Brain & Language 138 (2014) 38–50

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Developmental differences in the influence of phonological similarity on spoken word processing in Mandarin Chinese

Jeffrey G. Malins [a,b,1], Danqi Gao [c,1], Ran Tao [c], James R. Booth [d], Hua Shu [c,e], Marc F. Joanisse [a], Li Liu [c,e,*], Amy S. Desroches [f,*]

[a] Brain and Mind Institute, The University of Western Ontario, London N6A 5B7, Canada
[b] Haskins Laboratories, New Haven 06511, USA
[c] State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China
[d] Department of Communication Sciences and Disorders, Northwestern University, Evanston 60208, USA
[e] Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing 100875, China
[f] Department of Psychology, University of Winnipeg, Winnipeg R3B 2E9, Canada

ARTICLE INFO

Article history:
Accepted 2 September 2014

Keywords:
Spoken word recognition
Mandarin Chinese
Incremental processing
Top-down processing
Development
Lexical competition
ERPs

ABSTRACT

The developmental trajectory of spoken word recognition has been well established in Indo-European languages, but to date remains poorly characterized in Mandarin Chinese. In this study, typically developing children ($N = 17$; mean age 10; 5) and adults ($N = 17$; mean age 24) performed a picture–word matching task in Mandarin while we recorded ERPs. Mismatches diverged from expectations in different components of the Mandarin syllable; namely, word-initial phonemes, word-final phonemes, and tone. By comparing responses to different mismatch types, we uncovered evidence suggesting that both children and adults process words incrementally. However, we also observed key developmental differences in how subjects treated onset and rime mismatches. This was taken as evidence for a stronger influence of top-down processing on spoken word recognition in adults compared to children. This work therefore offers an important developmental component to theories of Mandarin spoken word recognition.

1. Introduction

Even though Mandarin Chinese is widely spoken, it remains relatively understudied compared to Indo-European languages such as English and Dutch. In light of this, psycholinguistic researchers have become increasingly interested in developing theories of language processing in Mandarin and other tonal languages (Li, Tan, Bates, & Tzeng, 2006). A key area of investigation concerns spoken word recognition; that is, how listeners map incoming acoustic information onto word knowledge in the brain in order to uncover the meaning of spoken forms. In recent years, a number of studies have examined spoken word recognition in Mandarin; however, they have tended to focus on adults (Lee, 2007; Malins & Joanisse, 2010, 2012; Shen, Deutsch, & Rayner, 2013; Ye & Connine, 1999; Zhao, Guo, Zhou, & Shu, 2011). It is our view that characterizing spoken word recognition in children might offer additional insights into these theories by establishing developmental constraints.

When studying the development of spoken word recognition, an important consideration is that over time, child readers change in their sensitivity to different types of phonological relationships (Anthony, Lonigan, Driscoll, Phillips, & Burgess, 2003; Ziegler & Goswami, 2005). This sensitivity is related to the concept of phonological awareness, or knowledge of the sounds that make up spoken words (McBride-Chang, 1996; Wagner & Torgesen, 1987). Phonological awareness can refer to phonemic awareness, or awareness of the individual consonants and vowels in words; alternatively, it can also refer to awareness of larger units such as rhymes (Bryant, MacLean, Bradley, & Crossland, 1990; Treiman, 1986). In a tonal language like Mandarin Chinese, there is an additional component to awareness of the sound structure of words: tone awareness. Tone can be defined as fluctuation in the pitch of a speaker's voice that is used in a lexically contrastive sense. For example, the syllable *tang* can mean 'soup' when pronounced in a high and level pitch (tone 1), while it can mean 'candy' when pronounced in a pitch that begins in the mid-range

* Corresponding authors at: State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China (L. Liu), Department of Psychology, University of Winnipeg, Winnipeg R3B 2E9, Canada (A.S. Desroches).

E-mail addresses: lilyliu.bnu@gmail.com (L. Liu), a.desroches@uwinnipeg.ca (A.S. Desroches).

[1] Both authors contributed equally to all aspects of this study, and should be considered co-first authors.

http://dx.doi.org/10.1016/j.bandl.2014.09.002
0093-934X/

of a speaker's register and rises throughout articulation of the syllable (tone 2). Tone awareness then refers to one's ability to recognize, identify, or manipulate the tones of a language.

Phonological awareness is typically measured using overt behavioral tasks such as the oddity task (Burnham et al., 2011; Chen et al., 2004). In this task, subjects are presented with triplets of syllables, two of which share a critical feature, and are asked to select the odd one out. Using these types of measures, the developmental progression of phonological awareness for the different components of the Mandarin syllable has been fairly well characterized (Chen et al., 2004; Shu, Peng, & McBride-Chang, 2008). These studies have provided evidence that awareness of larger units, such as syllables and tones, precedes awareness of smaller units such as onsets and rimes (Shu et al., 2008). This developmental progression from larger to smaller units has been similarly well documented in English and other Indo-European languages (see Ziegler & Goswami, 2005 for a review). However, a limitation of these types of measures is that they index metalinguistic knowledge of spoken words but fail to capture potential differences between individuals in the cognitive processing leading up to a behavioral response. For example, using eyetracking, Desroches, Joanisse, and Robertson (2006) showed that typically developing versus reading impaired English-speaking children differed in online processing of rhyming words, even though they failed to show a difference in accuracy on a rhyme judgment task.

These types of results indicate that online measures of spoken word recognition offer information that is not always apparent in behavioral measures. For this reason, we employed an online processing task in order to study the development of Mandarin spoken word recognition; namely, a task that required subjects to resolve different types of phonological competition. In this type of task, competition is incurred by setting up an expectation for a spoken word, which spreads activation to a set of lexical candidates that share a phonological relationship with the spoken word. As auditory input unfolds, these candidate items compete for recognition. The manner in which listeners resolve this competition is related to their sensitivity to different types of phonological relationships, among other factors (McMurray, Samelson, Lee, & Bruce Tomblin, 2010).

In the current study we examined how children versus adults resolved phonological competition based on shared onsets, rimes, or tones. Examining responses to these different types of phonological competition allowed us to draw conclusions regarding two key issues associated with the development of spoken word recognition in Mandarin. The first concerns whether spoken word recognition in children is incremental or holistic. The second concerns whether there are developmental differences in how children versus adults weight bottom-up and top-down processing.

Regarding the first issue, the distinction between incremental and holistic processing can be thought of in terms of the dominant theories of spoken word recognition. One critical way in which these theories differ from one another is in how they treat phonological similarity amongst spoken words. In theories like TRACE (McClelland & Elman, 1986) and Cohort (Marslen-Wilson, 1987), the temporal structure of a word affects the dynamics of lexical competition. Therefore, these theories predict different levels of competition amongst words sharing word-initial information (e.g., words belonging to the same 'cohort', such as *cat–cab*) versus words sharing word-final information (e.g., rhyming words such as *cat–hat*). A key reason for this is that Cohort and TRACE include explicit mismatch inhibition, so words diverging in onset from an expected word are suppressed on the basis of disconfirming auditory input (as discussed in Magnuson, Dixon, Tanenhaus, & Aslin, 2007). Cohort and TRACE differ in that Cohort predicts that onset-based competition is the only type of competition amongst spoken words, while TRACE allows for rhyme effects but predicts they are much weaker in nature than onset-based competition. Nevertheless, both models are similar in that they emphasize onset-based competition, and differences between the acoustic signal and word-level representations are mapped continuously as the spoken word unfolds. This type of processing can be considered 'incremental' (Magnuson, Tanenhaus, Aslin, & Dahan, 2003).[2] In contrast, the neighborhood activation model (NAM; Luce & Pisoni, 1998) instead emphasizes global similarity. In NAM, a competitor is defined as differing from a spoken word in one phoneme in any position, whether initial, medial, or final. Importantly, unlike Cohort and TRACE, NAM does not include mismatch inhibition amongst lexical competitors. This lack of mismatch inhibition means that NAM does not make differential predictions regarding early versus late effects of similarity. Therefore, like TRACE, it predicts that rhyming forms are part of the competitor set of a spoken word; however, they are considered to give rise to competitive effects that are equivalent to those of onset-based competitors (e.g., *cat* and *cab* overlap in two phonemes in the same way that *cat* and *hat* do, and are predicted to give rise to the same degree of lexical competition). This type of processing, in which a type of global similarity amongst spoken words is computed without taking temporal information into account, has sometimes been termed 'holistic' (Zhao et al., 2011).

We have previously reported evidence that spoken word recognition in Mandarin-speaking adults is incremental in nature (Malins & Joanisse, 2012). However, this has yet to be tested in Mandarin-speaking children. Given that Mandarin-speaking children appear to be differentially sensitive to onsets, rimes, and tones behaviorally (Shu et al., 2008), we predict this should also be reflected in online processing of spoken word forms. Therefore, we hypothesize that both children and adults should show evidence of incremental processing, and thus observed data should be more consistent with Cohort and TRACE as opposed to NAM in this regard.

The second theoretical issue relates to whether children and adults differ with respect to feedback connections that influence online spoken word recognition in Mandarin. Again, it is helpful to think of this issue in terms of the dominant theories of spoken word recognition. In feedforward models such as Cohort, spoken word recognition proceeds entirely from acoustic–phonetic information toward meaning, and can therefore be considered strictly bottom-up. In contrast, feedback models such as TRACE incorporate top-down connections from the lexical to phoneme layer (McClelland & Elman, 1986). In these types of theories, perception is still driven primarily by bottom-up or feedforward connections; top-down connections are thought to play a weaker role mostly confined to helping refine perception in challenging situations for the listener, such as noisy environments. Because top-down connections are confined to these types of situations, they can be thought of as highly experience-dependent; that is, top-down connections strengthen over time as an individual acquires word-specific knowledge (Davis & Johnsrude, 2007). Therefore, an important question arises: how do children and adults differ in the influence of top-down processing? Some insight into this question is highly useful from a theoretical perspective, as it helps add a developmental component to theories of spoken word recognition in Mandarin, which to date have mostly considered the adult state of processing.

One way to test for a developmental difference in top-down processing is to assess how children versus adults treat rhyming words. In models such as TRACE that allow for top-down

[2] Note that the term 'incremental' is used here to refer to spoken word processing that is continuous and dynamic. Thus this term should not be considered entirely synonymous with 'serial' or 'sequential'; rather, we would apply this term to both serial and parallel processing models so long as the temporal structure of a word is taken into account when calculating phonological similarity.