Neuropsychologia ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese

Jeffrey G. Malins *, Marc F. Joanisse

The University of Western Ontario, Department of Psychology and Program in Neuroscience, The Brain and Mind Institute, Natural Sciences Centre, London, Ontario, Canada N6A 5B7

ARTICLE INFO

Article history:
Received 12 July 2011
Received in revised form
2 May 2012
Accepted 4 May 2012

Keywords:
Lexical tone
Mandarin Chinese
Spoken word recognition
Event related potentials
Continuous mapping models
Syllable processing

ABSTRACT

We investigated the influences of phonological similarity on the time course of spoken word processing in Mandarin Chinese. Event related potentials were recorded while adult native speakers of Mandarin ($N=19$) judged whether auditory words matched or mismatched visually presented pictures. Mismatching words were of the following nature: segmental (e.g., picture: *hua1* 'flower'; sound: *hua4* 'painting'); cohort (e.g., picture: *hua1* 'flower'; sound: *hui1* 'gray'); rhyme (e.g., picture: *hua1* 'flower'; sound: *gua1* 'melon'); tonal (e.g., picture: *hua1* 'flower'; sound: *jing1* 'whale'); unrelated (e.g., picture: *hua1* 'flower'; sound: *lang2* 'wolf'). Expectancy violations in the segmental condition showed an early-going modulation of components (starting at 250 ms post-stimulus onset), suggesting that listeners used tonal information to constrain word recognition as soon as it became available, just like they did with phonemic information in the cohort condition. However, effects were less persistent and more left-lateralized in the segmental than cohort condition, suggesting dissociable cognitive processes underlie access to tonal versus phonemic information. Cohort versus rhyme mismatches showed distinct patterns of modulation which were very similar to what has been observed in English, suggesting onsets and rimes are weighted similarly across the two languages. Last, we did not observe effects for whole-syllable mismatches above and beyond those for mismatches in individual components, suggesting the syllable does not merit a special status in Mandarin spoken word recognition. These results are discussed with respect to modifications needed for existing models to accommodate the tonal languages spoken by a large proportion of the world's speakers.

1. Introduction

Even though there is an increasing interest in the psycholinguistics literature concerning processing Chinese languages (Li, Tan, Bates, & Tzeng, 2006), it remains the case that the predominant theories and models of speech perception have been developed primarily using data from Indo-European languages such as English and Dutch. These languages are different in nature from tonal languages such as Mandarin Chinese, and as a result, current theories might only account for speech perception in a subset of the world's speakers; namely, speakers of non-tonal languages. To address this problem, a number of studies have begun to focus on how and when lexical tone is processed in the brain, as tone is the key feature that distinguishes tonal languages such as Mandarin Chinese from non-tonal languages such as English (Gandour et al., 2003; Klein, Zatorre, Milner, & Zhao, 2001; Luo et al., 2006; Malins & Joanisse, 2010; Zhao, Guo, Zhou, & Shu, 2011). In addition, the structure of Mandarin syllables is also quite different from the structure of English syllables, and some studies have looked at how these differences in syllabic structure might also lead to differences in spoken word recognition between the two languages (Liu, Shu, & Wei, 2006; Wang & Cheng, 2008; Zhao et al., 2011). While these studies have outlined some of the ways by which theories of spoken word recognition might be updated to include Mandarin, there remain a number of outstanding issues in the field that still need to be resolved before this can be done. The current study uses ERPs to address some of these issues, asking the following three questions: (1) whether the processes underlying tonal versus phonemic access differ; (2) whether onsets and rimes play similar roles in Mandarin word recognition as they do in English; (3) whether, as some have suggested, Mandarin syllables are processed more holistically than their English counterparts.

1.1. Access to tonal information during spoken word processing

Lexical tone refers to variation in the fundamental frequency of a speaker's voice that is used to differentiate phonemically identical words. For instance, Mandarin Chinese includes four lexical tones: high and level (tone 1), mid-rising (tone 2), low-dipping

* Corresponding author. Tel.: +519 661 2111x81408; fax: +519 661 3961.
E-mail addresses: jmalins@uwo.ca (J.G. Malins), marcj@uwo.ca (M.F. Joanisse).

0028-3932/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2012.05.002

(tone 3), and falling (tone 4), and a syllable can carry different meanings depending on the tone that is used. For example, *ma1* means 'mother', yet *ma3* means 'horse' (numerals indicate the tone of the syllable). Thus in order to access the meaning of a spoken word, Mandarin listeners must process not only the phonemes that comprise the word (/m/ and /a/ in the previous example), but the tone of the syllable as well. Tonal information is denoted as suprasegmental, as tone is best described as being associated with an entire syllable rather than an individual phoneme, and its effect can span multiple phonemes; for example, in Mandarin a tonal contour can be realized in the same way over very different vowel clusters such as /iɑu/ versus /uei/. Tonal languages are particularly different from non-tonal languages in that a suprasegmental cue is being pervasively used to distinguish lexical items from each other. This is not to say that non-tonal languages do not make use of suprasegmental information; indeed, Dutch and Japanese are considered non-tonal yet do make considerable use of the suprasegmental features of stress (van Donselaar, Koster, & Cutler, 2005) and pitch accent (Cutler & Otake, 1999). However, tonal languages like Mandarin represent an extreme case of the use of suprasegmental information to constrain word recognition on a syllable-by-syllable basis; tonal information thus plays a key role in this group of languages and must be accounted for in a viable theory of spoken word recognition.

Critically, one of the questions of particular interest from a theoretical standpoint is whether the cognitive processes underlying tonal versus phonemic access are the same or different. A number of early studies using behavioral measures of word recognition such as accuracy and reaction time have suggested that tonal information is accessed later than phonemic information and is a weaker cue for word recognition (Repp & Lin, 1990; Taft & Chen, 1992; Cutler & Chen, 1997; Ye & Connine, 1999—Experiment 1); however, a later behavioral response does not necessarily indicate that the processes underlying tonal access are different from those involved in phonemic access. As discussed below, recent studies using online language measures have more closely investigated these underlying processes, with results suggesting some equivalence among the two, at least in terms of timing.

One type of online language measure that has been used extensively in recent years is the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In a typical study, subjects are presented with pictures of items, and then hear a spoken word that corresponds to one of the pictures. As the auditory word unfolds, eyetracking is used to monitor eye movements to the visual display. Of interest is both the time course of eye movements to the target picture (e.g., a picture of a *cat*), as well as looks to the other pictures that are phonological competitors to this target (e.g., a picture of a *cap*).

We have recently used this methodology to examine auditory word recognition in Mandarin (Malins & Joanisse, 2010). Subjects were presented with arrays of pictures that included target items along with segmental competitors, which shared all phonemes but differed in tone from targets (e.g., *shu3* 'mouse' versus *shu1* 'book'), as well as cohort competitors, which shared word-initial phonemes and tone with targets but differed in word-final vowels (e.g., *qiu2* 'ball' versus *qian2* 'money'). Acoustic analyses suggested the point of divergence of segmental competitors from targets (in tone) closely approximated the point of divergence of cohort competitors from targets (in phonemes), as tone is thought to be carried predominantly on the vowels of syllables (Howie, 1974). Likewise the eye movement data showed that competitor effects were very comparable between these segmental and cohort conditions, providing evidence that tones and phonemes are accessed simultaneously.

However, a lack of difference in processing time or levels of activation (as indexed by looks to items) is not necessarily an indicator that equivalent cognitive processes underlie these effects; indeed, it could be the case that tones versus phonemes are processed in parallel but within slightly different neural circuits. A technique that offers additional insight into these underlying processes is the use of event related potentials (ERPs). Specific components of the ERP waveform have been reliably associated with particular cognitive events, each of which has its own characteristic latency and scalp distribution (Newman, Forbes, & Connolly, 2012). Therefore the use of ERPs is ideally suited to uncovering potential differences in processing that may underlie access to tonal versus phonemic information during spoken word recognition.

Schirmer, Tang, Penney, Gunter, and Chen (2005) used ERPs to investigate tonal versus phonemic processing in Cantonese, a Chinese language that has six tones. They introduced semantic anomalies by altering either the tone or phonemes (specifically, vowels) of the terminal word of sentences. The authors found that tonal and phonemic violations modulated the N400 (an ERP component associated with accessing word-level information such as meaning; Kutas & Hillyard, 1984) at a similar latency and amplitude. They concluded that tonal and phonemic information are accessed in parallel in Cantonese and play comparable roles in constraining activation of phonologically similar words. However, sentence processing involves components beyond those involved in single word identification, and so the employed task may have tapped processes beyond ones that are specific to single word recognition.

Most recently, Zhao et al. (2011) examined recognition of spoken monosyllables in Mandarin by employing a novel picture / spoken-word/ picture task. In this task, Mandarin speakers were presented with a picture of an item in conjunction with its Chinese character before presentation of an auditory word. Following this, subjects were presented with a second picture, and asked whether or not the first and second pictures belonged to the same semantic category. ERPs recorded during passive processing of the auditory word showed that words that differed in tone from the name of the first picture modulated the N400 in the same fashion as words that differed in phonemes (specifically, rimes). Again, the authors reached the conclusion that tonal and phonemic information are accessed over a similar time course and play comparable roles in lexical access. However, Zhao et al. (2011) did not employ a task that required subjects to actively make decisions about the critical stimuli they were hearing, and so their task may not have been optimal for detecting potential differences underlying access to tonal versus phonemic information necessary for lexical processing. In addition, both Zhao et al. (2011) and Schirmer et al. (2005) did not look at any components earlier than the N400, leaving open the question of whether tonal information also influences earlier-going prelexical processes, and whether these might differentiate how and when speakers access segmental versus non-segmental information.

1.2. Syllable processing in Mandarin

In addition to asking whether equivalent processes underlie phonemic versus tonal access, we are also interested in whether the nature of Mandarin syllabic processing is fundamentally different from syllable processing in non-tonal languages such as English. Mandarin syllables are structured differently from English syllables in several ways. Mandarin permits only single onset consonants, whereas English syllables permit consonant clusters in onsets (e.g., /str/ in 'string'). Furthermore, Mandarin syllables do not terminate with consonants (with the exception of the nasals /n/ and /ŋ/), while English syllables are again much more permissive in this respect, allowing many single consonants as well as many clusters in this position. As a result of these

Please cite this article as: Malins, J. G., & Joanisse, M. F. Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in.... *Neuropsychologia* (2012), http://dx.doi.org/10.1016/j.neuropsychologia.2012.05.002