Syllabic tone articulation influences the identification and use of words during Chinese sentence reading: Evidence from ERP and eye movement recordings

Yingyi Luo[1,2] · Ming Yan[3] · Shaorong Yan[1] · Xiaolin Zhou[1] · Albrecht W. Inhoff[4]

Published online: 27 August 2015
Psychonomic Society, Inc. 2015

Abstract In two experiments, we examined the contribution of articulation-specific features to visual word recognition

Lukatela, Eaton, Sabadini, & Turvey, 2004; Wheat, Cornelissen, Frost, & Hansen, 2010). Furthermore, recordings

to Yan et al. (2014), when a target word is processed while being fixated, benefits of neutral-tone usage should be obtained for early ERP components, and costs might be found for later components. To further investigate neutral-tone usage not only before but also after a target word is identified during sentence reading, eye movements were recorded while sentences with neutral- and full-tone target words were read in Experiment 2. Readers of Standard Chinese were expected to spend less time viewing neutral- than full-tone words during first-pass reading, but this should not occur during target rereading if neutral-tone use during the later stage was relatively difficult. In addition, different types of spoken distractors were presented when the target words were viewed. Earlier work (Eiter & Inhoff, 2010; Inhoff, Connine, Eiter, Radach, & Heller, 2004) indicaword en,

lab at Minzu University of China while reading aloud 114 simple sentences in Standard Chinese. Each sentence contained a target word in the middle position (e.g., the word - . in the following example), with a frame, as in:

Ta /	shuo /	Target word [shi-huan] /	zhe-ge /	ci.
he /	say /	Target word [order around] /	this-CLASSIFIER /	word

'He said the word [order around].'

These readers sat before a computer monitor on which the test sentences were displayed using the custom-written recording tool AudiRec. A Shure 58 Microphone was placed about 10–15 cm in front of them. The sampling rate was 48 kHz, and the sampling format was one-channel 16-bit linear.

The duration and intensity of the target words were measured to establish their reliability in capturing the tone neutralization (Y. Wang, 2004). ProsodyPro, a Praat script (Xu, 2013), was used to perform the initial acoustic analysis in Praat (Boersma & Weenink, 2005). On the basis of the waveform and spectrogram of each sentence, segmentation labels were marked manually to identify the boundaries of the target syllable. The duration and intensity measurements for marked segments (i.e., target syllables) were then automatically extracted. The results showed that the second syllable had a shorter articulation duration for neutral-tone than for full-

ratings yielded identical numeric values of 1.4 (< 1) for sentences with full- and reduced-tone syllables, indicating that the sentences were relatively easy to read and that articulatory variation did not matter. The next sentence was presented 1, 000 ms after a sentence was rated. The sentences of each list were divided into two equal-sized blocks, and a rest period was offered in between these blocks. Five warm-up sentences were presented at the beginning of each block.

The electroencephalogram (C4(li)17(st)],

Trials contaminated by excessive movement artifacts (mean voltages exceeding ±70 V) were excluded before trials were averaged over the items of a particular condition. On average, 72 % of trials were accepted for the statistical analysis (41 trials for the neutral- and full-tone syllable targets). Loss of

less-negative amplitudes for neutral- than for full-tone words cannot be attributed to carryover from the prior context.

N250 Means for the N250 component are shown as a function of tone type and topographic location in Fig. 6. As can be seen in the depiction of the full waveforms in Fig. 4, the N250 component occurred between a trough at around 200 ms and a spike at around 400 ms, which is consistent with the N250

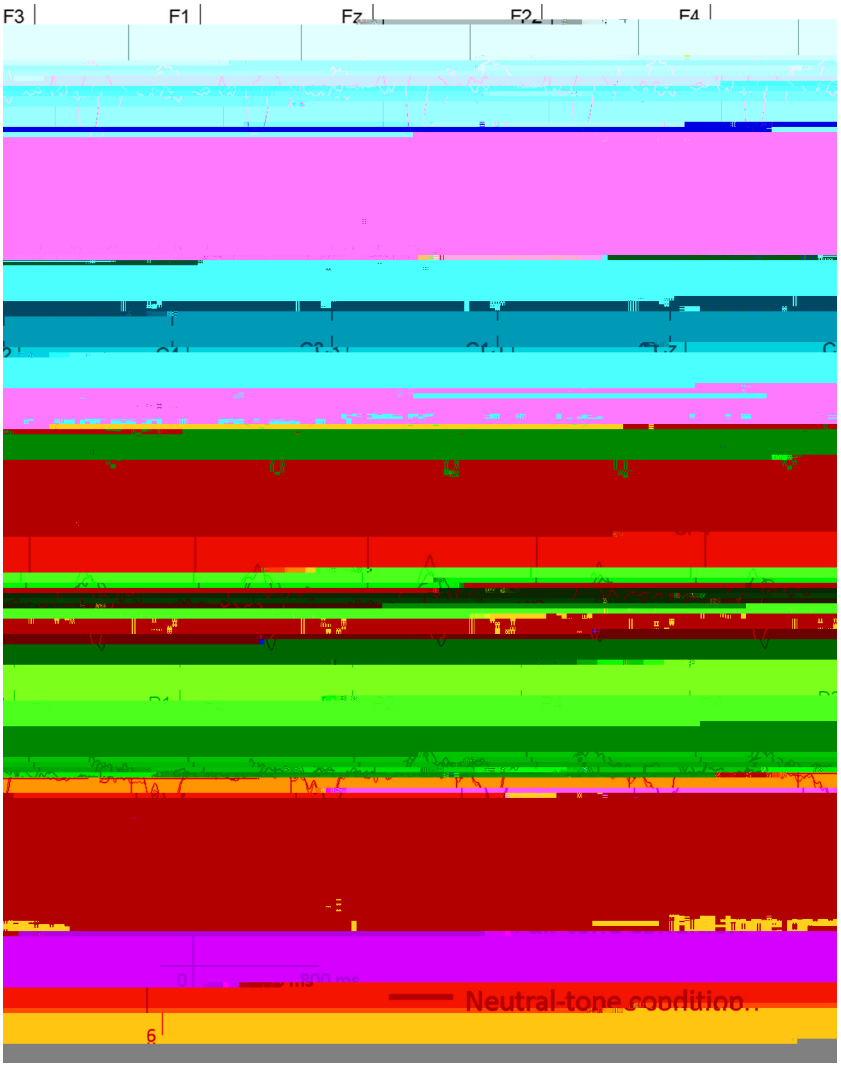

Fig. 4 Event-related potential waveforms for the target word

N400 Means for the N400 component are shown as a function of tone type and topographic location in Fig. 7. As can be seen, the mean amplitudes were distinctly more negative for neutral- than for full-tone targets, = –0.60 V, = 0.13, = 4.62. One topographic effect was reliable, with more negative center than lateral amplitudes, = –0.67 V, = 0.14, = –4.84. No other effect approached significance.

The inclusion of the N250 component as a covariate revealed a highly reliable covariate effect, = 0.34 V, = 0.04, = 8.12. The extraction of N250 variance further increased the size of the target type effect, = –0.72 V, = 0.13, = 5.73, and of the center–lateral difference, = –0.81 V, = 0.13, = 6.06.

P600 P600 amplitudes also yielded a significant effect of lexical tone, with more-negative amplitudes for neutral-tone words, = –0.32 V, = 0.14, = –2.19. In addition, one topographic effect was reliable, due to increases in negativity from anterior to posterior locations, = –0.14 V, = 0.05, = –2.80. No other effect was reliable. Inclusion of the N400 component in the statistical model revealed substantial carryover, = 0.80 V, = 0.03, = 28.09. When carryover was factored out, the effect of target type was no longer reliable, = 0.14 V, = 0.10, = 1.38.

Discussion

EEG recordings revealed less-negative N100 amplitudes for neutral- than for full-tone two-character Chinese compound words, and a corresponding N250 effect for anterior recording locations for native speakers of Standard Chinese. This effect was reversed for the N400 component, which was more negative for neutral-tone targets. Despite the well-matching lexical properties between the two conditions, the reading of pretarget words yielded discrepant ERP responses for the neutral- and the full-tone conditions at the late time window. However, the N100 effect observed on target words could not be solely due to spillover effects from the processing of the

The temporal properties of the two early effects of syllabic tone articulation, a broadly distributed N100 effect and an anterior N250 effect, are in general agreement with the timeline of prior work that examined ERPs in response to phonemic and supraphonemic manipulations with English text. As we noted earlier, briefly presented phonetic or syllabic primes that matched or mismatched the beginning phonetic or syllabic segment of English target words yielded robust N100 and N250 effects, with less negativity for matching than for mismatching prime–target pairs (Ashby, 2010; Ashby & Martin, 2008; Ashby et al., 2009).

The topographic distribution of the two early N100 and N250 ERP effects in Experiment 1 can also be reconciled with prior work. The N80–180 effect was not confined to specific recording sites, and broadly distributed early sub- and supraphonemic ERP effects have been reported in the literature (Ashby,

required more effort at a later point in time. This seeming reversal of the neutral-tone effects over time is also consistent with Yan et al. (2014), in which the processing of neutral-tone target words diminished the uptake of information from the next words(s), relative to full-tone words.

What accounts for the reversal of our neutral-tone effects? In their comprehensive review of N400 effects, Kutas and Federmeier (2009, 2011) noted that N400 amplitudes increased with items' difficulty and lack of familiarity. Related work suggests that this component may index late stages of word processing, when integrated multimodal lexical representations are constructed from phonological and orthographic forms (Laszlo & Federmeier, 2011), and when semantic processing converges upon a specific word meaning (Wlotko & Federmeier, 2012). Hence, one viable account for more negative N400 amplitudes for neutral-tone words may be conflicting articulations during early and later stages of neutral-tone target processing. Whereas the tonal features of full-tone targets did not differ at the morpheme and whole-word levels, the tonal features of neutral-tone words were level-specific. Integration of the two corresponding representations could thus have been more difficult for neutral-tone targets with incongruent second-syllable articulations at the character and word levels than for full-tone targets with congruent second-syllable articulations at the two levels, resulting in a larger N400 for neutral-tone words.

Our finding for lexical tone neutralization in Standard Chinese is not in accordance with the canonical frameworks of Chinese compound word recognition, in which the form representations of compound words consist purely of those of the individual morphemes (e.g., Taft & Zhu, 1995; Perfetti et al., 2005; Zhou & Marslen-Wilson,

effect did not reach significance for first-fixation durations (–4 ms), = –0.018 ms, = 0.013, = 1.41, it was reliable for gaze durations (–18 ms) and total viewing durations (–31 ms), = –0.046 ms, = 0.020, = 2.24, and = –0.072 ms, = 0.025, = 2.88, respectively. Additional analyses of regressions to the target, shown in Fig. 8, also revealed a significant effect of lexical tone, with fewer regressions to neutral- than to full-tone targets (6.3 % and 8.8 %, respectively), = –.352 [logits], = .171, = 2.06. Neutral-tone targets also received fewer first-pass fixations than full-tone targets (1.20 and 1.24, respectively). Although small, the effect was marginally significant, = .188 [logits], = .103, = 1.83, < .1. It replicates a corresponding finding in Yan et al. (2014) and is consistent with the effect of lexical stress in Ashby and Clifton (2005), in which words with one stress received fewer fixations than those with two stresses.

The analysis of AD effects revealed numerically shorter durations in the identical than in the two nonidentical AD syllable conditions, but the size of the syllable match effect was quite small for first-fixation durations and gaze durations (3 and 5 ms, respectively), and did not approach significance (both values < 1.5). The estimated effect size was larger and marginally reliable for total viewing durations (11 ms), = –.009 ms, = .005, = 1.85, < .1. No other AD effect approached significance, all values < 1.5.

Posttarget word region The target's lexical tone did not influence any of the three posttarget viewing duration measures, all values < 1.4. AD type influenced posttarget viewing, with longer—not shorter—viewing durations in the identical condition than in the two nonidentical AD conditions. The corresponding syllable match contrast was significant for first-fixation durations (8 ms), gaze durations (14 ms), and total

= 0005

was used to determine the time course and nature of the tonal

indicating that perception of irrelevant—but overlapping or similar—

structure may be more difficult and take more time. An analogous account has been offered to explain lexical stress effects during English word recognition. Specifically, longer viewing durations and more refixations for words with more stressed syllables were tentatively attributed to an increase in the difficulty with which suprasegmental phonological units were

Perfetti, C. A., & Tan, L. (1999). The constituency model of Chinese word identification. In J. Wang, A. Inhoff, & H.-S. Chen (Eds.), (pp. 115–134). Mahwah, NJ: Erlbaum.

R Development Core Team. (2014). R: A language and environment for statistical computing (Version 3.0.3). Vienna, Austria: R Foundation