-0.50(0.66)	-2.49(0.63)

Table D5

ANOVA for PMN Peak Latency Data for Condition X Time X Site Analysis for Experiment 2

Source	df	MSE	F
COND	4, 56	7821.25	9.78*
SITE	16, 224	320.72	1.87
COND X SITE	64, 896	268.22	1.13

* $p < .05$

Table D6

ANOVA for N400 Amplitude Data for Condition X Time X Site Analysis for Experiment 2

Source	df	MSE	F
COND	4, 56	133.66	36.78*
TIME	3, 42	29.68	28.90*
SITE	16, 224	20.45	1.97
COND X TIME	12, 168	17.15	3.74*
COND X SITE	64, 896	7.16	4.88*
TIME X SITE	48, 672	0.92	12.36*
COND X TIME X SITE	192, 2688	0.53	2.75*

* $p < .05$

Table D7

Mean Amplitudes (in μV) (and Standard Deviations) for the N400 component for the Condition x Time x Region Analysis for Experiment 2.

350-400ms

	Frontal	Central	Parietal	Temporal	Occipital
Congruent	2.08(0.48)	3.14(0.51)	3.62(0.53)	2.05(0.41)	2.62(0.55)
Onset Change	-1.72(0.41)	-2.16(0.48)	-1.55(0.49)	-0.27(0.38)	-1.57(0.44)
Rime Change	-1.53(0.47)	-2.51(0.65)	-2.63(0.59)	-1.74(0.45)	-2.83(0.53)
Tonal Change	0.16(0.59)	0.16(0.81)	0.07(0.80)	-0.01(0.46)	-0.17(0.70)
Incongruent	-2.24(0.64)	-3.57(0.65)	-3.42(0.97)	-1.47(0.53)	-2.76(0.81)

400-450ms

	Frontal	Central	Parietal	Temporal	Occipital
Congruent	2.78(0.49)	3.96(0.55)	4.38(0.50)	2.72(0.41)	3.19(0.55)
Onset Change	-1.38(0.48)	-1.41(0.64)	-0.11(0.55)	0.66(0.50)	0.31(0.46)
Rime Change	-1.06(0.53)	-1.57(0.58)	-1.54(0.56)	-0.94(0.46)	-2.00(0.59)
Tonal Change	0.61(0.67)	0.72(0.91)	0.67(0.86)	0.30(0.55)	0.16(0.74)
Incongruent	-2.84(0.51)	-3.98(0.50)	-3.68(0.75)	-1.60(0.47)	-3.03(0.61)

Table D7 con't

450-500ms

	Frontal	Central	Parietal	Temporal	Occipital
Congruent	3.37(0.54)	4.65(0.60)	4.66(0.60)	3.08(0.41)	3.12(0.58)
Onset Change	-1.47(0.46)	-1.11(0.71)	0.30(0.67)	0.81(0.49)	0.77(0.56)
Rime Change	-0.36(0.52)	-0.11(0.61)	0.22(0.64)	0.20(0.49)	-0.24(0.58)
Tonal Change	0.80(0.59)	1.56(0.83)	1.67(0.90)	1.00(0.59)	1.10(0.95)
Incongruent	-2.43(0.54)	-2.72(0.69)	-2.14(0.99)	-0.94(0.67)	-1.63(0.91)

500-550ms

	Frontal	Central	Parietal	Temporal	Occipital
Congruent	2.94(0.56)	4.46(0.60)	4.66(0.67)	3.05(0.39)	3.12(0.69)
Onset Change	-0.86(0.46)	0.31(0.64)	1.90(0.63)	1.73(0.57)	2.22(0.59)
Rime Change	0.12(0.55)	0.87(0.51)	1.95(0.63)	1.28(0.46)	1.67(0.58)
Tonal Change	0.70(0.58)	2.41(0.81)	3.62(0.86)	2.15(0.55)	3.25(0.85)
Incongruent	-2.20(0.56)	-1.71(0.68)	-1.01(0.87)	-0.30(0.59)	-0.73(0.78)

Table D8

ANOVA for N400 Amplitude Data for Condition X Time X Region X Hemisphere Analysis for Experiment 2

Source	df	MSE	F
COND	4, 56	74.19	35.50*
TIME	3, 42	18.31	35.90*
REGION	4, 56	20.12	3.64*
HEMIS	1, 14	40.57	0.70
COND X TIME	12, 168	9.00	3.92*
COND X REGION	16, 224	8.53	5.41*
TIME X REGION	12, 168	1.22	18.39*
COND X TIME X REGION	48, 672	0.72	3.71*
COND X HEMIS	4, 56	8.06	1.82
TIME X HEMIS	3, 42	1.00	2.82
COND X TIME X HEMIS	12, 168	0.94	1.38
REGION X HEMIS	4, 56	7.67	0.31
COND X REGION X HEMIS	16, 224	1.88	2.14
TIME X REGION X HEMIS	12, 168	0.21	1.99
COND X TIME X REGION X HEMIS	48, 672	0.12	1.44

* $p < .05$

APPENDIX E

Table E1

ANOVA for N270 Amplitude Data for Condition X Time X Site Analysis for Experiment 3

Source	df	MSE	F
COND	4, 56	97.52	45.64*
TIME	2, 28	295.07	5.17*
SITE	16, 224	68.07	5.03*
COND X TIME	8, 112	14.46	59.01*
COND X SITE	64, 896	3.08	10.88*
TIME X SITE	32, 448	7.18	7.94*
COND X TIME X SITE	128, 1792	0.44	10.31*

* $p < .05$

Table E2

Mean Amplitudes (in μV) (and Standard Deviations) for the N270 component for the Condition x Time x Site Analysis for Experiment 3.

200-250ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Fz	5.53(1.09)	4.35(0.94)	3.93(0.92)	3.85(1.09)	2.24(0.94)
F3	4.80(0.94)	4.11(0.78)	3.76(0.87)	3.76(1.05)	2.38(0.79)
F4	4.85(1.01)	3.78(0.92)	3.57(0.85)	3.58(0.97)	2.05(0.95)
F7	2.82(0.73)	3.28(0.60)	2.38(0.90)	2.48(0.79)	1.77(0.74)
F8	2.40(0.62)	2.03(0.64)	2.09(0.67)	2.15(0.64)	2.08(0.86)
Cz	6.86(1.22)	4.64(0.97)	4.78(0.97)	4.28(1.25)	2.59(1.07)
C3	5.57(1.07)	4.24(0.88)	4.56(0.91)	4.09(1.06)	2.83(0.88)
C4	6.21(1.20)	4.37(0.97)	4.75(0.94)	4.33(1.10)	3.01(1.04)
Pz	5.18(0.98)	3.47(0.92)	4.02(0.83)	3.55(0.89)	2.29(1.01)
P3	4.92(0.74)	3.54(0.77)	4.20(0.68)	3.40(0.73)	2.50(0.86)
P4	6.23(0.92)	4.27(0.98)	5.23(0.79)	4.49(0.81)	3.45(1.06)
T3	2.41(0.52)	1.88(0.44)	1.96(0.47)	1.10(0.66)	1.51(0.51)
T4	2.31(0.62)	1.45(0.58)	1.75(0.54)	1.85(0.55)	1.32(0.63)
T5	2.33(0.57)	1.47(0.74)	1.81(0.78)	0.90(0.70)	0.70(0.78)
T6	3.54(0.56)	1.91(0.66)	2.83(0.68)	2.24(0.56)	1.74(0.75)
O1	3.72(0.83)	2.25(1.10)	3.03(1.14)	2.19(1.23)	1.97(1.22)
O2	4.63(0.75)	2.72(1.18)	4.13(1.08)	3.28(1.10)	2.62(1.25)

Table E2 con't

250-300ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Fz	7.96(1.40)	3.98(1.31)	3.62(1.28)	3.33(1.30)	2.12(1.01)
F3	7.50(1.28)	4.05(1.18)	3.81(1.24)	3.51(1.31)	2.62(0.95)
F4	7.26(1.25)	3.91(1.20)	3.49(1.10)	3.59(1.13)	2.31(0.97)
F7	4.92(0.84)	4.23(0.87)	3.36(1.05)	3.40(1.02)	3.08(0.69)
F8	4.69(0.79)	3.13(0.85)	3.01(0.82)	3.31(0.85)	3.67(0.88)
Cz	9.86(1.78)	3.15(1.48)	3.01(1.57)	2.80(1.56)	1.08(1.26)
C3	8.94(1.64)	3.45(1.21)	3.58(1.43)	3.26(1.39)	2.05(1.11)
C4	9.32(1.66)	3.59(1.38)	3.60(1.38)	3.65(1.35)	2.20(1.17)
Pz	9.03(1.60)	2.24(1.28)	2.20(1.44)	2.44(1.25)	0.82(1.20)
P3	8.74(1.44)	2.68(1.06)	2.82(1.30)	2.79(1.12)	1.45(1.00)
P4	9.70(1.54)	3.02(1.32)	3.57(1.36)	3.54(1.18)	2.11(1.26)
T3	4.66(0.88)	1.67(0.86)	1.95(1.02)	1.31(1.15)	1.90(0.88)
T4	4.59(0.79)	1.84(0.89)	1.86(0.79)	2.22(0.78)	1.64(0.85)
T5	5.29(0.99)	0.86(0.71)	0.70(0.87)	0.88(0.70)	0.14(0.50)
T6	6.06(0.87)	1.07(0.80)	1.76(0.86)	1.74(0.65)	0.98(0.74)
O1	4.53(1.23)	-1.43(1.10)	-1.80(1.25)	-1.11(1.18)	-2.36(1.06)
O2	4.80(1.01)	-1.93(0.94)	-1.08(1.06)	-0.97(0.91)	-2.52(0.91)

Table E2 con't

300-350ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Fz	9.84(1.45)	1.24(1.06)	1.21(1.21)	0.95(1.37)	-1.11(1.10)
F3	9.37(1.34)	1.48(1.05)	1.68(1.22)	1.14(1.35)	-0.28(1.05)
F4	9.21(1.37)	1.38(0.95)	1.29(1.03)	1.58(1.12)	-0.74(1.08)
F7	5.87(0.74)	2.20(0.90)	1.37(1.09)	0.99(1.12)	0.65(0.81)
F8	5.71(0.87)	1.08(0.82)	1.25(0.81)	2.02(0.92)	1.31(0.93)
Cz	12.40(1.95)	0.15(1.25)	-0.29(1.58)	-0.21(1.73)	-3.32(1.35)
C3	11.31(1.74)	0.64(1.15)	0.74(1.42)	0.37(1.53)	-1.78(1.16)
C4	11.88(1.98)	0.84(1.26)	0.84(1.38)	1.09(1.52)	-1.62(1.32)
Pz	11.59(2.01)	-1.26(1.39)	-1.48(1.59)	-0.91(1.69)	-3.77(1.50)
P3	11.15(1.77)	-0.39(1.17)	-0.47(1.43)	-0.40(1.49)	-2.98(1.32)
P4	11.88(2.07)	-0.12(1.44)	0.31(1.52)	0.42(1.58)	-2.03(1.63)
T3	6.13(0.94)	-0.30(0.84)	-0.46(1.11)	-1.44(1.26)	-1.05(0.96)
T4	6.27(1.13)	-0.37(0.90)	-0.09(0.83)	0.32(0.97)	-1.07(0.98)
T5	7.62(1.32)	-0.68(0.75)	-1.02(0.99)	-1.21(0.90)	-2.60(0.69)
T6	7.82(1.37)	-0.47(0.97)	0.23(0.94)	0.12(0.89)	-1.16(1.15)
O1	6.57(1.59)	-4.19(1.11)	-4.58(1.41)	-3.62(1.29)	-5.94(1.48)
O2	6.04(1.42)	-5.03(1.16)	-4.04(1.29)	-3.85(1.23)	-6.10(1.51)

Table E3

Mean Amplitudes (in µV) (and Standard Deviations) for the N270 component for the Condition x Time x Region Analysis for Experiment 3.

200-250ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Frontal	3.72(0.79)	3.30(0.70)	2.95(0.74)	2.99(0.80)	2.07(0.77)
Central	5.89(1.13)	4.30(0.90)	4.66(0.90)	4.21(1.07)	2.92(0.94)
Parietal	5.57(0.82)	3.91(0.87)	4.72(0.72)	3.94(0.75)	2.98(0.95)
Temporal	2.65(0.42)	1.68(0.45)	2.09(0.43)	1.52(0.42)	1.32(0.45)
Occipital	4.18(0.73)	2.49(1.10)	3.58(1.08)	2.74(1.14)	2.29(1.22)

250-300ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Frontal	6.09(1.00)	3.83(0.98)	3.42(0.96)	3.45(1.02)	2.92(0.79)
Central	9.13(1.64)	3.52(1.28)	3.59(1.39)	3.46(1.35)	2.12(1.12)
Parietal	9.22(1.49)	2.85(1.17)	3.20(1.32)	3.16(1.14)	1.78(1.12)
Temporal	5.15(0.82)	1.36(0.67)	1.57(0.76)	1.54(0.69)	1.17(0.61)
Occipital	4.67(1.09)	-1.68(0.99)	-1.44(1.14)	-1.04(1.02)	-2.44(0.95)

Table E3 con't

300-350ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
Frontal	7.54(1.04)	1.54(0.85)	1.40(0.91)	1.43(1.05)	0.24(0.87)
Central	11.60(1.84)	0.74(1.17)	0.79(1.38)	0.73(1.50)	-1.70(1.21)
Parietal	11.51(1.91)	-0.25(1.29)	-0.08(1.46)	0.01(1.53)	-2.51(1.45)
Temporal	6.96(1.13)	-0.46(0.75)	-0.34(0.87)	-0.55(0.92)	-1.47(0.82)
Occipital	6.30(1.48)	-4.61(1.10)	-4.31(1.33)	-3.74(1.22)	-6.02(1.44)

Table E4

ANOVA for N270 Amplitude Data for Condition X Time X Region X Hemisphere Analysis for Experiment 3

Source	df	MSE	F
COND	4, 56	65.37	41.85*
TIME	2, 28	171.93	6.51*
REGION	4, 56	155.47	6.69*
HEMIS	1, 14	41.13	1.52
COND XTIME	8, 112	8.97	59.88*
COND X REGION	16, 224	4.69	10.43*
TIME X REGION	8, 112	12.30	13.99*
COND X TIME X REGION	32, 448	0.68	11.90*
COND X HEMIS	4, 56	2.35	2.75*
TIME X HEMIS	2, 28	8.22	0.31
COND X TIME X HEMIS	8, 112	0.47	1.11
REGION X HEMIS	4, 56	11.69	1.31
COND X REGION X HEMIS	16, 224	0.86	1.39
TIME X REGION X HEMIS	8, 112	0.94	3.10*
COND X TIME X REGION X HEMIS	32, 448	0.09	2.25

* $p < .05$

Table E5

ANOVA for N270 Peak Latency Data for Condition X Time X Site Analysis for Experiment 3

Source	df	MSE	F
COND	4, 56	2469.20	21.38*
SITE	16, 224	324.85	2.37*
COND X SITE	64, 896	227.20	0.87

* $p < .05$

Table E6

ANOVA for N400 Amplitude Data for Condition X Time X Site Analysis for Experiment 3

Source	df	MSE	F
COND	4, 56	95.26	61.65*
TIME	3, 42	95.10	48.50*
SITE	16, 224	112.24	10.07*
COND X TIME	12, 168	23.74	45.70*
COND X SITE	64, 896	5.68	6.55*
TIME X SITE	48, 672	4.59	8.89*
COND X TIME X SITE	192, 2688	0.57	8.52*

* $p < .05$

Table E7

Mean Amplitudes (in µV) (and Standard Deviations) for the N400 component for the Condition x Time x Region x Hemisphere Analysis for Experiment 3.

350-400ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
L. Frontal	7.85(0.83)	1.20(1.11)	1.13(1.17)	0.74(1.16)	-0.74(1.15)
R. Frontal	7.75(0.83)	1.05(1.00)	1.05(0.78)	1.61(0.96)	0.04(0.92)
L. Central	12.28(1.33)	0.39(1.31)	0.83(1.20)	0.59(1.37)	-2.25(1.24)
R. Central	12.86(1.43)	0.89(1.37)	1.14(1.16)	1.08(1.41)	-1.93(1.29)
L. Parietal	11.68(1.35)	-0.73(1.14)	-0.72(1.15)	-0.77(1.34)	-3.63(1.27)
R. Parietal	12.11(1.49)	-0.48(1.32)	-0.34(1.16)	-0.49(1.37)	-3.32(1.50)
L. Temporal	8.36(0.95)	-0.30(0.66)	-0.24(0.87)	-1.33(0.98)	-1.96(0.75)
R. Temporal	7.84(0.89)	-0.44(0.82)	-0.26(0.58)	-0.39(0.77)	-1.57(0.84)
L. Occipital	6.91(1.33)	-4.47(1.20)	-4.92(1.44)	-4.35(1.35)	-6.86(1.63)
R. Occipital	5.76(1.18)	-5.78(1.23)	-5.28(1.34)	-5.23(1.35)	-7.84(1.55)

Table E7 con't

400-450ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
L. Frontal	7.29(0.65)	2.52(1.10)	3.44(1.15)	2.27(1.20)	1.39(1.15)
R. Frontal	6.89(0.61)	1.93(1.00)	2.80(0.80)	3.10(0.93)	1.96(0.89)
L. Central	11.04(0.98)	2.44(1.13)	4.20(1.08)	3.37(1.26)	1.04(1.21)
R. Central	11.01(1.11)	2.52(1.11)	3.95(0.96)	3.79(1.19)	1.26(1.07)
L. Parietal	9.82(1.03)	1.25(1.02)	2.59(0.99)	2.15(1.24)	-0.56(1.23)
R. Parietal	9.25(1.15)	0.88(1.11)	2.35(0.91)	2.23(1.15)	-0.53(1.41)
L. Temporal	7.32(0.82)	1.04(0.69)	2.13(0.89)	0.90(1.00)	0.28(0.86)
R. Temporal	6.15(0.79)	0.17(0.71)	1.28(0.51)	1.34(0.69)	-0.01(0.87)
L. Occipital	3.80(1.34)	-4.28(1.31)	-3.96(1.51)	-2.59(1.67)	-5.90(1.87)
R. Occipital	2.14(1.08)	-5.82(1.24)	-4.21(1.32)	-3.70(1.42)	-7.04(1.74)

Table E7 con't

450-500ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
L. Frontal	6.02(0.80)	5.05(0.91)	5.80(1.05)	5.04(1.11)	3.50(1.05)
R. Frontal	6.12(0.59)	4.42(0.92)	5.17(0.79)	5.65(0.92)	4.17(0.78)
L. Central	9.36(1.02)	6.60(0.98)	8.39(0.82)	8.42(1.11)	5.12(1.03)
R. Central	9.95(1.05)	6.39(1.10)	8.04(0.77)	8.65(1.10)	5.19(0.98)
L. Parietal	8.35(0.88)	5.30(0.96)	7.20(0.67)	7.47(0.92)	4.15(0.93)
R. Parietal	8.12(0.91)	4.53(1.08)	6.44(0.61)	7.13(0.95)	3.66(1.16)
L. Temporal	5.95(0.66)	4.15(0.63)	5.53(0.64)	5.05(0.80)	3.65(0.77)
R. Temporal	5.19(0.63)	2.23(0.70)	3.75(0.44)	4.26(0.68)	2.51(0.72)
L. Occipital	3.85(0.82)	0.05(1.01)	1.43(1.24)	2.36(1.09)	-0.87(1.41)
R. Occipital	2.59(0.78)	-1.44(0.99)	0.90(1.11)	1.22(1.05)	-1.88(1.40)

Table E7 con't

450-500ms

	Congruent	Onset Change	Rime Change	Tonal Change	Incongruent
L. Frontal	4.22(0.79)	4.02(0.82)	4.25(1.04)	4.83(0.87)	2.28(1.00)
R. Frontal	4.83(0.66)	3.97(0.88)	4.12(0.89)	5.52(0.79)	3.62(0.81)
L. Central	7.14(0.90)	5.88(0.84)	6.99(0.76)	8.17(0.70)	3.82(0.89)
R. Central	7.83(0.93)	6.25(0.92)	6.68(0.84)	8.64(0.75)	4.61(0.83)
L. Parietal	6.47(0.71)	5.41(0.68)	6.47(0.68)	7.89(0.59)	3.63(0.69)
R. Parietal	6.02(0.69)	5.15(0.80)	5.82(0.68)	7.77(0.64)	3.90(0.71)
L. Temporal	4.30(0.53)	4.38(0.44)	5.02(0.66)	5.77(0.63)	3.29(0.53)
R. Temporal	3.85(0.48)	3.21(0.65)	3.78(0.56)	5.22(0.64)	3.37(0.59)
L. Occipital	2.65(0.62)	0.99(0.86)	1.32(1.21)	3.49(0.99)	-0.06(1.05)
R. Occipital	1.54(0.51)	0.03(0.82)	1.16(0.96)	2.68(0.99)	-0.55(0.97)

Table E8

ANOVA for N400 Amplitude Data for Condition X Time X Region X Hemisphere Analysis for Experiment 3

Source	df	MSE	F
COND	4, 56	64.41	57.16*
TIME	3, 42	56.47	50.42*
REGION	4, 56	244.79	14.39*
HEMIS	1, 14	26.67	1.87
COND X TIME	12, 168	15.03	44.03*
COND X REGION	16, 224	9.02	6.28*
TIME X REGION	12, 168	9.10	7.62*
COND X TIME X REGION	48, 672	0.82	9.28*
COND X HEMIS	4, 56	5.65	1.89
TIME X HEMIS	3, 42	7.04	1.85
COND X TIME X HEMIS	12, 168	0.73	2.26
REGION X HEMIS	4, 56	18.11	2.22
COND X REGION X HEMIS	16, 224	1.27	2.60*
TIME X REGION X HEMIS	12, 168	0.57	4.08*
COND X TIME X REGION X HEMIS	48, 672	0.09	2.38*

* $p < .05$

References

Allison, T., Wood, C.C., & McCarthy, G.M. (1986). The central nervous system. In *Psychophysiology: systems, processes, and applications.* M.G.H. Coles, E. Donchin, & S.W. Porges (Eds.). Guilford, New York.

Anderson, J.E., & Holcomb, P.J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology*, 32, 177-190.

Aull, A.M. (1984). Lexical stress and its application in large vocabulary speech recognition. *Unpublished master's thesis*, M.I.T. Massachusetts.

Baddeley, A. (1982). Reading and working memory. *Bulletin of the British Psychological Society*, 414-417.

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception & Performance*, 10, 340-357.

Bansal, R.K. (1966). The intelligibility of Indian English. *Unpublished doctoral dissertation*, University of London, UK.

Bard, E.G., Shillcock, R.C., & Altmann, G.T.M. (1988). The recognition of words after their acoustic offsets in spontaneous speech: effects of subsequent context. *Perception & Psychophysics*, 44, 395-408.

Barron, R.W. (1978). Access to the meaning of printed words: some implications for reading and learning to read. In F.B. Murray (Ed.), *The recognition of words: IRA series on the development of the reading process*, 34-56. Newark, DE: International Reading Association.

Basso, A., Casati, G. & Vignolo, L.A. (1977). phonemic identification defect in aphasia. *Cortex*, 13: 85-95.

Berger, H. (1929). Uber das elektrekephalogramm des menschen. *Arch Psychiatry Nervenkr,* 87, 527-570.

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other nonlinguistic contexts. *Psychophysiology*, 24, 14-25.

Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain and Language*, 31, 308-327.

Bentin, S., Kutas, M., & Hillyard, S.A. (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Psychophysiology*, 30, 161-169.

Bentin, S., McCarthy, G., & Wood, C.C. (1985). Event-related potentials associated with semantic priming. *Electroencephalography and Clinical Neurophysiology*, 60, 343-355.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M.H., et al. (1999). ERP Manifestations of Processing Printed Words at Different Psycholinguistic levels: Time Course and Scalp Distribution. *Journal of Cognitive Neuroscience*, 11(3) 235-260.

Bentin, S., Kutas, M., & Hillyard, S. A. (1995). Semantic processing and memory for attended and unattended words in dichotic listening: behavioral and electrophysiological evidence. *Journal of experimental psychology: Human perception and performance*, 21, 54-67.

Blank, M. (1978). Review of 'Toward an understanding of dyslexia: Psychological factors in specific reading disability'. In: A.L. Benton & D. Pearl (Eds.), *Dyslexia: An appraisal of current knowledge* (pp 115-122). New York: Oxford University Press.

Blumstein, S.E., Baker, E. & Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia*, 15: 19-30.

Bolinger, D. L. (1965). The atomization of meaning. *Language*, 41, 555-573.

Bolinger, D. (1989). Intonation and its uses: Melody in grammar and discourse. Stanford: Stanford University press.

Bond, Z.S., & Garnes, S. (1980). Misperception of fluent speech. In R. Cole (Ed.), *Perception and Production of Fluent Speech*, pp.115-132. Hillsdale, NJ: Erlbaum.

Bond, Z.S., & Small, L.H. (1983). Voicing, vowel, and stress mispronunciations in continuous speech. *Perception & Psychophysics*, 34, 470-474.

Bradley, D.C. (1980). Lexical representation of derivational relation. In M. Aronoff & M.L. Keen (Eds.), *Juncture*. Saratoga, CA: Anma Libri.

Bradley, D.C., & Forster, K.I. (1987). A reader's view of listening. *Cognition,* 25.

Brandeis, D., & Lehmann, D. (1986) Event-related potentials of the brain and cognitive processes: approaches and applications. *Neuropsychologia*, 24(1): 151-168.

Bransford, J.D., & Johnson, M.K. (1972). Contextual Prerequisites for Understanding: Some Investigations of Comprehension and Recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717-726.

Brown, C.M., & Hagoort, P. (1993). The processing nature of the N400: evidence from masked priming. *Journal of Cognitive Neuroscience*, 5(1), 34-44.

Carlson, R., Grandstrom, B., Lindblom, B., et al. (1973). Some timing and fundamental frequency characteristics of Swedish sentences: Data, rules, and perceptual evaluation. *Speech Transmission Laboratory Quarterly Progress Status Report*, 4, 11-19.

Carpenter, P.A., & Daneman, M. (1981). Lexical retrieval and error recovery in reading: a model based on eye fixations. *Journal of Verbal Learning and Verbal Behavior*, 20, 137-160.

Carrell, T.D., Smith, L.B., & Pisoni, D.B. (1981). Some perceptual dependencies in speeded classification of vowel color and pitch. *Perception & psychophysics* 29: 1-10.

CELEX Dutch database (Release N31) (CD-Rom) (1990). Nijmegen, Netherlands: Centre for Lexical Informaiton (Producer and Distributor).

Chao, Y.R. (1948). Mandarin primer. Cambridge: Harvard University press.

Chao, Y.R. (1968). A grammar of spoken chinese. Berkeley, CA: University of California Press.

Chao, Y.R. (1980). Chinese tones and English stress. In *the Melody of language*, 41-44. Baltimore, MD: University Park Press.

Chen, H.C., Flores d'Arcais, G.B., & Cheung, S.L. (1995). Orthographic and phonological activation in recognizing Chinese characters. *Psychological research*, 58, 144-153

Chen, H.C. (1998). Semantics without phonology: evidence from replication of Perfetti and Tan (1998). *Paper presented at the 39th Annual Meeting of the Psychonomic Society*, Dallas, Texas, USA

Ching, Y.C. (1985). Lipreading Cantonese with voice pitch. Paper presented to the 109th meeting, Acoustical Society of America, Austin. *Abstract Journal of the Acoustical Society of America*, 77, Supplement 1, S39-40.

Ching, Y.C. (1988). Voice pitch information for the deaf. *Proceedings of the first Asian-Pacific Regional Conference on Deafness*, 340-343. Hong Kong.

Chomsky N., & Halle, M. (1968). The sound pattern of English. New York: Harper & Row.

Chua, F. (1999). Phonological recoding in Chinese logograph recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 876-891.

Cluff, M.S., & Luce, P.A. (1990). Similarity neighborhoods of spoken two-syllable words: retroactive effects on multiple activation. *Journal of experimental psychology: human perception & performance*, 16, 551-563

Cole, R.A., & Jakimik, J. (1978). Understanding speech: how words are heard. In G. Underwood (Ed.), *Strategies of information processing*, 67-116. London: Academic Press.

Cole, R.A., & Jakimik, J. (1980a). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech*, 133-163. Hillsdale, NJ: Erlbaum.

Cole, R.A., & Jakimik, J.(1980b). How are syllables used to recognize words? *Journal of the Acoustical Society of America*, 67, 965-970.

Coles, M.G.H., & Rugg, M.D. (1995). Event-related brain potentials: an introduction. M.D. Rugg & M.G.H. Coles (Eds.) *Electrophysiology of Mind*. Oxford.

Coltheart, M., Davelaar, E., Jonasson, J.T., et al. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535-555). New York: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., et al. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, 100, 589-608

Coltheart, M., Patterson, K.E., & Marshall, J.C. (1980). *Deep Dyslexia*. London: Routledge and Kegan Paul.

Collins, A.M., & Quillian, M.R. (1969). Retrieval from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.

Connine, C.M., Blasko, D.G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: temporal and linguistic constraints. *Journal of memory and language*. 30: 234-250

Connine C.M., Blasko, D.G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of memory & language*, 32: 193-210

Connine, C.M., Clifton, C., JR., & Cutler, A. (1987). Effects of lexical stress and phonetic categorization. *Phonetica*, 44, 133-146.

Connolly, J.F., Byrne, J.M., & Dywan, C.A. (1995). Assessing adult receptive vocabulary with event-related potentials: an investigation of cross-modal and cross-form priming. *Journal of Clinical and Experimental Neuropsychology,* 17(4), 548-565.

Connolly, J.F., & Kleinman, K.M. (1978). A single channel method for recording vertical and lateral eye movements. *Electroencephalography and clinical neurophysiology*, 45, 128-129.

Connolly, J.F., Service, E., D'Arcy, R.C.N., Kujala, A., & Alho K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *Neuroreport: For Rapid Communication of Neuroscience Research,* 12, 237-243.

Connolly, J.F., Stewart, S.H., & Phillips, N.A. (1990). The effects of processing requirements on neurophysiological responses to spoken sentences. *Brain and language*. 39, 302-318.

Connolly, J.F., Phillips, N.A., Stewart S.H., et al. (1992). Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. *Brain and language*. 43: 1-18.

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(3): 256-266.

Connolly, J.F, Phillips, N.A., & Forbes K.A.K. (1995). The effects of phonological and semantic features of sentence-ending words on visual event-related brain potentials. *Electroencephalography and Clinical Neurophysiology,* 94, 276-287.

Cutler, A. (1986). Forbear is a homophone: lexical prosody does not constrain lexical access. *Language and speech*, 29: 201-220

Cutler, A. (1989). Auditory lexical access: where do we start? In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 342- 356). Cambridge, MA: MIT Press

Cutler, A. (1992). Auditory lexical access: where do we start? *Lexical representation & process*. The MIT press. pp343-356

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: evidence from juncture misperception. *Journal of memory & language*, 31, 218-236.

Cutler, A. & Carter D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer speech and language,* 2: 133-142

Cutler, A., & Chen, H.C. (1995). Phonological similarity effects in Cantonese word recognition. *Proceedings of the Thirteenth International Congress of Phonetic Sciences*, 106-109. Stockholm.

Cutler, A., & Chen, H.C., (1997). Lexical tone in Cantonese spoken word processing. *Perception and psychophysics.* 59 (2): 165-179.

Cutler, A., & Clifton, C., Jr. (1984). The use of prosodic information in word recognition. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and Performance X: Control of language processes.* Hillsdale, NJ: Erlbaum.

Cutler, A., & Foss, D.J. (1977). On the role of sentence stress in sentence processing. *Language and speech,* 20, 1-10.

Culter, A., Mehler, J., Norris, D., et al. (1983). A language- specific comprehension strategy. *Nature*, 304: 159-160

Culter, A., Mehler, J., Norris, D., et al. (1986). The syllable's differing role in the segmentation of French and English. *Journal of memory and language*, 25: 385-400

Culter A., & Norris D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of experimental Psychology:Human perception and performance*, 14(1), 113-121

Daneman, M., Reingold, E., & Davidson, M. (1995). Time course of phonological activation during reading: evidence from eye fixations. *Journal of experimental psychology: learning, memory and cognition*, 21, 884-898

Darwin, C.J. (1975). On the dynamic use of prosody in speech perception. In A. Cohen & S.G. Nooteboom (Eds.), *Structure and process in speech perception*, 178-193. Berlin: Springer- Verlag.

Divenyi, P.L. & Robinson, A.J. (1989) Nonlinguistic auditory capabilities in aphasia, *Brain and Language*, 37, 290-326

Donchin, E., Ritter, W., & McCallum, W.C. (1978). Cognitive Psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, & S.H. Koslow (Eds.), *Event-related brain potentials in man* (pp. 349-441). New York: Academic Press.

Dupoux, E., Pallier, C., Sebastian, N., et al. (1997). A distressing deafness in French?

Journal of Memory and Language, 36, 406- 421.

Edfelt, A.W. (1960). Silent speech and silent reading. Chicago: University of Chicago Press.

Elman, J.L., & McClelland, J.L. (1986). Exploiting lawful variability in the speech waveform. In J.S. Perkell & D.H. Klatt (Eds.), *Invariance and variability in speech processes*. Hillsdale, NJ: Erlbaum.

Eimas, P.D., Tartter, V.C., Miller J.L. et al. (1978). Asymmetric dependencies in processing phonetic features, *Perception and Psychophysics* 23: 12-20.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology*, 45A, 353-372.

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology*, 47A, 365-382.

Fiez, J.A., & Petersen, S.E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of sciences*, 95: 914-921

Fischler, I., Bloom, P.A., Childers, D.G., et al. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, 20, 400-409.

Fischler, I., Bloom, P.A., Childers, D.G., et al. (1984). Brain potentials during sentence verification: late negativity and long-term memory strength. *Neuropsychologia*, 22, 559-568.

Fischler, I., Childers, D.G., Achariyapaopan, T., et al. (1985). Brain Potentials during sentence verification: automatic aspects of comprehension. *Biological Psychology*, 21, 83-106.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological Bulletin*. 123(1): 71-99

Forbes, K.A.K. (1993). Memory performance, event-related brain potentials and phonological recoding during silent reading. *Unpublished master dissertation*, Canada.

Forbes, K.A.K. (1998). Event-related brain potentials (ERPs) measure the influences of orthographic, phonological and semantic representations during silent reading. *Unpublished doctoral dissertation*, Canada.

Forster, K.I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology*, 33A, 465-495.

Fox, R.A., & Unkefer, J. (1985). The effect of the lexical status on the perception of tone. *Journal of Chinese Linguistics*, 13, 69-90.

Frauenfelder, U.H., Segui J., & Dijkstra T. (1990). Lexical effects in phonemic processing: facilitatory or inhibitory? *Journal of experimental psychology: human perception & performance*, 16: 77-91

Frauenfelder, U.H., & Tyler, L.K. (1987). The process of spoken word recognition: an introduction. *Cognition*, 25, 1-20.

Fromkin, V., & Rodman, R. (1993). An introduction to language (5th ed.) New York: Harcourt Brace Jovanovich.

Fry, D.B. (1955). Duration and intensity of physical correlates of linguistic stress. *Journal of the acoustical society of America*, 27, 765-768

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

Gandour, J. (1988). Perception and production of tone in aphasia. *Brain and Language,* 35, 201-240.

Gandour, J., Ponglorpisit S., Potisuk S., et al. (1997). Interaction between tone and intonation in Thai after unilateral brain damage., 58, 174-196.

Gandour, J., Potisuk, S., Dechongkit, S., et al. (1992a). Tonal coarticulation in Thai disyllabic utterances: a preliminary study. *Linguistics of the Tibeto-Burman Area*, 15 (1): 93-110

Gandour, J., Potisuk, S., Dechongkit, S., et al. (1992b). Anticipatory tonal coarticulation in Thai noun compounds. *Linguistics of the Tibeto-Burman Area*, 15 (1): 111-124.

Gandour, J., Wong, D., & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport,* 9, 2115-2119.

Gandour, J.(2000). Aphasia in tone languages. *Aphasia in Atypical population.*

Gandour, J., Wong D., Hsieh L., et al. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, 12(1): 207-222

Garding, E. (1967). Internal structure in swedish. Lund, Sweden: Gleerups.

Garner, W.R. (1970). the stimulus in information processing. *American psychologist*, 25: 350-358

Garner, W.R., & Felfoldy G.L.(1970). Integrality of stimulus dimensions in various types of information processing, *Cognitive psychology*, 1: 225-241.

Givon, T. (1975). Serial verbs and syntactic change: Niger-Congo. In C.N. Li (Ed.) *Word order and word order change*. Austin/ London: University of Texas Press.

Goldinger, S.D. (1999). Only the shadower knows: comment on Hamburger and Slowiaczek (1996). *Psychonomic Bulletin & Review*, 6, 347- 351.

Goldinger, S.D., Luce, P.A., & Pisoni, D.B. (1989). Priming lexical neighbors of spoken words: effects of competition and inhibition. *Journal of Memory and Language*, 28, 501-518.

Goldinger, S.D., Luce, P.A., Pisoni, D.B., et al. (1992). Form-based priming in spoken word recognition: the roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 18(6), 1211-1238

Goldsmith J. (1976). An overview of autosegmental phonology. *Linguistic Analysis,* 2(1): 23-68.

Goodglass, H. (1993). Understanding aphasia. San Diego, CA: Academic Press.

Goodglass, H., Wingfield, A., Hyde, M.R., et al. (1997). The importance of word-initial phonology: Error patterns in prolonged naming efforts by aphasic patients. *Journal of the International Neuropsychological Society*, 3, 128-138.

Goodman J. & Huttenlocher J. (1988). Do we know how people identify spoken words? *Journal of memory & language*, 27: 684-698

Gottfried T.L. & Suiter T.L. (1997). Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics,* 25: 207-231

Gow, D.W., & Gordon, P.C. (1993). Coming to terms with stress: Effects of stress location in sentence processing. *Journal of Psycholinguistic Research*, 22, 545-578.

Graf, P., Mandler, G., & Haden, P. (1982). Simulating amnesic symptoms in normal participants. *Science*, 218, 1243-1244.

Greenhouse, S.W. & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24: 95-112.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Percpetion & Psychophysics*, 28, 267-283

Grosjean, F. (1985). The recognition of words after their acoustic offset: evidence and implications. *Perception & Psychophysics*, 38, 299-310

Grosjean F. (1996). Gating. *Language and Cognitive Processes*, 11: 597-604

Grosjean, F., & Gee, J.P. (1987). Prosodic structure and spoken word recognition. *Cognition*, 25, 135-155.

Hagoort, P., & Brown, C.M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia*, 38, 1518-1530.

Hamburger M., & Slowiaczek, L.M. (1996). Phonological priming reflects lexical competition. *Psychonomic Bulletin & Review*, 3(4), 520-525.

Helenius, P., Salmelin, R., Service E., & Connolly, J.F. (1999). Semantic Cortical Activation in Dyslexic Readers. *Journal of Cognitive Neuroscience*, 11(5) 535-550.

Henderson, L. (1982). Orthography and word recognition in reading. London: Academic Press.

Henderson, L., Wallis, J., & Knight, D. (1984). Morphemic structure and lexical access. Bouma, H., Bouwhuis, D.G. (Eds.), *Attention and Performance X: Control of Language Processes*. Hillsdale, New Jersey. Pp 211-226.

Hillinger, M.L. (1980). Priming effects with phonemically similar words: The encoding-bias hypothesis reconsidered. *Memory & Cognition*, 8, 115-123.

Hua Z., & Dodd B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language,* 27: 3-42.

Holcomb, P.J., & Anderson, J.E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes*, 8, 379-412.

Holcomb, P.J. & Neville, H.J. (1990). Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. *Language Cognitive Processes*, 5, 281-312.

Holcomb, P.J. & Neville, H.J. (1991). The electrophysiology of spoken sentence processing. *Psychobiology*, 19: 286-300.

Holcomb, P.J. (1993). Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. *Psychophysiology*, 30: 47-61.

Huttenlocher, D.P., & Zue, V.W. (1983). Phonotactic and lexical constraints in speech recognition. *Proceedings of the American Association for Artificial Intelligence*, 172-176.

Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 110, 306-340.

Jakimik, J. (1979). The interaction of sound and knowledge in word recognition from fluent speech. *Unpublished doctoral dissertation*, Carnegie Mellon University, Pittsburgh.

Jakimik, J., Cole, R.A., & Rudnicky, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language*, 24, 165-178.

Jared, D., & Seidenberg, M.S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of experimental psychology: general*, 120, 358-394

Jasper, H.H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalography and clinical neurophysiology*, 10: 371-375.

Johnson-Laird, P.N., Herrmann, D.J., & Chaffin, R. (1984). Only connections: A critique of semantic networks. *Psychological Bulletin*, 96, 292-315.

Jongenburger, W. (1996). The role of lexical stress during spoken word processing. *Unpublished doctoral dissertation*, Leiden University, The Netherlands.

Jongenburger, W., & van Heuven, V.J. (1995). The role of linguistic stress in the time course of word recognition in stress-accent languages. *Proceedings of the Fourth European Conference on Speech Communication and Technology*, 1695- 1698. Madrid.

Katz, J.J., & Fodor, J.A. (1963). The structure of a semantic theory. *Language*, 39, 170-210.

Klatt, D.H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical Representation and Process,* MIT press, Cambridge, MA.

Klatt, D.H., & Stevens, K.H. (1973). On the automatic recognition of continuous speech: implications from a spectrogram reading experiment. *IEEE transactions on audio & electroacoustics,* 21, 210-217.

Knight, R.T. (1990). Neural Mechanisms of Event-related Potentials: Evidence from Human Lesion Studies. In J.W. Rohrbaugh, R. Parasuraman, R. Johnson (Eds.) *Event-related Brain Potentials: Basic Issues and Applications.*

Kozhevnikov, V., & Chistovich, L. (1966). Speech: Articulation and Perception (Translation JPRS 30,543). Washington, DC: U.S. Department of Commerce, Joint Publications Research Service.

Kramer, A.F. & Donchin, E., (1987). Brain potentials as indices of orthographic and phonological interaction during word matching. *Journal of Experimental Psychology: Learning, memory, & Cognition.* 13: 76-86.

Kutas, M. (1985). Event-related brain potentials in bilinguals during reading. Presented at the *13th Annual Meeting of the International Neuropsychological Society*, San Diego, California.

Kutas, M. (1997). Views on how the electrical activity that the brain generates reflects the functions of different language structures. *Psychophysiology,* 34: 383-398.

Kutas, M. (1998). Current thinking on language structures. *Current Psychology of Cognition*, 17(4-5), 951-969.

Kutas, M., & Federmeier, K.D., (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, 4, 463-470

Kutas, M., & Hillyard, S.A (1980a). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science,* 207: 203-5.

Kutas, M., & Hillyard, S.A. (1980b). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, 11, 99-116.

Kutas, M., & Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, 11, 539-550.

Kutas, M., & Hillyard, S.A. (1984). Brain Potentials during reading reflect word expectancy and semantic association. *Nature*, 307, 161-163.

Kutas, M., Lindamood, T., & Hillyard, S.A. (1984). Word expectancy and event-related brain potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory states and Processes* (pp. 217-238). Hillsdale, NJ: Erlbaum.

Kutas M., & Iragui, V. (1998). The N400 in a semantic categorization task across 6 decades. *Electroencephalography and Clinical Neurophysiology*, 108, 456-471

Kutas, M., Neville, H.J., & Holcomb, P.J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening, and signing. In R.J. Ellingson, N.M.F. Murray, & A.M. Hallidy (Eds.), *Electroencephalo-graphy and Clinical Neurophysiology* Supplement 39: the London Symposia (pp. 325-330). Amsterdam: Elsevier.

Kutas, M., & Van Petten, C. (1988). Event-related brain potential studies of language. *Advances in Psychophysiology*, 3: 139-187.

Kutas, M., & Van Petten C.K. (1994). Psycholinguistics Electrified: Event-Related Brain Potential Investigations. *Handbook of Psycholinguistics*. Academic Press. Inc.

Kutas, M., & Hillyard, S.A. (1989). An electrophysiological probe of incidental semantic association. *Journal of cognitive neuroscience*. 1: 38-49

Leck, K.J., Weekes, B.S. & Chen, M.J. (1995). Visual and phonological pathways to the lexicon: evidence from Chinese readers. *Memory & cognition*, 23, 468-476

Lee Y.S., Vakoch D.A., & Wurm L.H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research,* 25 (5): 527-542.

Lee, L., & Nusbaum, H.C. (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception and Psychophysics,* 53 (2), 157-165.

Lee, Y.S., Vakoch, D.A., & Wurm, L.H. (1996). Tone perception in Cantonese and Mandarin: a cross-linguistic comparison. *Journal of Psycholinguistic Research,* 125, 527-542.

Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. *Phonetica*, 5 (suppl.)

Lehiste, I. (1970). Suprasegmentals. Cambridge, Mass.: MIT Press

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the acoustical society of America,* 51, 2018-2024.

Lesch, M.F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of experimental psychology: learning, memory, and cognition*, 19: 285-294

Levelt, W.J.M. (1989). Speaking. Cambridge, MA: MIT Press.

Levelt, W.J.M. (1992). Accessing words in speech production: stages, processes and representation. *Cognition*, 42, 1-22.

Levelt, W.J.M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50, 239-269.

Levelt, W.J.M., Roelofs, A., & Meryer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences,* 22(1), 1-75

Lewine, J.D., & Orrison, W.W. (1995). Clinical Electroencephalography and Event-related Potentials. *Functional Brain Imaging.*

Liberman, A.M. (1970). Some characteristics of perception in the speech mode. In P. Hamburg, K. Pribram, & A.J. Standard (Eds.), *Perception and its disorders* (pp. 238-254). Baltimore: Williams & Wilkins.

Lieberman, P. (1965). On the acoustic basis of perception of stress by linguists. *Word*, 21, 40-54.

Lin, H.B. (1988). Contextual stability of Taiwanese tones. Unpublished doctorate thesis, University of Connecticut, USA.

Lindfield, K.C., Wingfield A., & Goodglass H. (1999). The role of prosody in the mental lexicon. *Brain and Language*. 68, 312-317.

Luce, P.A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics,* 39, 155-158.

Luce, P.A., & Cluff, M.S. (1998). Delayed commitment in spoken word recognition: evidence from cross-modal priming. *Perception & Psychophysics,* 60 (3), 484-490.

Lukatela, G., Lukatela, K., & Turvey, M.T. (1993). Further evidence for phonological constraints on visual lexical access: TOWED primes FROG. *Perception & psychophysics*, 53, 461-466.

Lukatela, G., & Turvey, M.T. (1994). Visual lexical access is initially phonological: 1. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of experimental psychology: general*, 123, 107-128

Marslen-Wilson W.D.(1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244: 522-523

Marslen-Wilson W. D. (1984). Function and process in spoken word recognition. In H.Bouma & D.G. Bouwhuis (Eds.), *Attention & performance X: Control of language processes* (pp. 125-150). Hillsdale, NJ:Erlbaum

Marslen-Wilson W.D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25: 71-102

Marslen-Wilson, W.D. (1989). Access and integration: processing sound onto meaning. Marslen-Wilson W. D.(Ed.). *Lexical Representation and Process*, M.I.T. Press.

Marslen-Wilson, W.D. (1994). Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychological review,* 101(4) 653-675.

Marslen-Wilson, W.D., Moss, H.E., & Halen, S.V. (1996). Perceptual Distance and Competition in Lexical Access. *Journal of Experimental Psychology: Human Perception and Performance,* 22(6) 1376-1392.

Marslen-Wilson, W.D., & Tyler, L.K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.

Marslen-Wilson, W.D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychological Review*, 101(4): 653-675.

Marslen-Wilson W.D., & Welsh A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10: 29-63

Marslen-Wilson W.D., & Zwitserlood P. (1989). Accessing spoken words: the importance of word onsets. *Journal of experimental psychology: human perception & performance*, 15: 576-585

Massaro, D.W. (1975). Acoustic features in speech perception. In D.W. Massaro (Ed.), *Understanding Language: An Information Processing Analysis of Speech Perception, Reading, and Psycholinguistics* (pp. 77-124). New York: Academic Press.

Mattys S.L. (1997). The use of time during lexical processing and segmentation: a review. *Psychonomic Bulletin & Review,* 4(3), 310-329

Mattys, S.L., (2000). The perception of primary and secondary stress in English. *Perception & Psychophysics,* 62 (2), 253-265.

Mattys, S.L., & Samuel, A.G. (2000). Implications of stress-pattern differences in spoken-word recognition, *Journal of Memory & Language,* 42, 571-596.

Mattys, S., Jusczyk, P.W., Luce, P.A., et al. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465-494.

Mattys, S., & Samuel, A.G. (2000). Implications of stress-pattern differences in spoken-word recognition. *Journal of Memory and Language*, 42, 571-596.

McCallum, W.C., Farmer, S.F., & Pocock, P.K. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 59, 477-488.

McCarthy, G., & Nobre, A.C. (1993). Modulation of semantic processing by spatial selective attention. *Electroencephalography and clinical Neurophysiology*, 88, 210-219.

McCarthy, G. & Wood, C.C. (1985). Scalp distribution of event related potentials: an ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, 62: 203-208

McClelland, J.L. & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.

McClelland, J.L., & Elman, J.L. (1986). Interactive processes in speech recognition: The TRACE model. In J.L. McClelland & D.E. Rumelhart (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Cambridge, MA: Bradford.

McCusker, L.X., Hillinger, M.L., & Bias, R.G. (1981). Phonological recoding and reading. *Psychological Bulletin*, 89, 217-245.

McQueen J.M., & Cutler A. (1992). Words within words: lexical statistics and lexical access. In *Proceedings of the International Conference on Spoken Language Processing* (Vol. 1, pp 221-224). Banff, Alberta, Canada: University of Alberta.

McQueen J.M., Cutler A., Briscoe T., et al. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language & cognitive processes.* 10: 309-331

McQueen J.M., Norris D., & Cutler A. (1994). Competition in spoken word recognition: spotting words in other words. *Journal of experimental* *psychology: learning, memory, and cognition,* 20(3), 621-638

Mehler, J. (1981). The role of syllables in speech processing: infant and adult data. *Philosophical Transactions of the Royal Society*, London, b295: 333-352

Mehler J., Dommergues J.Y., Frauenfelder U. et al. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20, 298-305

Meyer, D.E., Schvaneveldt, R.W., & Ruddy, M.G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory & Cognition*, 2, 309-321.

Miller J.L. (1978) Interactions in processing segmental and suprasegmental features of speech. *Perception and Psychophysics,* 24(2), 175-180.

Miller, J.L., & Eimas, P.D. (1995). Speech perception: from signal to word. *Annual Review of Psychology,* 46, 467-492.

Moen, I., & Sundet, K. (1996). Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage. *Brain and language,* 53, 267-281.

Monsell, S., & Hirsh, K.W. (1998). Competitor priming in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1495-1520.

Morgan, J.L. (1996). A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language*, 35, 666-688.

Nakatani, L.H., & Dukes, K.D. (1977). Locus of segmental cues for word juncture. *Journal of the acoustical society of America*, 62, 714-719.

Nakatani, L.H., & Schaffer, J.A. (1978). Hearing "words" without words: prosodic cues for word perception. *Journal of the acoustical society of America*, 63, 234-245.

Neely, J.H. (1991). Semantic priming effects in visual word recognition: a selective review of current findings and theories. Besner, D., Humphreys, G.W. (Eds.) *Basic Processes in Reading: Visual Word Recognition.* Hillsdale, New Jersey.

Newman, R.L. (2000). Defining phonology's role in silent reading with event-related brain potentials. *Unpublished master's dissertation*, Canada.

Nobre, A.C., Allison, A.T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature,* 372 (6503), 260-263.

Nooteboom, S.G. (1979). The time course of speech perception. In W.J. Barry & K.J. Kohler (Eds.), *'Time' in the production and perception of speech.* University of Kiel.

Nooteboom, S.G. (1981). Lexical retrieval from fragments of spoken words: Beginnings versus endings. *Journal of Phonetics*, 9, 407-424.

Nooteboom, S.G., Brokx, J.P.L., & De Rooij, J.J. (1978). Contribution of prosody to speech perception. In W.J.M. Levelt & G.B. Flores d'Arcais (Eds.), *Studies in the Perception of Language*, pp. 75-107. New York: Wiley.

Norris, D.(1981). Autonomous processes in comprehension. *Cognition*, 11: 97-101

Norris, D.G. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, 52, 189-234

Norris, D., McQueen, J.M., & Culter, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of experimental Psychology: Learning, memory, and cognition.* 21: 1209-1228.

Nusbaum, H.C. & Slowiaczek, L.M. (1983). An activation model of the Cohort theory of auditory word recognition. Paper presented at *the Sixteenth Annual Meeting of the Society for Mathematical Psychology.* Boulder, CO.

Osterhout, L. & Holcomb, P.J. (1993). Event-related potentials and syntactic anomaly: evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, 8, 413-437.

Osterhout L. and Holcomb P.J., (1995). Event-related potentials and language comprehension. *Electrophysiology of Mind: Event-related brain potentials and Cognition.* Oxford Science publications.

Ottevanger I.B. (1984). The detection of mispronunciations and the influence of context. In M.P.R. van den Broecke & A. Cohen (Eds.), P*roceedings of the Tenth International Congress of Phonetic Sciences* (Vol. IIb, pp. 486-490). Dordrecht: Foris

Patterson, K., & Coltheart, V. (1987). Phonological processes in reading: a tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 421-447). Hillsdale, NJ: Erlbaum.

Perfetti, C.A., & Tan, L.H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of experimental psychology: learning, memory, and cognition*, 24 (1)101-118

Perfetti, C.A., & Zhang, S. (1991). Phonological processes in reading Chinese words. *Journal of experimental psychology: learning, memory, and cognition*, 17, 633-643

Perfetti, C.A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 21: 24-33.

Peter, M., & Turvey, M.T. (1994). Phonological codes are earlier sources of constraint in visual semantic categorization. *Perception & psychophysics*, 55, 497-504

Pisoni, D.B., & Luce, P.A. (1987). Acoustic-Phonetic Representations in Word Recognition. *Cognition*, 25, 21-52.

Pisoni, D.B., Nusbaum, H.C., Luce, P.A., et al. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, 4, 75-95.

Pitt, M.A., & Samuel, A.G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 564-573.

Poldrack, R.A., Wagner, A.D., Prull, M.W., et al. (1999). Functional specialization for semantic and phonological processing in the left inferior frontal cortex. *Neuroimage*, 10, 15-35

Poldrack R.A., Temple e., Protopapas A., et al. (2001). Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. *Journal of Cognitive Neuroscience,* 13(5):687-697

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology,* 41(2) 103-146.

Polich, J.M., McCarthy, G., Wang, W.S., et al. (1983). When words collide: orthographic and phonological interference during word processing. *Biological Psychology*, 16, 155-180.

Praamstra, P., & Stegeman D.F. (1993). Phonological effects on the auditory n400 event-related brain potential. *Cognitive Brain Research*, 1, 73-86

Pugh, K. R., Shaywitz, B.A., Shaywitz, S.E., et al. (1996). Cerebral organization of component processes in reading. *Brain*, 119: 1221-1238.

Rack, J.P., Snowling, M.J. & Olson, R.K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly,* 27: 28-53

Radeau, M., & Colin, C. (1996). Task effect in phonological priming between spoken words. Paper presented at the *37th Annual Meeting of the Psychonomics Society*, Chicago.

Radeau, M., Morais, J., & Dewier, A. (1989). Phonological priming in spoken word recognition: Task effects. *Memory & Cognition*, 17, 525-535.

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception & Performance*, 21, 1297-1311

Rayner, K., Sereno, S.C., Lesch, M.F., et al. (1995). Phonological codes are automatically activated during reading: evidence from an eye movement priming paradigm. *Psychological Science*, 6, 26-30

Rayner K., Pollatsek, A. & Binder, K.S. (1998). Phonological codes and eye movements in reading. *Journal of experimental psychology: learning, memory, & cognition,* 24(2), 476-497

Reddy, R. (1976). Speech recognition by machine: a review. *Proceedings of the IEEE*, 64, 501-531.

Repp B.H., & Lin H.B. (1990). Integration of segmental and tonal information in speech perception: a cross-linguistic study. *Journal of Phonetics,* 18: 481-495.

Rips, L.J., Shoben, E.J., & Smith, E.E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 12, 1-20.

Robinson, G.M. (1977). Rhythmic organization in speech processing. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 83-91.

Rubinstein, H., & Garfield, L., & Millikan, J.A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning & Verbal Behavior*, 9, 487-494.

Rubinstein, H., Lewis, S.S., & Rubinstein, M.A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior,* 10, 646-657.

Rugg, M.D. (1984). Event-related potentials in phonological matching tasks. *Brain & language*. 23: 225-240.

Rugg, M.D. (1987). Dissociation of semantic Priming, word and non-word repetition effects by event-related potentials. *The Quarterly Journal of Experimental Psychology*, 39A, 123-148.

Rugg, M.D., & Barrett, S.E. (1987). Event-related potentials and the interaction between orthographic and phonological information in a rhyme-judgement task. *Brain and language* 32: 336-361.

Rugg, M.D., & Nagy, M.E. (1987). Lexical contribution to non-word repetition effects: evidence from event-related potentials. *Memory and Cognition*, 15, 473-481.

Rumsey, J.M., Horwitz, B., Donohue, B.C., et al. (1997). Phonological and orthographic components of word recognition. A PET-rCBF study. *Brain*, 120: 739-759

Samuel A. G. (1987). Lexical uniqueness effects on phonemic restoration. *Journal of memory & language*, 26: 36-56

Sanquist, T.F., Rohrbaugh, J.W., Syndulko, K., et al. (1980). Electrocortical signs of levels of processing: perceptual analysis and recognition memory. *Psychophysiology*, 17: 568-576.

Segui J., Dupoux E., & Mehler J. (1990). The role of the syllable in speech segmentation, phoneme identification, and lexical access. In G.T.M. Altmann (Ed.), *Cognitive models of speech processing: psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review*, 96, 523-568

Seidenberg, M.S., Waters, G.S., Sanders, M., et al. (1984). Pre and post lexical loci of contextual effects on word recognition. *Memory & Cognition*, 12, 315-328.

Shaywitz, S.E. (1996). Dyslexia. *Scientific American*, 275, 98- 104.

Shaywitz, S.E., Shaywitz, B.A., Pugh, K.R., et al. (1998). Functional Disruption in the Organization of the Brain for Reading in Dyslexia. *Proceedings of the National Academy of Sciences*, 95, 2636-2641.

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

Shen, J.X. (1992). Examples and categorization of slips of the tongue in Chinese. *Chinese Language*, 4: 306-316.

Shen, D., & Forster, K. (1999). Masked phonological priming in reading Chinese words depends on the task. *Language and Cognitive Processes*, 14, 429-460.

Shields, J.L., McHugh, A., & Martin, J.G. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology*, 102, 250-255.

Shipley-Brown, f., Dingwall, W.O., & Berlin, C.I., et al (1988). Hemispheric processing of affective and linguistic intonation contours in normal subjects. *Brain and language*, 33: 16-26.

Shillcock R. (1990). Lexical hypotheses in continuous speech. In G.T.M. Altmann (Ed.), *Cognitive models of speech processing: psycholinguistic and computational perspectives* (pp.24-49). Cambridge, MA: MIT press

Siegel, L.S. (1985). Psycholinguistic aspects of reading disability. In: L.S. Siegel & F.J. Morrison (Eds.), *Cognitive development in atypical children* (pp. 45-65). New York: Academic Press.

Slowiaczek, L.M. (1987). On the role of lexical stress in auditory word recognition. *Unpublished Doctoral dissertation*, Indiana University.

Slowiaczek, L.M. (1990). Effects of lexical stress in auditory word recognition. *Language and Speech*, 33, 47-68.

Slowiaczek, L.M., & Hamburger, M.B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18, 1239-1250.

Slowiaczek, L.M., McQueen, J.M., Soltano, E.G., et al. (2000). Phonological representations in prelexical speech processing: evidence from form-based priming. *Journal of Memory and Language*, 43, 530-560.

Slowiaczek, L.M., & Nusbaum, H.C. (1985). Effects of speech rate and pitch contour on the perception of synthetic speech. *Human Factors*, 27, 701-712.

Slowiaczek L.M., Nusbaum H.C. & Pisoni D.B. (1987). Phonological priming in auditory word recognition. *Journal of experimental psychology: learning, memory, & cognition,* 13: 64-75

Slowiaczek, L.M., & Pisoni, D.B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory & Cognition*, 14, 230-237.

Small, L.H., & Bond, Z.S. (1982). Effects of mispronunciation on lexical access in continuous speech perception. Paper presented at the meeting of the American Speech-Language Hearing Association, Toronto.

Smith, F. (1985). Reading without nonsense (2nd ed.). New York: Teachers College Press.

Smith, E.E., Shoben, E.J., & Rips, L.J. (1974). Structure and process in semantic memory: a feature model for semantic decision. *Psychological Review*, 18, 214-241.

Smith, M.E., & Halgren, E., (1987). Event-related potentials during lexical decision: effects of repetition, word frequency, pronounceability, and concreteness. In R. Johnson Jr., J.W. Rohrbaugh and R. Parasuramn, *Current Trends in Event-related potential research*, (EEG suppl. 40). Elsevier, Amsterdam.

So D., & Siegel, L.S. (1997). Learning to read Chinese: semantic, syntactic, phonological and working memory skills in normally achieving and poor Chinese readers. *Reading and writing*, 9(1): 1-21

Soli, S.D. (1980). Some effects of acoustic attributes of speech on the processing of phonetic feature information. *Journal of experimental psychology: human perception and performance*, 6: 622-638.

Solomon, R.L., & Howes, D.H., (1951). Word frequency, personal values and visual duration thresholds. *Psychological Review*, 58, 256-270.

Sorin, C. (1983). Evaluation of F_0 contribution to speech intelligibility. Paper presented at the *10th International Congress of Phonetic Science*. Utrecht, August.

Speer, S.R., Crowder, R.G., & Thomas, L.M. (1993). Prosodic structure and sentence recognition. *Journal of Memory and Language*, 32, 336-358.

Spinks, J.A., Liu, Y., Perfetti, C.A., et al. (2000). Reading Chinese characters for meaning: the role of phonological information. *Cognition*, 76, B1-B11.

Square-Storer, P., Darley, F.L. & Sommers, R.K. (1988). Nonspeech and speech processing skills in patients with aphasia and apraxia of speech. *Brain and language*, 33: 65-85.

St. George, M., Mannes, S., & Hoffman, J.E. (1994). Global Semantic Expectancy and Language Comprehension. *Journal of Cognitive Neuroscience*, 6, 70-83.

Stanovich, K.E. (1991). Word recognition: changing perspective. In R. Barr, M.L. Kamil, P. Mosenthal & P.D. Pearson (eds), *Handbook of reading research*, Vol 2 (pp. 418-452). New York: Longman

Studdert-Kennedy, M. (1975). Speech perception. In N.J. Lass (Ed.), *Contemporary* *issues in experimental phonetics* (pp. 243-293). New York: Academic Press.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645-659.

Swinney, D. (1981). The process of language comprehension: an approach to examining issues in cognition and language. *Cognition,* 10: 307-312

Tabossi P. (1993). Connections, competitions, and cohorts: comments on the chapters by Marslen-Wilson; Norris: and Bard and Shillcock. In G.T.M. Altmann (Ed.), *Cognitive Models of Speech Processing* (pp. 277-294). Hillsdale, NJ: Erlbaum.

Taft L.A. (1984). Prosodic constraints and lexical parsing strategies. *Unpublished Doctoral dissertation*, University of Massachusetts, Amherst.

Taft M., & Chen. H.C. (1992). Judging homophony in Chinese: the influence of tones. In H.C. Chen & OJL Tzeng (Eds). *Language Processing in Chinese,* 151-172. Amsterdam: Elsevier.

Taft M. & Hambly G. (1986). Exploring the cohort model of spoken word recognition. *Cognition*, 22, 259-282.

Taft M. & Van Graan F. (1998). Lack of phonological mediation in a semantic categorization task. Journal of memory and language, 38, 203-224

Tan, L.H., Hoosain, R. & Peng, D.L (1995). Role of early presemantic phonological code in Chinese character identification, *Journal of experimental psychology: learning, memory, and cognition*, 21, 43-54

Tan, L.H. & Hoosain, R. & Siok, W.W.T. (1996). The activation of phonological codes before access to character meaning in written Chinese. *Journal of experimental psychology: learning, memory, and cognition*, 22, 865-883

Tan L.H. & Perfetti C.A. (1997). Visual Chinese character recognition: does phonological information mediate access to meaning? *Journal of Memory and Language*, 37, 41-57.

Tan L.H. & Perfetti C.A. (1998). Phonological codes as early sources of constraint in Chinese word identification: a review of current discoveries and theoretical accounts. *Reading and Writing: an Interdisciplinary Journal*, 10, 165-200.

Tomiak G.R., Mullennix, J.W. & Sawusch, J.R. (1987). Integral processing of phonemes: Evidence for a phonetic mode of perception, *Journal of the Acoustical Society of America,* 81: 755-764

Tyler L.K., & Wessels J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception & Psychophysics*, 34: 409-420

Tzeng, O.J.L., Hung, D.L., & Wang, W.S.Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 621-630.

Umeda, N. (1977). Consonant duration in American English. *Journal of the acoustical society of America*, 61, 846-858.

Van den Brink, D., Brown, C.M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken word recognition: N200 versus N400 effects. Journal of Cognitive Neuroscience, 13(7) 967-985.

Van Berkum, J., Hagoort, P.,& Brown, C. (1999). Semantic Integration in Sentences and Discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, 11(6), 657-671.

Van Lancker, D., & Fromkin, V.A. (1973). Hemispheric specialization for pitch and tone: Evidence from Thai. *Journal of Phonetics,* 1, 101-109.

Van Orden, G.C. (1987). A ROWS is a ROSE: spelling, sound, and reading. *Memory & cognition*, 15, 181-198

Van Orden, G.C. (1990). Phonological mediations is fundamental to reading. In D. Besner & G. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition*. Hillsdale, NJ: Erlbaum.

Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes,* 8(4).485-531

Van Petten C., Coulson S., Rubin S., et al (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental psychology: Learning, memory, and cognition,* 25 (2):394-417.

Van Petten, C., & Kutas M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, 18(4), 380-393.

Van Petten, C., Kutas, M., Kluender, R., et al. (1991). Fractionating the word repetition effect with event-related potentials, *Journal of Cognitive Neuroscience*, 3, 131-150.

Vroomen, J., & De Gelder, B. (1997). Trochaic rhythm in speech segmentation. Paper presented at the *38th Meeting of the Psychonomic Society*, Philadelphia.

Vroomen, J., Tuomainen, J., & De Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory & Language*, 38, 133-149

Wagner, R.K., & Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101, 192-212.

Waibel, A. (1988). Prosody and speech recognition. London: Pitman.

Wan, I.P. (1997). Tone errors in Mandarin Chinese. *Unpublished Master thesis*, State university of New York, Buffalo.

Wang, J.T. (1997). The interplay of orthography, phonology, semantics and tone in reading Mandarin: an ERP investigation. *Unpublished thesis for Medical Degree in Neurology*, P.R. China.

Wang, J.T., Young, G.B., Connolly, J.F. (2003). Prognostic value of Evoked Responses and Event-Related brain Potentials in coma. To be submitted to Canadian Journal of Neurological Sciences.

West, R.F., & Stanovich, K.E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 385-399.

Wingfield, A., Goodglass, H., & Lindfield, K.C. (1997). Word recognition from acoustic onsets and acoustic offsets: effects of cohort size and syllabic stress. *Applied Psycholinguistics*, 18: 85-100.

Wong, K.F., & Chen, H.C. (1999). Orthographic and phonological processing in reading Chinese text: evidence from eye fixations. *Language and Cognitive Processes*, 14, 461- 480.

Wood, C.C. (1974). Parallel processing of auditory and phonetic information in speech discrimination. *Perception and Psychophysics,* 15: 501-508

Wood, C.C., & Day, R.S. (1975). Failure of selective attention to phonetic segments in consonant-vowel syllables. *Perception & Psychophysics,* 17: 346-350.

Xu, Y. (1991). Depth of phonological recoding in short term memory. *Memory & Cognition*, 19(3), 263-273.

Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of the acoustical society of America,* 95(4): 2240-2253

Xu, Y., Pollatsek, A., & Potter, M. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 838-857

Yang, Y.F. (1991). Ear differences in distinguishing consonant features and lexical tones. *Acta Psychologica Sinica,* 2, 131-137.

Ye Y., & Connine C. M. (1999). Processing spoken Chinese: the role of tone information. *Language and Cognitive Processes,* 14(5/6), 609-630.

Ziegler, J.C., & Jacobs, A.M. (1995). Phonological information provides early sources of constraint in the processing of letter strings. *Journal of memory and Language,* 34, 567-593.

Ziegler, J.C., Van Orden, G.C., & Jacobs, A.M. (1997). Phonology can help or hurt the perception of print. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 845-860.

Zhang, S. & Perfetti, C.A. (1993). The tongue-twister effects in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1082- 1093.

Zhou X. & Marslen-Wilson W. (1999). Phonology, orthography, and semantic activation in reading Chinese. *Journal of memory and language*, 41, 579-606

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken word processing. *Cognition*, 32, 25-64.