Brain and Mind Institute Researchers' Publications   Brain and Mind Institute

11-1-2014

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

Jeffrey G Malins
Brain and Mind Institute, The University of Western Ontario, London N6A 5B7, Canada; Haskins Laboratories, New Haven 06511, USA

Danqi Gao
State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China

Ran Tao
State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China

James R Booth
Department of Communication Sciences and Disorders, Northwestern University, Evanston 60208, USA

Hua Shu
State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China; Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing 100875, China

See next page for additional authors
Follow this and additional works at: https://ir.lib.uwo.ca/brainpub

Part of the Neurosciences Commons, and the Psychology Commons

Authors

Jeffrey G Malins, Danqi Gao, Ran Tao, James R Booth, Hua Shu, Marc F Joanisse, Li Liu, and Amy S
Desroches

Developmental Differences in the Influence of Phonological Similarity on Spoken Word Processing in Mandarin Chinese

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

[a]Brain and Mind Institute, The University of Western Ontario, London, Canada, N6A 5B7

[b]Haskins Laboratories, New Haven, USA, 06511

[c]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China, 100875

[d]Department of Communication Sciences and Disorders, Northwestern University, Evanston, USA, 60208

[e]Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing, China, 100875

[f]Department of Psychology, University of Winnipeg, Winnipeg, Canada, R3B 2E9

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.

Keywords

spoken word recognition; Mandarin Chinese; incremental processing; top-down processing; development; lexical competition; ERPs

[*]Correspondence should be addressed to either: Amy S. Desroches, Telephone: 1-204-258-2922, a.desroches@uwinnipeg.ca or Li Liu, Telephone: 86-10-58802913, lilyliu.bnu@gmail.com.
[†]Both authors contributed equally to all aspects of this study, and should be considered co-first authors.

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 (Ye & Connine, 1999; Lee, 2007; Malins & Joanisse, 2010; Zhao, Guo, Zhou, & Shu, 2011; Malins & Joanisse, 2012; Shen, Deutsch, & Rayner, 2013). 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 (Wagner & Torgesen, 1987; McBride-Chang, 1996). 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 (Treiman, 1986; Bryant, MacLean, Bradley, & Crossland, 1990). 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 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, Robertson, and Joanisse (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, & 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)[1]. 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

---

[1]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.

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 towards 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 connections, expecting to hear a certain word form leads to partial activation of the phonemes that comprise it, including word-final phonemes. This results in a key prediction: facilitated processing of rhyming versus non-rhyming forms. The idea here is because word-final phonemes are partially active, when they finally arrive in the acoustic signal they are processed more easily than word-final phonemes in non-rhyming forms, which have not received this prior activation. Evidence for facilitated processing of rhyming forms has been

observed in numerous studies of Indo-European languages in both adults (Praamstra, Meyer, & Levelt, 1994; Radeau, Besson, Fontenau, & Castro, 1998; Desroches, Newman, & Joanisse, 2009) and children (Coch, Grossi, Coffey-Corina, Holcomb, & Neville, 2002; Desroches, Newman, Robertson, & Joanisse, 2013; Malins et al., 2013). In contrast, rhyme effects in Mandarin speakers have been elusive. With respect to adults, rhyme effects have been observed for disyllabic words (Liu, Shu, & Wei, 2006), but there has been mixed evidence in monosyllables (Malins & Joanisse, 2010; Malins & Joanisse, 2012; Zhao et al, 2011). In children, however, no experiment has yet recorded online measures of spoken word recognition to test for rhyme effects during rapid processing. Therefore it remains unclear whether there are developmental differences in these types of effects. Nevertheless, we hypothesize that rhyme effects will be stronger in adults compared to children, as recent work using functional magnetic resonance imaging (fMRI) has shown that top-down modulation is weaker in children compared to adults during tasks of spoken word recognition (Bitan et al., 2006; Bitan, Cheon, Lu, Burman, & Booth, 2009). It could be the case that this effect is especially prominent in Mandarin (Cao et al., 2011).

In the present study, we analyzed event related potentials (ERPs), an online processing measure which has not yet been used to assess spoken word recognition in Mandarin-speaking children. ERPs are advantageous as they provide a continuous measure of speech perception as spoken words unfold, and furthermore certain components of the ERP waveform have been reliably associated with discrete cognitive processes (Newman, Connolly, & Forbes, 2012). We employed a picture-word matching paradigm that we have previously used to investigate spoken word recognition in adult native Mandarin speakers (Malins & Joanisse, 2012). In this task, subjects are presented with pictures of items, which set up expectations of subsequent auditory input (Connolly, Byrne, & Dywan, 1995; Desroches et al., 2009). After viewing a picture for a specified interval, subjects hear an auditory word that either matches or mismatches expectations, and their task is to indicate whether or not the auditory word matches the picture. These mismatching forms share some phonological relationship with the expected word, such as onset or rime, but differ in other components, such as tone. In the current study, we included the following mismatch types: onset mismatches (see: *tang2* 'candy', hear: *lang2* 'wolf'); rime mismatches (see: *tang2* 'candy', hear: *tao2* 'peach'); tone mismatches (see: *tang2* 'candy', hear: *tang1* 'soup'); segmental mismatches, which differed in all phonemes but not tone from expected words (see: *tang2* 'candy', hear: *niu2* 'cow'); total mismatches, which differed in all phonemes and tone from expected words (see: *tang2* 'candy', hear: *xia1* 'shrimp')[2].

We were interested in how these mismatches would modulate two components thought to index distinct stages of the word recognition process: the phonological mapping negativity (PMN; Connolly & Phillips, 1994; Connolly et al., 1995; Newman & Connolly, 2009), thought to index pre-lexical phoneme mapping, and early and late portions of the N400

---

[2]Note that these condition names are different from those we have used in previous papers, such as Malins and Joanisse (2010; 2012). Specifically, the labels for tone and segmental mismatches have been inverted; additionally, 'cohort' mismatches are called 'rime' mismatches in the present study, and 'rhyme' mismatches are now called 'onset' mismatches. Therefore, when we discuss 'rhyming forms' and 'rhyme effects' we are in fact referring to the onset mismatch condition. We apologize for any confusion this causes for readers; however, we hope the current condition labels are more intuitive and also allow this study to make greater contact with others in the field.

(Kutas & Hillyard, 1984; Connolly & Phillips, 1994; Desroches et al., 2009), thought to index whole-word recognition. Importantly, it has been shown that these two components are functionally dissociable in terms of the underlying cortical regions that give rise to these ERP responses (Trébuchon, Démonet, Chauvel, & Liégeois-Chauvel, 2013). Furthermore, previous research has indicated that analysis of these components reveals temporally dissociable effects for different mismatch types, such as onset versus rime competitors (Desroches et al., 2009; Archibald & Joanisse, 2011; Malins & Joanisse, 2012; Desroches et al., 2013; Malins et al., 2013). This approach is therefore well-suited to the current investigation, as we aimed to identify which specific aspects of spoken word recognition differed between children and adults.

To summarize, we were interested in whether Mandarin-speaking children process spoken words incrementally, as well as the extent of the influence of bottom-up and top-down processing during word recognition in children versus adults. We hypothesized that both children and adults process spoken words incrementally, and therefore predicted that they would show a similar pattern of responses to words differing from expectations in either onset, rime, or tone. Namely, words diverging from expectations in word-initial information (i.e., onset mismatches) were expected to modulate the PMN response to a greater extent than words overlapping expectations in word-initial information (i.e., rime and tone mismatches). In addition, words differing from expectations in onsets versus rimes versus tones were expected to lead to differential modulation of the N400 and late N400. Next, as we hypothesized that top-down processing is weighted more heavily in adults compared to children, we predicted that the adults should show greater evidence for facilitated processing of rhyming forms. We expected this to result in a greater attenuation of the late N400 for onset mismatches in the adults compared to the children.

2. Methods

2.1. Subjects

Testing took place at the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University in Beijing, China. Overall, 29 children and 21 adults participated. Children were recruited from local schools in Beijing, and adults were all undergraduate students recruited from Beijing Normal University. Twelve children and four adults were removed due to either excessive EEG artifacts or equipment failure, for a total of 17 children and 17 adults included in the experiment (children: 12 female, mean age 10;5, see Table 1 for full details; adults: 11 female, mean age 24). As it has been shown that there is a relationship between phonology and reading ability in Chinese children (Cheung et al., 2009; Liu, Shu, & Yang, 2009), we wished to screen out atypical readers to make sure that our results were not contaminated by differences in reading ability across children. To do this, we screened the children using a test of reading ability as well as a test of phoneme awareness; the latter was included because deficits in phoneme awareness have been implicated in reading impairment in Chinese (Ho, Law, & Ng, 2000; Newman, Tardif, Huang, & Shu, 2011). All children were identified as typically developing readers based on scoring at or above the expected average for their grade level on these two tests. The reading test was a single character reading task taken from Lei et al. (2011), which required children

to read aloud 150 Chinese characters, all of which are expected to be learned by the sixth grade in Beijing (Shu, Chen, Anderson, Wu, & Xuan, 2003). The characters were arranged from the simplest to the most challenging based on complexity and the grade level at which a character is first learned. The phoneme awareness test was a phoneme deletion task, taken from Li et al. (2012), in which children were asked to produce a new syllable after the deletion of a target phoneme in either the initial, medial, or final position. There were 28 trials in this task. In addition, cognitive ability was assessed using the Chinese version of the Wechsler Intelligence Scale for Children (WISC-R) (Wechsler, 1974), which verified that all children had typical or above average IQ (IQ > 90). Scoring procedures were based on local norms (Lin & Zhang, 1986).

2.2. Stimuli and Procedures

All stimuli were monosyllabic Mandarin words referring to common, highly imageable objects. Sixteen sets of items (as listed in the Supplementary Appendix) were created by changing single Mandarin words in ways that created phonologically related competitor words; for example, for the critical word *tang2* 'candy', we derived the following mismatching forms: an onset mismatch: *lang2* 'wolf'; a rime mismatch: *tao2* 'peach'; a tone mismatch: *tang1* 'soup'; a segmental mismatch, different in all segments but not tone: *niu2* 'cow'; and a total mismatch, differing in all segments and tone: *xia1* 'shrimp'. By deriving stimuli from these critical words, we were able to control as much as possible attendant psycholinguistic factors influencing single-word recognition, as each item participated equally in all conditions. In addition, as much as possible, mismatching items were selected from the other sets in the experiment in order to have a closed set of stimuli, and care was taken to ensure that critical words and competitors were not closely related to one another semantically. Furthermore, items in the five mismatch conditions did not differ from one another in logarithmic frequency [$F(4,60) = .686$, $p = .58$, $\eta_p^2 = .04$].

Auditory stimuli were single-syllable words spoken in isolation by an adult male speaker of Mandarin, digitized and recorded to disk at 16 bits quantization, with a 44,100 Hz sampling rate, and volume normalized to −10 dB of maximum amplitude. Each item was spoken seven times, and the middle five tokens were selected for use in the experiment. Items in the five mismatch conditions did not differ from one another in duration [$F(4,60) = 1.51$, $p = .22$, $\eta_p^2 = .09$]. Culturally appropriate pictures depicting these items were selected with the assistance of native Mandarin speakers. To verify that subjects were familiar with the names of these pictures, we asked subjects at the beginning of the experimental session to name aloud each of the pictures by saying the first Mandarin syllable/character that came to mind upon viewing the picture. The naming task allowed us to ensure they were comfortable with the names we had selected for each picture and associated them with monosyllabic words rather than potential disyllabic alternatives. Subjects showed high levels of accuracy on this task [adults: 89.5% (SE = .77%); children: 83.8% (SE = 1.42%)], indicating that the selected names for the pictures were familiar and appropriate for the subjects. This was further supported by high levels of accuracy for the subsequent picture-word matching task in both groups. In cases where the given name differed from the intended name, subjects were told the intended name, as in prior picture naming studies (Brooks & MacWhinney, 2000).

In the experimental trials, subjects performed a picture-word matching task while we recorded continuous EEG activity from the surface of their scalp. In each trial, subjects viewed a fixation cross for 250 ms, following which a picture of an item appeared on 21-inch CRT monitor for 1500 ms. Next, while the item remained on screen, subjects heard an auditory word presented binaurally via Sennheiser HD201 headphones. Following responses, a blank screen was presented for 1000 ms prior to the onset of the next trial.

Subjects were asked to indicate via button press whether the auditory word matched or mismatched the picture present on the screen. Across all trials in the experiment, the match to mismatch ratio was one to one, and the five mismatch trial types (onset, rime, tone, segmental, and total) were presented with equal probability. Subjects completed 320 trials (160 match trials; 32 trials for each of the five mismatch types). For all mismatch trials, the roles of pictures and words were counterbalanced (e.g., for the pair *tang2-tang1*, in one trial *tang2* was the picture and *tang1* the sound, and in another trial *tang1* was the picture and *tang2* the sound). Thus, we ensured that every time a picture or sound item was presented, it was equally likely to be a match or a mismatch trial. Testing was divided into four blocks of trials with short rests between blocks. Block order was counterbalanced across subjects, which allowed us to control for whether each item was a match or mismatch the first time it was encountered. Prior to the experimental trials, subjects performed a training block of six trials containing items not presented in the actual experiment.

All materials and procedures were approved by the Institutional Review Board at Beijing Normal University. When testing children, each child's caregivers gave written consent to participate in the study, and each family received a small gift as well as a performance report after attending the experimental session.

2.3. EEG Data Acquisition

EEG data were collected at a sampling rate of 1000 Hz using Acquire 4.2 (Neurosoft Inc., El Paso, TX) and 32-channel caps with sintered Ag/AgCl electrodes (Quik-Caps; Neurosoft Inc., El Paso, TX) oriented according to the international 10–20 system and referenced to the nose tip. Across subjects, the impedance of each channel was kept below 5 kΩ.

Data were amplified at a gain of 500 using a SynAmps amplifier and filtered online using 60 Hz notch and .1 – 100 Hz bandpass filters. ERPs were then segmented into epochs spanning from 200 ms pre-stimulus to 750 ms post-stimulus onset, time-locked to the onset of the auditory stimulus. Following this, data were filtered offline using a 24 dB zero phase shift digital bandpass filter (0.1 – 30 Hz), and baseline corrected to the mean voltage of the pre-stimulus interval. Trials containing blinks and other artifacts were removed using a maximum voltage criterion of ±100 μV for the portion of the waveform subjected to statistical analysis (0 - 600 ms) on the following electrodes: Fz, F3, F4, F7, F8, Cz, C3, C4, T7, T8, Pz, P3, P4, P7, P8. In addition, incorrect trials were rejected. Following artifact rejection, the average number of accepted trials per condition for the children was 121/160 for the match trials (minimum 103), 24/32 for onset mismatches (minimum 17), 23/32 for rime mismatches (minimum 17), 23/32 for tone mismatches (minimum 15), 25/32 for segmental mismatches (minimum 16), and 24/32 for the total mismatch condition (minimum 16), while for adults the average number of accepted trials per condition was 149/160 for the

match trials (minimum 133), 30/32 for onset mismatches (minimum 27), 29/32 for rime mismatches (minimum 26), 29/32 for tone mismatches (minimum 24), 30/32 for segmental mismatches (minimum 26), and 29/32 for the total mismatch condition (minimum 26). During data recording, there were several bad channels; namely, F7 for two of the children, T8 for two of the children, and C3 for one of the children and eight of the adults. Values for these channels were calculated using the mean of the surrounding electrodes (for F7: F3 and FT7; for T8: FT8, C4, and TP8; for C3: T7, FC3, Cz, and CP3).

Because the ratio of accepted trials differed between groups [$t(32) = -6.41$, $p < .001$, Cohen's $d = 2.20$], we calculated signal to noise ratios for the fifteen electrodes subjected to data analysis. This was done for each electrode by dividing the root mean square of the voltage for the period following the auditory stimulus (0 – 600 ms) by the root mean square of the pre-stimulus baseline period (−200 – 0 ms). Mean SNR was then calculated per subject across these fifteen electrodes. Comparison between groups showed that mean SNRs were not different between the children and the adults [$t(32) = .252$, $p = .80$, Cohen's $d = .09$]. These results indicate that any observed differences between adults and children were not likely due to differences in the signal to noise ratio of our ERP recordings.

2.4. Analysis of ERP Data

Because our aim was to compare the adults and children, all statistical analyses were performed using difference waves. Our rationale was that performing between-groups statistics on relative voltage differences between the word type conditions rather than on absolute voltages helps offset potential differences in the absolute magnitude of ERP waveforms. These difference waves were generated by subtracting each group's match condition from respective mismatch conditions.

For both groups, we analyzed the following ERP components: the PMN, the N400, and the late N400, as in prior picture-word matching studies (Desroches et al., 2009; Archibald & Joanisse, 2011; Malins & Joanisse, 2012; Desroches et al., 2013; Malins et al., 2013). Windows of analysis were determined by visual inspection, and were defined as follows: PMN (260-320 ms), N400 (350-500 ms), and late N400 (500-600 ms). We calculated the magnitude of each ERP response for each respective mismatch condition as the difference in mean voltage compared to match over the specified time window. For each analysis window in each group, we selected fifteen electrodes (Fz, F3, F4, F7, F8, Cz, C3, C4, T7, T8, Pz, P3, P4, P7, P8), which provided full scalp coverage in a way that would allow us to differentiate the components of interest (Newman, Connolly, Service, & McIvor, 2003). We then analyzed mean amplitude values using a 'column' approach to ERP analysis similar to that described in Holcomb and Grainger (2006). That is, we conducted separate repeated measures analyses of variance for lateral, medial, and midline electrodes. Midline electrodes were further divided into anterior (Fz), central (Cz), and posterior (Pz) regions, yielding a 3-way mixed design ANOVA with two within-subjects factors (3 region, 5 mismatch type), and a between-subjects factor of group (2; adults and children). Medial electrodes were also divided into anterior (F3/F4), central (C3/C4), and posterior (P3/P4) regions, and there was an additional within-subjects factor of hemisphere (left or right). Similarly, for lateral electrodes, we divided electrodes according to region (anterior: F7/F8; central: T7/T8;

posterior P7/P8), and hemisphere. Thus, lateral and medial electrodes were analyzed with a four-way mixed design ANOVA (within-subjects factors: hemisphere, region, and word type; between-subjects factor: group). All analyses of variance were conducted using conservative degrees of freedom (Greenhouse & Geisser, 1959).

For each component, we focused on interactions between word type and group; if any of these were significant, we performed step-down repeated measures ANOVAs separately for each group. If there were no significant interactions between word type and group, we then focused on interactions between word type and the within-subjects factors of region and hemisphere. If none of these were significant, we then focused on main effects of word type.

Hypothesis 1: Both Children and Adults Process Spoken Words Incrementally—To test this hypothesis, we followed up on significant interactions and main effects in the PMN, N400, and late N400 windows with Bonferroni corrected pairwise comparisons between the following mismatch types: onset, rime, and tone. The reason for this is that these three conditions differed as to which component of the syllable diverged from expectations. Furthermore, because these mismatches are signaled at different times during the unfolding of the syllable, we predicted that we would see differences among these conditions in the PMN, N400, and late N400 windows, which can be considered evidence of incremental processing. As we hypothesized that both children and adults process spoken words incrementally, we predicted both groups would show a similar pattern of mismatch effects for these three conditions.

Hypothesis 2: Top-Down Processing is Weighted More Heavily in Adults Compared to Children—Because we hypothesized that top-down processing is weighted more heavily in adults compared to children, we predicted that we would observe a developmental difference in terms of how subjects treated rhyming words. Namely, we expected to observe a greater attenuation of the late N400 for onset mismatches (that is, words differing in onset from expectations but sharing word-final phonemes with expectations; e.g., *tang2-lang2*) in adults compared to children (Praamstra et al., 1994; Radeau et al., 1998; Coch et al., 2002; Desroches et al., 2006; Desroches et al., 2009; Desroches et al., 2013; Malins et al., 2013). For this reason, we followed up on significant interactions in the late N400 window with post-hoc analyses that contrasted the onset mismatch condition between groups. Additionally, to rule out the possibility that late N400 amplitude was simply attenuated across the board for all conditions in the children versus adults, we tested for the specificity of this effect by performing similar analyses for the other four mismatch conditions.

3. Results

3.1. Behavioral Data

Table 2 presents behavioral performance for the picture-word matching task for both the adults and children. Trials were removed from behavioral analysis if decision latencies were over 2.5 standard deviations above or below condition-wise means, or were less than 100 ms; a total of 3.01% of all trials for the children, and 2.70% for the adults, were removed. Furthermore, decision latencies were only analyzed for correct trials. For decision latencies,

a mixed repeated measures ANOVA with word type as a within-subjects factor (6) and group as a between-subjects factor (2), showed that there was a significant interaction between group and word type [$F(5,170) = 3.052$, $p = .02$, $\eta_p^2 = .08$], as well as a main effect of group [$F(1,34) = 31.388$, $p = .001$, $\eta_p^2 = .48$], as decision latencies were generally longer for the children. To follow up on this interaction, we performed pairwise comparisons between each respective mismatch condition and the match condition within each group (presented in Table 2). For accuracy, there was a main effect of word type [$F(5,170) = 4.116$; $p = .01$; $\eta_p^2 = .11$] as well as a main effect of group [$F(1,34) = 15.743$; $p = .001$; $\eta_p^2 = .32$], as accuracy was slightly lower in the children (although still well over 90% in all word type conditions); however, there was no significant interaction between word type and group [$F(5,170) = .664$; $p = .59$; $\eta_p^2 = .02$]. The values in Table 2 clearly show that both the adults and the children performed well on the task.

3.2. ERP Data

Difference waveforms for the adults and children for the different mismatch conditions are shown in Figure 1, while condition-wise waveforms are plotted in Supplementary Figures S1 though S6. Supplementary Table S1 presents the ANOVAs for effects of age group, word type, and electrode position for each component.

3.2.1. Hypothesis 1: Both children and adults process spoken words incrementally—To test this hypothesis, we examined responses to the different mismatch conditions in all three of the PMN, N400, and N400 windows. For the PMN, as is shown in Table S1, analyses did not reveal an interaction between group and word type in any column. This suggests the relationship between the magnitude of the PMN response and mismatch type was comparable across the adults and children. However, when the two groups were collapsed, there was a significant interaction between region and word type in all three columns [midline: $F(8,256) = 4.195$, $p = .002$, $\eta_p^2 = .12$; medial: $F(8,256) = 2.804$, $p = .03$, $\eta_p^2 = .08$; lateral: $F(8,256) = 2.552$, $p = .05$, $\eta_p^2 = .07$], as well as an interaction between hemisphere and word type in the lateral column [$F(4,128) = 2.793$, $p = .04$, $\eta_p^2 = .08$]. Bonferroni corrected pairwise comparisons between the onset, rime, and tone mismatch conditions (presented in Table 3) revealed that the PMN response was larger in the onset mismatch condition than in the rime and tone mismatch conditions.

For the N400, there was no significant interaction between word type and group in the midline and medial columns, nor interactions between word type and region or hemisphere. However, there was a main effect of word type in both columns [midline: $F(4,128) = 3.180$, $p = .03$, $\eta_p^2 = .09$; medial: $F(4,128) = 4.286$, $p = .01$, $\eta_p^2 = .12$]. Bonferroni corrected pairwise comparisons (Table 3) revealed that the N400 was larger in amplitude for rime and tone mismatches compared to onset mismatches.

The spatial extent of this effect was slightly larger in the adults, as there was a three-way interaction between region, word type, and group in the lateral column [$F(8,256) = 2.368$, $p = .05$, $\eta_p^2 = .07$]. To follow up on this, we performed separate repeated measures ANOVAs for each group, which revealed a main effect of word type in the adults [$F(4,64) = 9.041$, $p = .001$, $\eta_p^2 = .36$], but no significant interaction between word type and region. In contrast,

the children showed neither a significant interaction between word type and region nor a main effect of word type. Pairwise comparisons (Table 3) revealed that the adults showed a larger N400 for rime and tone mismatches compared to onset mismatches, the same pattern observed for both the children and the adults in the midline and medial columns.

Last, for the late N400, there were significant three-way interactions between region, word type, and group in the midline and medial columns [midline: $F(8,256) = 2.632$, $p = .03$, $\eta_p^2 = .08$; medial column: $F(8,256) = 2.510$, $p = .04$, $\eta_p^2 = .07$]. Separate repeated measures ANOVAs within each group revealed an interaction between region and word type in the children [midline: $F(8,128) = 4.828$, $p = .002$, $\eta_p^2 = .23$; medial column: $F(8,128) = 3.499$, $p = .01$, $\eta_p^2 = .18$], and a main effect of word type in the adults [midline: $F(4,64) = 2.710$, $p = .05$, $\eta_p^2 = .15$; medial column: $F(4,64) = 3.357$, $p = .02$, $\eta_p^2 = .17$]. For both groups, Bonferroni corrected pairwise comparisons (Table 3) revealed larger late N400 amplitudes for the rime mismatch condition compared to the onset mismatch condition, and the rime mismatch condition compared to the tone mismatch condition. In the lateral column, there was a main effect of word type [$F(8,128) = 3.589$, $p = .01$, $\eta_p^2 = .10$], and similar same condition-wise effects were observed as in the midline and medial columns.

Overall, these results suggest that the children and adults showed similar patterns of mismatch effects: first, a larger PMN response for onset mismatches compared to rime and tone mismatches; second, larger N400 responses for rime and tone mismatches compared to onset mismatches; third, larger late N400 responses for rime mismatches compared to onset and tone mismatches. These effects are apparent in Figure 2.

3.2.2. Hypothesis 2: Top-down processing is weighted more heavily in adults compared to children—As there were significant three-way interactions between group, region, and word type in the late N400 window, we focused on the onset mismatch condition and tested for group by region interactions in the midline and medial columns. For the medial column, we collapsed across hemisphere; this analysis revealed a significant group by region interaction for onset mismatches [$F(2,64) = 5.638$, $p = .006$, $\eta_p^2 = .15$]. We explored the nature of this interaction by performing Bonferroni corrected pairwise t-tests between adults and children in the three electrode regions. These post-hoc analyses revealed that in medial frontal sites F3 and F4, late N400 amplitude for onset mismatches was reduced in the adults compared to the children [$t(32) = 2.787$, $p = .03$, Cohen's $d = .96$]. This effect is illustrated in Figure 3A.

To test for the specificity of this effect, we also performed similar analyses for the other four mismatch conditions. For tone mismatches, analyses did not reveal a significant interaction between region and group nor a main effect of group on late N400 amplitude. However, there was a main effect of group for rime mismatches in the medial column [$F(1,32) = 5.984$, $p = .02$, $\eta_p^2 = .16$]; as illustrated in Figure 3B, late N400 amplitudes were larger in the children compared to the adults in most medial sites. Additionally, there was a significant interaction between group and region for the segmental and total mismatch conditions in both the midline [segmental: $F(2,64) = 5.936$, $p = .01$, $\eta_p^2 = .16$; total: $F(2,64) = 4.923$, $p = .02$, $\eta_p^2 = .13$] and medial columns [segmental: $F(2,64) = 5.357$, $p = .01$, $\eta_p^2 = .14$; total: $F(2,64) = 4.624$, $p = .03$, $\eta_p^2 = .13$]. However, Bonferroni corrected post-hoc t-

tests did not reveal differences between adults and children in late N400 amplitude in any region for either of these two word type conditions ($p > .15$ for all tests).

4. Discussion

We investigated whether Mandarin-speaking children process spoken words incrementally, and whether top-down processing influences spoken word recognition in children to the same extent that it does in adults. To satisfy this aim, we compared typically developing children and adults on a task that required subjects to determine whether or not auditory words matched visual pictures (Connolly et al., 1995; Desroches et al., 2009). Our rationale is that when viewing a particular object, a subject generates a name for it and activates its phonological form. Viewing a picture therefore sets up a strong expectation for subsequent auditory input; however, during this process competitor words are also partially activated. Next, when auditory information arrives, a subject maps auditory input onto the previously activated phonological form to perform the matching judgment. This task requires the subject to use incoming acoustic information to disambiguate targets from competitors. Critically, we designed stimuli such that we expected different levels of phonological and/or lexical competition based on the nature of the relationship between expected words and mismatches in the different word type conditions. As a result, our analyses focused on two components thought to be associated with phonological and lexical processing during auditory word recognition; namely, the phonological mapping negativity (PMN) response, which is associated with processing word-initial phonemes (Connolly & Phillips, 1994; Newman & Connolly, 2009), and early and late portions of the N400, a component associated with processing at the word level (Kutas & Hillyard, 1984).

4.1. Hypothesis 1: Both Children and Adults Process Spoken Words Incrementally

The first hypothesis concerned whether or not children process spoken words incrementally. By incremental processing, we are referring to competition amongst spoken forms that takes the temporal structure of words into account (Magnuson et al., 2003). This type of processing can be considered the opposite of holistic processing, which computes phonological similarity amongst spoken words in a global fashion (Luce & Pisoni, 1998). We compared responses to mismatches that diverged from expectations at different time points during the unfolding of a spoken word. Specifically, we compared responses to the onset, rime, and tone mismatch conditions in the children and adults. Onset mismatches (e.g., *tang2-lang2*) were signaled very early during the unfolding of spoken forms, as word-initial phonemes differed from expectations. Conversely, word-initial phonemes in the rime (e.g., *tang2-tao2*) and tone (e.g., *tang2-tang1*) mismatch conditions matched expectations, and therefore differences from expectations were not signaled until later on in the spoken form. In the case of the rime mismatch condition, these were different vowels, whereas in the case of the tone mismatch condition, this was a different tone, which is carried on the vowel cluster (Howie, 1974). We expected that if subjects process words incrementally, the PMN response, which is sensitive to pre-lexical phoneme mapping, should be more negative in the onset mismatch condition compared to the rime and tone mismatch conditions (Connolly & Phillips, 1994; Newman & Connolly, 2009; Desroches et al., 2009; Archibald

& Joanisse, 2011; Desroches et al., 2013). As can be seen in Figure 2, this is exactly what was observed in both children and adults.

Evidence for incremental processing was also observed for both adults and children in the N400 and late N400 windows. First, the rime and tone mismatch conditions both showed larger N400 responses compared to the onset mismatch condition. We think that the size of the N400 and late N400 response is associated with the amount of indecision and/or interference a listener experiences when recognizing an auditory word, which can be considered a reflection of the extent of competition experienced from other lexical items (O'Rourke & Holcomb, 2002). A larger N400 in the rime and tone mismatch conditions thus suggests that individuals experienced more competition in this window, likely because this was when the mismatch was first signaled in these conditions. In contrast, divergence from expectation was signaled earlier in the onset mismatch condition, which led to the suppression of representations of competitor words that were incompatible with the spoken input. Additionally, it is likely that the N400 response for the rime and tone mismatch conditions was conflated with a PMN response in this window, and so the apparent amplitude of the N400 appeared larger as a result. However, we hesitate to make any precise claims about this, as the point of mismatch between expected words and rime and tone mismatches was variable across stimulus sets, so PMN latency was expected to be variable in both of these conditions.

Second, the rime mismatch condition showed a larger late N400 amplitude compared to onset and tone mismatches for both the adults and the children. A large late N400 amplitude in the rime mismatch condition has been observed previously in this task in both Mandarin (Malins & Joanisse, 2012) and English (Desroches et al., 2009); it is thought to reflect a delay in the N400 response incurred as a result of word-initial overlap with an expected word form (Connolly & Phillips, 1994). Interestingly though, both rime and tone mismatches overlapped expected word forms in word-initial information, yet these two conditions differed from one another in this window. It is thus possible that listeners resolved competition on the basis of tone at a slightly earlier point in time that they resolved competition on the basis of word-final vowels. This view is supported by a previous gating study in which subjects showed superior identification of tones compared to word-final vowels following presentation of only the initial segments of Mandarin monosyllables (Wu & Shu, 2003).

As can be seen in Supplementary Figures S1 and S4, in which the tone mismatch condition showed a larger negativity compared to match in the PMN and N400 windows, tonal competition was resolved relatively early in both groups. We would therefore argue that both children and adults were able to efficiently use tonal information to constrain word recognition as soon as it became available. This conclusion is further motivated by previous work, including our own, highlighting the pivotal role of tone in Mandarin spoken word recognition (Brown-Schmidt & Canseco-Gonzalez, 2004; Malins & Joanisse, 2010; Zhao et al., 2011; Malins & Joanisse, 2012). The current study adds to the existing literature by suggesting that tonal processing systems are sufficiently developed in ten-year old children such that they resolve tonal competition over a similar time course as adults. This assertion

fits with earlier results showing that tone awareness in Mandarin-speaking children increases dramatically following literacy instruction (Shu, Peng, McBride-Chang, 2008).

In summary then, as is shown in Figure 2, there was a complete dissociation in responses to onset, rime, and tone mismatches in terms of which components of the waveform were modulated. Importantly, this pattern held for both the adults and the children, and can be taken as evidence of incremental processing of spoken words in both groups. The current study thus replicates prior work we have done that has offered evidence for incremental processing in English-speaking adults (Desroches et al., 2009), English-speaking children (Desroches et al., 2013; Malins et al., 2013), and Mandarin-speaking adults (Malins & Joanisse, 2012), but this is the first study to establish these effects in Mandarin-speaking children.

4.2. Hypothesis 2: Top-Down Processing is Weighted More Heavily in Adults Compared to Children

The second hypothesis concerned the relative influence of top-down processing on spoken word recognition in children compared to adults. As discussed previously, top-down connections are thought to be highly-experience dependent, and therefore become stronger over time as an individual acquires word-specific knowledge. For this reason, we hypothesized that adults would show evidence of a greater influence of top-down processing on spoken word recognition in Mandarin; we expected this to manifest as a developmental difference in how children versus adults treat rhyming words. The logic is as follows: theories that allow for top-down feedback from the lexical to phonological layer predict that the phonemes that comprise a word are partially activated when expecting to hear a given word form, which can result in facilitated processing of rhyming forms (onset mismatches in the present study). The explanation for this effect is that because word-final phonemes in rhyming words receive some partial activation, they are processed more easily than word-final phonemes in phonologically unrelated word forms, which do not benefit from this prior activation. In ERP studies, this prior activation results in an attenuation of the (late) N400 component for rhyming compared to non-rhyming forms, reflecting the lesser extent of processing that is required to resolve mismatches that overlap in word-final information. This attenuation has been observed in ERP studies of Indo-European languages in both adults (Praamstra et al., 1994; Radeau et al., 1998; Desroches et al., 2009) and children (Coch et al., 2002; Desroches et al., 2006; Desroches et al., 2013; Malins et al., 2013), offering support for the viability of models such as TRACE that allow for top-down feedback.

As is shown in Figure 3, the late N400 was attenuated in Mandarin-speaking adults to a greater extent than it was in Mandarin-speaking children. Importantly, this group-wise difference in processing onset mismatches suggests that top-down processing does not influence spoken word recognition in children to the same extent that it does in adults. This finding complements earlier work using fMRI showing that top-down modulation from prefrontal areas of the brain is weaker in children compared to adults during spoken word recognition tasks (Bitan et al., 2006; Bitan et al., 2009; Cao et al., 2011). However, there is another explanation for this difference between groups. It could be that the children

activated onset mismatches to the same extent as adults, yet had more trouble suppressing these on the basis of bottom-up acoustic input. A consequence of this is that these word forms were still active in the late N400 window for the children, and so they were still experiencing competition; meanwhile, the adults had suppressed these forms and instead only experienced facilitation in processing word-final phonemes. This is supported by the observation that children and adults also differed in late N400 amplitude for rime mismatches, suggesting differences in bottom-up processing between the two groups. Importantly though, this difference in bottom-up processing was not observed across the board for all mismatch conditions; indeed, adults and children did not differ in late N400 amplitude for the tone mismatch condition, even though word-initial overlap in this condition likely incurred strong bottom-up competition. Furthermore, late N400 amplitude did not differ between adults and children for either the segmental or total mismatch conditions. We would therefore argue that while the effect of onset competitors could be explained solely by differences between children and adults in bottom-up processing, it is our view that it is instead the result of a developmental difference in the weighting of top-down information.

It should be noted that the location at which we observed an attenuation of the late N400 for onset mismatches was actually different than that reported in studies such as Liu et al. (2006), in which a rhyme effect in Mandarin was observed predominantly in centroparietal sites. The reason we observed a difference between groups in frontal sites in the present study is that the spatial extent of the late N400 for onset mismatches was confined to only centroparietal electrodes in the adults, while it was widespread across the scalp in the children, including frontal sites. Therefore, when we performed a direct subtraction between groups, frontal sites were identified as the location where the group-wise difference was most prominent.

It is also worth noting that the lack of difference between adults and children for segmental mismatches (*tang2-niu2*) is interesting from a theoretical perspective, given that developmental differences could have been expected in this condition due to mismatches sharing a degree of word-final overlap with expectations. In the segmental mismatch condition, the tonal contour matched that of expected word forms, and it is conceivable that there could have been facilitated processing of these words based on prior activation of tonal units. However, we observed very little evidence for this given that ERP responses in the segmental mismatch condition were almost identical to those in the total mismatch condition in both children and adults (Figure 1A and 1C). We propose that tonal overlap alone is too weak a manipulation to drive this facilitatory effect; indeed, we did not observe it in our prior study in which we performed a direct statistical comparison between segmental and total mismatches (Malins & Joanisse, 2012). In contrast, mismatches in the onset mismatch condition overlapped expectations in both tone and word-final phonemes, and it is possible that the conjunction of these two components was responsible for the facilitatory effect. Since none of the word type conditions in the current study overlapped expectations in only word-final phonemes and not tone, we cannot dissociate these two types of information; this could be an avenue of investigation for future studies.

Finally, it is worth considering that behaviorally, children much younger than those tested in the current study show ceiling performance on tests of rhyme awareness (Siok & Fletcher, 2001; Shu et al., 2008). Therefore the present data suggest that children ten years of age differ from adults in how they treat rhyming words during rapid processing, even though they do not differ from adults in overt sensitivity to rhyming relationships. This view is consistent with the behavioral data reported in Table 2, as we observed similar patterns of effects for onset mismatches between the adults and children, even though ERP responses to these words were different between groups. This finding highlights the utility of ERP investigations in uncovering potential differences in spoken word recognition between groups of subjects that are not always apparent in behavioral measures (Spivey, 2007).

4.3. Implications for Theories of Spoken Word Recognition

These findings have important implications for theories and models of spoken word recognition in Mandarin. First, these results are most consistent with theories such as Cohort and TRACE that emphasize incremental processing, and are least consistent with theories such as NAM that emphasize global similarity. Furthermore, they suggest that children as young as ten years of age show dissociable patterns of responses to violations in onsets versus rimes versus tones, suggesting they are sensitive to each of these components of the Mandarin syllable. Second, the developmental difference in rhyme effects emphasizes the importance of adding a learning component to theories of spoken word recognition in Mandarin. Critically, this finding supports continuous mapping models such as TRACE that allow for top-down feedback from the lexical to phonological layer, but expands upon these theories by suggesting that these top-down connections strengthen as a result of experience, and are not fully matured in children ten years of age.

More generally, these results also serve to highlight several similarities in spoken language processing that hold across different language groups. Namely, the current data suggest that principles such as incremental processing and an influence of top-down information on spoken word recognition might apply more broadly across human languages than some might propose. With this in mind, it is our view that it is vitally important to study a variety of languages when developing theories of spoken word recognition. These investigations have the potential to not only point out how disparate languages differ, but also to illuminate what they share in common, thereby enabling theories to be more universal in nature.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Anthony JL, Lonigan CJ, Driscoll K, Phillips BM, Burgess SR. Phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations. Reading Research Quarterly. 2003; 38(4):470–487.

Archibald LM, Joanisse MF. Electrophysiological responses to coarticulatory and word level miscues. Journal of Experimental Psychology: Human Perception and Performance. 2011; 37(4):1275–1291. [PubMed: 21574742]

Bitan T, Burman DD, Lu D, Cone NE, Gitelman DR, Mesulam M, Booth JR. Weaker top–down modulation from the left inferior frontal gyrus in children. NeuroImage. 2006; 33(3):991–998. [PubMed: 16978881]

Bitan T, Cheon J, Lu D, Burman DD, Booth JR. Developmental increase in top–down and bottom–up processing in a phonological task: An effective connectivity, fMRI Study. Journal of Cognitive Neuroscience. 2009; 21(6):1135–1145. [PubMed: 18702576]

Brooks PJ, MacWhinney B. Phonological priming in children's picture naming. Journal of Child Language. 2000; 27(2):335–366. [PubMed: 10967891]

Brown-Schmidt S, Canseco-Gonzalez E. Who do you love, your mother or your horse? An event-related brain potential analysis of tone processing in Mandarin Chinese. Journal of Psycholinguistic Research. 2004; 33(2):103–135. [PubMed: 15098511]

Bryant PE, MacLean M, Bradley LL, Crossland J. Rhyme and alliteration, phoneme detection, and learning to read. Developmental Psychology. 1990; 26(3):429.

Burnham D, Kim J, Davis C, Ciocca V, Schoknecht C, Kasisopa B, Luksaneeyanawin S. Are tones phones? Journal of Experimental Child Psychology. 2011; 108(4):693–712. [PubMed: 21087775]

Cao F, Khalid K, Lee R, Brennan C, Yang Y, Li K, Bolger DJ, Booth JR. Development of brain networks involved in spoken word recognition of Mandarin Chinese. NeuroImage. 2011; 57(3):750–759. [PubMed: 20884355]

Chen X, Anderson RC, Li W, Hao M, Wu X, Shu H. Phonological awareness of bilingual and monolingual Chinese children. Journal of Educational Psychology. 2004; 96(1):142.

Cheung H, Chung KK, Wong SW, McBride-Chang C, Penney TB, Ho CS. Perception of tone and aspiration contrasts in Chinese children with dyslexia. Journal of Child Psychology and Psychiatry. 2009; 50(6):726–733. [PubMed: 19175808]

Coch D, Grossi G, Coffey-Corina S, Holcomb PJ, Neville HJ. A developmental investigation of ERP auditory rhyming effects. Developmental Science. 2002; 5(4):467–489.

Connolly JF, Byrne JM, Dywan CA. Assessing adult receptive vocabulary with event-related potentials: an investigation of cross-modal and cross-form priming. Journal of Clinical and Experimental Neuropsychology. 1995; 17(4):548–565. [PubMed: 7593475]

Connolly JF, Phillips NA. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. Journal of Cognitive Neuroscience. 1994; 6(3):256–266. [PubMed: 23964975]

Davis MH, Johnsrude IS. Hearing speech sounds: top-down influences on the interface between audition and speech perception. Hearing Research. 2007; 229:132–147. [PubMed: 17317056]

Desroches AS, Joanisse MF, Robertson EK. Specific phonological impairments in dyslexia revealed by eyetracking. Cognition. 2006; 100(3):B32–B42. [PubMed: 16288732]

Desroches AS, Newman RL, Joanisse MF. Investigating the time course of spoken word recognition: Electrophysiological evidence for the influences of phonological similarity. Journal of Cognitive Neuroscience. 2009; 21(10):1893–1906. [PubMed: 18855555]

Desroches AS, Newman RL, Robertson EK, Joanisse MF. Electrophysiological indices of phonological impairments in dyslexia. Journal of Speech, Language and Hearing Research. 2013; 56(1):250–264.

Greenhouse SW, Geisser S. On methods in the analysis of profile data. Psychometrika. 1959; 24:95–112.

Ho CSH, Law TPS, Ng PM. The phonological deficit hypothesis in Chinese developmental dyslexia. Reading and Writing. 2000; 13(1-2):57–79.

Holcomb PJ, Grainger J. On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. Journal of Cognitive Neuroscience. 2006; 18(10): 1631–1643. [PubMed: 17014368]

Howie JM. On the domain of tone in Mandarin. Phonetica. 1974; 30(3):129–148.

Kutas M, Hillyard SA. Brain potentials during reading reflect word expectancy and semantic association. Nature. 1984; 307(5947):161–163. [PubMed: 6690995]

Lee CY. Does horse activate mother? Processing lexical tone in form priming. Language and Speech. 2007; 50(1):101–123. [PubMed: 17518105]

Lei L, Pan J, Liu H, McBride-Chang C, Li H, Zhang Y, Shu H. Developmental trajectories of reading development and impairment from ages 3 to 8 years in Chinese children. Journal of Child Psychology and Psychiatry. 2011; 52(2):212–220. [PubMed: 20854364]

Li H, Shu H, McBride-Chang C, Liu H, Peng H. Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills. Journal of Research in Reading. 2012; 35(3):287–307.

Li, P.; Tan, LH.; Bates, E.; Tzeng, OJL., editors. The Handbook of East Asian psycholinguistics volume 1: Chinese. Cambridge University Press; New York, USA: 2006.

Lin, CD.; Zhang, HC. The Chinese revision of WISC-R [in Chinese]. Beijing Teachers College Press; Beijing, China: 1986.

Liu W, Shu H, Yang Y. Speech perception deficits by Chinese children with phonological dyslexia. Journal of Experimental Child Psychology. 2009; 103(3):338–354. [PubMed: 19380151]

Liu Y, Shu H, Wei J. Spoken word recognition in context: Evidence from Chinese ERP analyses. Brain and Language. 2006; 96(1):37–48. [PubMed: 16194567]

Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. Ear and Hearing. 1998; 19(1):1–36. [PubMed: 9504270]

Magnuson JS, Dixon JA, Tanenhaus MK, Aslin RN. The dynamics of lexical competition during spoken word recognition. Cognitive Science. 2007; 31(1):133–156. [PubMed: 21635290]

Magnuson JS, Tanenhaus MK, Aslin RN, Dahan D. The time course of spoken word learning and recognition: studies with artificial lexicons. Journal of Experimental Psychology: General. 2003; 132(2):202–227. [PubMed: 12825637]

Malins JG, Desroches AS, Robertson EK, Newman RL, Archibald L, Joanisse MF. ERPs reveal the temporal dynamics of auditory word recognition in Specific Language Impairment. Developmental Cognitive Neuroscience. 2013; 5:134–148. [PubMed: 23523986]

Malins JG, Joanisse MF. Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese. Neuropsychologia. 2012; 50:2032–2043. [PubMed: 22595659]

Malins JG, Joanisse MF. The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. Journal of Memory and Language. 2010; 62(4):407–420.

Marslen-Wilson WD. Functional parallelism in spoken word-recognition. Cognition. 1987; 25(1):71–102. [PubMed: 3581730]

McBride-Chang C. Models of speech perception and phonological processing in reading. Child Development. 1996; 67(4):1836–1856. [PubMed: 8890511]

McClelland JL, Elman JL. The TRACE model of speech perception. Cognitive Psychology. 1986; 18(1):1–86. [PubMed: 3753912]

McMurray B, Samelson VM, Lee SH, Bruce Tomblin J. Individual differences in online spoken word recognition: Implications for SLI. Cognitive Psychology. 2010; 60(1):1–39. [PubMed: 19836014]

Modern Chinese Frequency Dictionary. Beijing Language Institute Publisher; Beijing: 1986. in Chinese

Newman EH, Tardif T, Huang J, Shu H. Phonemes matter: The role of phoneme-level awareness in emergent Chinese readers. Journal of Experimental Child Psychology. 2011; 108(2):242–259. [PubMed: 20980019]

Newman RL, Connolly JF. Electrophysiological markers of pre-lexical speech processing: Evidence for bottom–up and top–down effects on spoken word recognition. Biological Psychology. 2009; 80(1):114–121. [PubMed: 18524453]

Newman RL, Connolly JF, Service E, McIvor K. Influence of phonological expectations during a phoneme deletion task: Evidence from event-related brain potentials. Psychophysiology. 2003; 40(4):640–647. [PubMed: 14570171]

Newman, RL.; Forbes, K.; Connolly, JF. Event-related potentials and magnetic fields associated with spoken word recognition. In: Spivey, M.; Joanisse, M.; McRae, K., editors. Cambridge handbook of psycholinguistics. Cambridge University Press; New York, NY: 2012.

O'Rourke TB, Holcomb PJ. Electrophysiological evidence for the efficiency of spoken word recognition. Biological Psychology. 2002; 60(2):121–150. [PubMed: 12270588]

Praamstra P, Meyer AS, Levelt WJ. Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. Journal of Cognitive Neuroscience. 1994; 6(3):204–219. [PubMed: 23964972]

Radeau M, Besson M, Fonteneau E, Castro SL. Semantic, repetition and rime priming between spoken words: Behavioural and electrophysiological evidence. Biological Psychology. 1998; 48:183–204. [PubMed: 9700017]

Shen J, Deutsch D, Rayner K. On-line perception of Mandarin tones 2 and 3: Evidence from eye movements. The Journal of the Acoustical Society of America. 2013; 133:3016–3029. [PubMed: 23654405]

Shu H, Chen X, Anderson RC, Wu N, Xuan Y. Properties of school Chinese: Implications for learning to read. Child Development. 2003; 74(1):27–47. [PubMed: 12625434]

Shu H, Peng H, McBride-Chang C. Phonological awareness in young Chinese children. Developmental Science. 2008; 11(1):171–181. [PubMed: 18171377]

Siok WT, Fletcher P. The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. Developmental Psychology. 2001; 37(6):886–899. [PubMed: 11699761]

Spivey, M. The continuity of mind. Oxford University Press; New York: 2007.

Trébuchon A, Démonet JF, Chauvel P, Liégeois-Chauvel C. Ventral and dorsal pathways of speech perception: An intracerebral ERP study. Brain and Language. 2013; 127(2):273–283. [PubMed: 24028995]

Treiman R. The division between onsets and rimes in English syllables. Journal of Memory and Language. 1986; 25(4):476–491.

Wagner RK, Torgesen JK. The nature of phonological processing and its causal role in the acquisition of reading skills. Psychological Bulletin. 1987; 101(2):192–212.

Wechsler, D. Wechsler intelligence scale for children-revised. The Psychological Corporation; San Antonio, TX: 1974.

Wu N, Shu H. The gating paradigm and spoken word recognition of Chinese [in Chinese]. Acta Psychologica Sinica. 2003; 35(5):582–590.

Ye Y, Connine CM. Processing spoken Chinese: The role of tone information. Language and Cognitive Processes. 1999; 14(5-6):609–630.

Zhao J, Guo J, Zhou F, Shu H. Time course of Chinese monosyllabic spoken word recognition: Evidence from ERP analyses. Neuropsychologia. 2011; 49(7):1761–1770. [PubMed: 21382389]

Ziegler JC, Goswami U. Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. Psychological Bulletin. 2005; 131(1):3–29. [PubMed: 15631549]

Highlights

- We examined the development of spoken word recognition in Mandarin using ERPs

- Children and adults resolved different types of phonological competition

- Responses to mismatches in onset, rime, and tone were dissociable in both groups

- The adults showed greater evidence for facilitated processing of rhyming words

- The influence of top-down information was stronger in adults compared to children
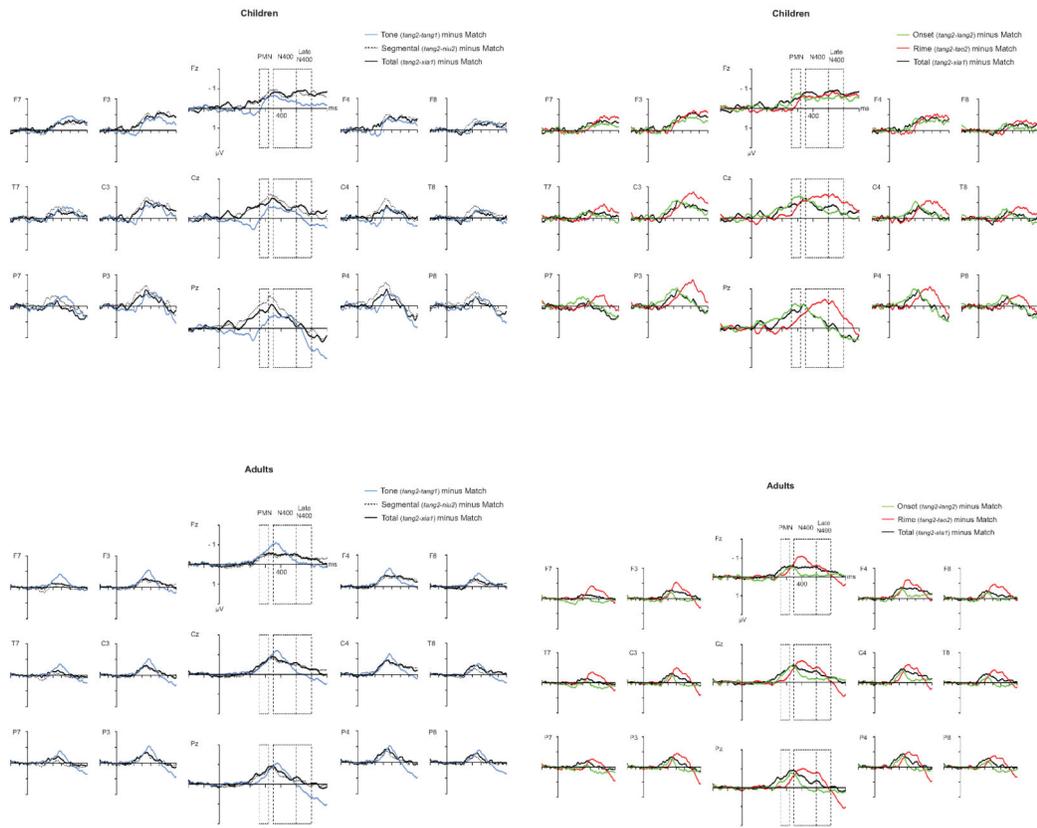
Figure 1.
Difference waveforms for the five mismatch conditions for the children (A, B) and adults (C, D). Waveforms were generated by subtracting the match condition from each respective mismatch condition in each group of subjects. The boxes delineate the PMN, N400, and late N400 windows that were assessed via statistical analysis.
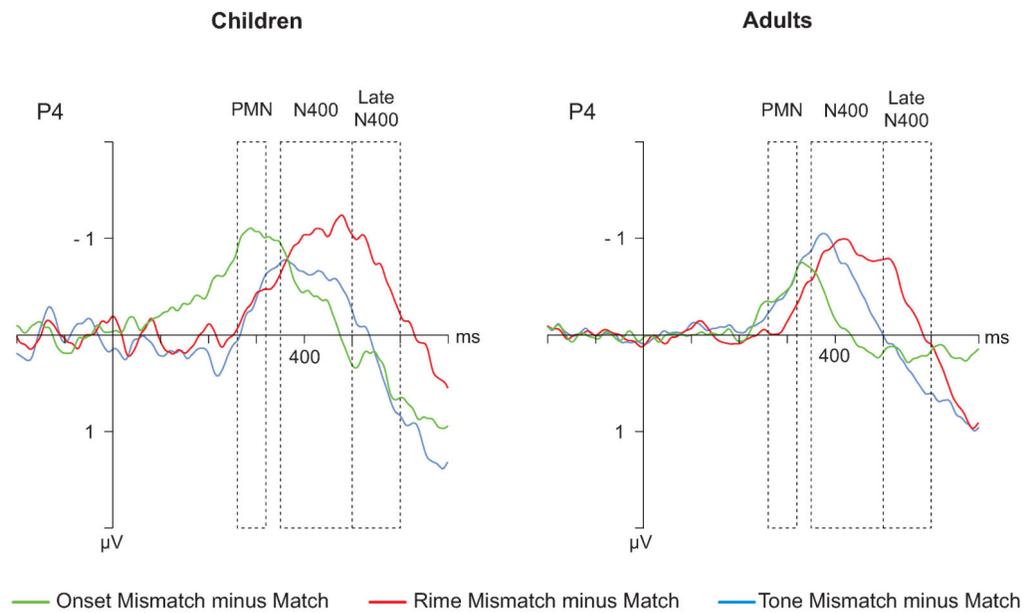
Figure 2.
Responses to onset, rime, and tone mismatches in the children and adults. Across these three conditions, there was a dissociation in terms of which components of the waveform were modulated compared to the match condition. A similar pattern of effects was observed between the children and adults.

Onset Mismatches minus Match for Children vs. Adults

Rime Mismatches minus Match for Children vs. Adults
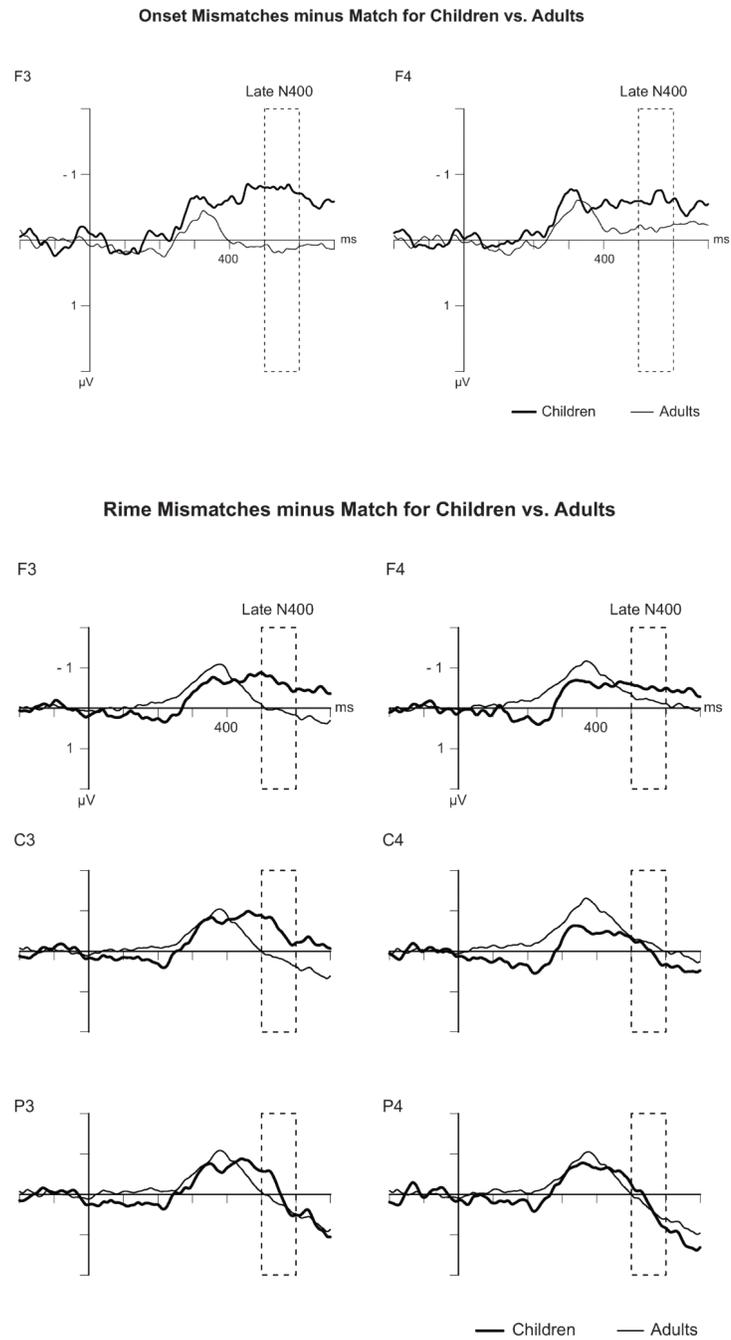
Figure 3.
Responses to onset (A) and rime (B) mismatches in the children and adults. The late N400 was attenuated in the adults compared to the children in both word type conditions.

Table 1

Descriptive Statistics for the Group of Typically Developing Children (*N* = 17; 12 Female)

Measure	Mean (SD)
Age (months)	124.8 (13.7)
Verbal IQ[a]	114.5 (13.1)
Performance IQ[a]	114.9 (11.3)
Phoneme deletion[b]	22.7 (3.3)
Single character reading[b]	118.0 (14.1)

IQ = Intelligence Quotient

[a]Standard scores.

[b]Raw score out of a maximum of 28 for phoneme deletion and 150 for single character reading.

Table 2

Mean Decision Latency and Mean Percent Accuracy for the Matching Judgment Relative to Word Onset

Word Type Condition	Adults (N = 17)		Children (N = 17)	
	Decision Latency (ms)	Percent Accuracy	Decision Latency (ms)	Percent Accuracy
Match	744.1 (24.4)	99.3 (0.11)	996.3 (30.2)	94.5 (0.90)
Onset mismatch	792.4 (22.0)*	99.2 (0.34)	1041.0 (33.1)*	96.2 (1.01)
Rime mismatch	850.8 (23.0)*	96.8 (0.95)*	1137.2 (41.3)*	93.8 (1.61)
Tone mismatch	835.0 (22.3)*	97.3 (0.84)*	1097.1 (45.7)*	94.5 (1.31)
Segmental mismatch	769.5 (19.0)	99.8 (0.19)*	1055.5 (43.4)*	96.4 (0.89)
Total mismatch	761.3 (19.5)	99.3 (0.43)	1058.2 (42.7)*	97.0 (0.79)*

Note. Values in parentheses represent standard errors. Asterisks represent pairwise comparisons between the match condition and respective mismatch conditions that were significant at $p < .05$.

Table 3

Results of Pairwise Comparisons Relevant to Hypothesis 1

Component	Comparison	Group	Location	df	t	Cohen's d
PMN	Onset > Tone	Both	Pz	33	−3.11[*]	.53
			P3 and P4	33	−2.63[**]	.45
	Onset > Rime	Both	Cz	33	−3.06[*]	.52
			C3 and C4	33	−3.30[*]	.57
			Pz	33	−3.23[*]	.55
			P3 and P4	33	−3.50[*]	.60
N400	Rime > Onset	Both	midline	33	−4.06[***]	.70
			medial column	33	−5.19[***]	.89
		Adults	lateral column	16	−6.27[***]	1.52
	Tone > Onset	Both	medial column	33	−3.20[**]	.55
		Adults	lateral column	16	−5.29[***]	1.28
Late N400	Rime > Onset	Children	Pz	16	−3.43[*]	.83
			P3 and P4	16	−4.37[**]	1.06
		Adults	medial column	16	−3.24[*]	.79
		Both	lateral column	33	−3.76[**]	.64
	Rime > Tone	Children	Pz	16	−3.51[*]	.85
		Adults	medial column	16	−2.65[*]	.64

[*]
$p < .05$ (two-tailed; Bonferroni corrected)

[**]
$p < .01$

[***]
$p < .001$