available at www.sciencedirect.com
ScienceDirect
www.elsevier.com/locate/brainres
BRAIN RESEARCH

Research Report

Pitch accent and lexical tone processing in Chinese discourse comprehension: An ERP study

Xiaoqing Li[a,b], *Yufang Yang*[a], *Peter Hagoort*[b,c,*]

[a]*The Institute of Psychology, Chinese Academy of Sciences, China*
[b]*F.C. Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, The Netherlands*
[c]*Max Planck Institute for Psycholinguistics, The Netherlands*

ARTICLE INFO

Article history:
Accepted 5 May 2008
Available online 20 May 2008

Keywords:
Pitch accent
Lexical tone
Discourse processing
Semantic information
ERP
Chinese

ABSTRACT

In the present study, event-related brain potentials (ERP) were recorded to investigate the role of pitch accent and lexical tone in spoken discourse comprehension. Chinese was used as material to explore the potential difference in the nature and time course of brain responses to sentence meaning as indicated by pitch accent and to lexical meaning as indicated by tone. In both cases, the pitch contour of critical words was varied. The results showed that both inconsistent pitch accent and inconsistent lexical tone yielded N400 effects, and there was no interaction between them. The negativity evoked by inconsistent pitch accent had the some topography as that evoked by inconsistent lexical tone violation, with a maximum over central–parietal electrodes. Furthermore, the effect for the combined violations was the sum of effects for pure pitch accent and pure lexical tone violation. However, the effect for the lexical tone violation appeared approximately 90 ms earlier than the effect of the pitch accent violation. It is suggested that there might be a correspondence between the neural mechanism underlying pitch accent and lexical meaning processing in context. They both reflect the integration of the current information into a discourse context, independent of whether the current information was sentence meaning indicated by accentuation, or lexical meaning indicated by tone. In addition, lexical meaning was processed earlier than sentence meaning conveyed by pitch accent during spoken language processing.

1. Introduction

In the cognitive neuroscience of language, the role of prosody and its neural underpinning are still relatively underdeveloped areas. Prosody refers to suprasegmental properties of the speech signal. These include prosodic structure, pitch accent and intonation. As one of the three aspects of prosody, pitch accent reflects the relative prominence of a particular syllable, word, or phrase in a certain prosodic structure realized by modulation of pitch or by greater intensity. At the discourse level, one important function of pitch accent is to signal the state of the information structure (IS) in discourse. The ongoing research on IS has produced several theoretical definitions (e.g., Lambrecht, 1996; Prince, 1981). The main description of IS refers to the given/new or background/focus status. Even though there is not a one to one correspondence between focus and new information, the focus of a sentence is usually the new information. Considerable research has shown that in

* *Corresponding author.* F.C. Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands.
E-mail address: peter.hagoort@fcdonders.ru.nl (P. Hagoort).

0006-8993/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2008.05.031

spoken language, there is some correspondence between pitch accent and information structure. Speakers tend to place a pitch accent on new information (focus), while leaving given information unaccented or even de-accented (e.g., Bolinger, 1961; Halliday, 1967; Selkirk, 1995; Kadmon, 2001).

Although the mapping between pitch accent and information structure is complex, listeners do seem to use prosodic information about information structure in interpreting utterances. Several psycholinguistic studies using behavioral measures (acceptability judgments, reaction times) revealed that speech comprehension is facilitated when new information is accented and given information de-accented (Cutler, 1976; Bock and Mazzella, 1983; Terken and Noteboom, 1987; Nooteboom and Kruyt, 1987; Birch and Clifton, 1995; Dahan et al., 2002).

Recently several EEG studies investigated the correspondence between pitch accent and IS during spoken language processing. Hruska et al. (2001) investigated the mapping between IS and pitch accent with German as materials. They presented auditory sentences in discourse contexts that asked a narrow focus question about either a noun or a verb in the target sentences. The results revealed that when a focused word in the target sentence lacked a pitch accent, a broadly-distributed negativity was observed, which appeared to peak within 200–400 ms after the onset of the focused word. However, when a non-focused word had a pitch accent, no ERP response was observed. Subsequently, Johnson et al. (2003) investigated ERP effects of the correspondence between IS and pitch accent, with dialogues in English as materials. The results showed that there was an earlier negativity to a focused word, but no ERP response to the presence of extraneous pitch accent on given information. These effects were quite similar to the effects reported by Hruska et al. (2001).

Magne et al. (2005) used the ERP method to investigate the on-line use of pitch accent to process contrastive focus in French. The material also consisted of a question and an answer. The critical words were in either sentence-medial or -final position ("Did he give his fiancée a ring or a bracelet? He gave a ring to his fiancée."). The results revealed that when an incongruous pitch accent (both missing pitch accent on new information and superfluous pitch accent on given information) occurred in sentence-medial position, positive-going effects were elicited which belonged to the P300 family of components; the sentence-final incongruous pitch accent elicited a negative-going effect which resembled the N400 in the 300- to 600-ms latency range.

All of these studies targeted on-line processing of pitch accent and IS in speech processing. However, the electrophysiological results were not congruous. Whereas the studies of Hruska et al. (2001) and Johnson et al. (2003) yielded negativities, Magne et al. (2005) observed a positive-going effect in sentence-medial position, and a negative-going effect in sentence-final position. Hruska et al. (2001) and Johnson et al. (2003) only found an ERP effect for a missing pitch accent on new information; Magne et al. (2005) found an ERP effect for both missing pitch accent on new information and superfluous pitch accent on given information.

Last and most importantly, one question about the processing of pitch accent still remains unanswered. It is still unclear what the mechanisms are that underly pitch accent processing. During spoken language comprehension, listeners will identify and integrate the upcoming information into the ongoing discourse context. In spoken language, information is not only carried by words, but also by the speech prosody. Pitch accent on a word in sentence context also conveys sentence-level meaning, related to the focus distribution of the sentence (Gussenhoven, 1983; Selkirk, 1995). Does the processing of sentence meaning indicated by pitch accent involve the same neural mechanism as the processing of lexical meaning?

Previous studies have established that when the lexical meaning of words mismatches with its context, an N400 effect will result. Previous studies on the N400 have led to a fairly general consensus that within the language domain this ERP effect is sensitive to the degree of semantic fit between an incoming word and the relevant interpretive context, regardless of whether the latter was delineated by a larger piece of discourse or just the first part of an unfolding isolated sentence (Kutas and Van Petten, 1994; Van Berkum, Brown, and Hagoort, 1999; Hagoort and Brown, 2000; Van Berkum, Zwitserlood, Hagoort, and Brown, 2003). If pitch accent is processed in the same manner as lexical meaning, the mismatch between pitch accent and IS indicated by discourse context will evoke the same kind of ERP effect as the mismatch between lexical meaning and discourse context.

Chinese is suitable to investigate this question. Different from English, Chinese is a tone language. Tone refers to the pitch contour on a word that can distinguish lexical meaning. Dichotic-listening studies (Van Lanker and Fromkin, 1973), studies on patients suffering from aphasia (e.g., Gandour, 1998), and ERP studies (Brown-Schmidt and Canseco-Gonzalez, 2004; Schirmer et al., 2005) all established that in a tone language, tone is essential to ascertain word meaning; tone is processed linguistically just as phonemic information. In Chinese, the pitch contour is not only the acoustic cue of accentuation (Wang, Lü, and Yang, 2002), but also the acoustic cue of tone. Chinese provides a unique possibility to explore the same acoustic feature (pitch) for marking different levels of meaning, namely sentence meaning indicated by pitch accent vs. lexical meaning indicated by tone.

Therefore, we conducted an event-related potential study in Mandarin Chinese to further investigate the role of pitch accent in on-line spoken language comprehension, and, most importantly, to explore a potential difference in the nature and time course of the brain response to sentence meaning indicated by pitch accent and lexical meaning indicated by tone. Native Chinese speakers were presented with short spoken dialogues. In every dialogue, the pitch accent on the critical word was either appropriate or inappropriate relative to its information state in the discourse context. In addition, the lexical meaning of the critical word was made either consistent or inconsistent with the discourse context, by just varying its tone. This resulted in a total of four experimental conditions: "Correct", "Pitch accent violation", "Tone violation", and "Combined violation". We investigated the similarity in the neural mechanisms underlying pitch accent processing and the processing of lexical meaning in the discourse context by just varying the pitch contour of the critical words.

2. Results

Fig. 1 displays the grand average ERPs time-locked to the acoustic onset of the critical words in the continuous spoken discourse. Although there were no clear exogenous components in these ERPs, there were clear differential effects of Pitch accent and Tone. Relative to the Correct condition, the other three conditions all elicited a widely distributed negative deflection. The negative deflection peaked at about 400 ms, and reached its maximum over centro-posterior scalp sites. Given their latency and topography (see Fig. 2), we classified the negative deflections as N400 effects.

Analyses of variance (ANOVAs) were conducted using mean amplitudes in the following three latency ranges: a standard N400 latency range of 300–500 ms, 150–300 ms, and 600–900 ms after the acoustic onset of the critical word (CW). An overall repeated measures ANOVAs was conducted with Pitch accent (consistent, inconsistent), Tone (consistent, inconsistent), Laterality (left, midline, right) and Anteriority (F3/Fz/F4; C3/Cz/C4; P3/Pz/P4) as factors. Second, to explore the effects of the pitch accent violation, the tone violation, and the combined violation separately, three planned comparisons were performed. Third, we tested whether the effect of the combined violation was additive (i.e. the summation of the effects of Pitch accent violation and Tone

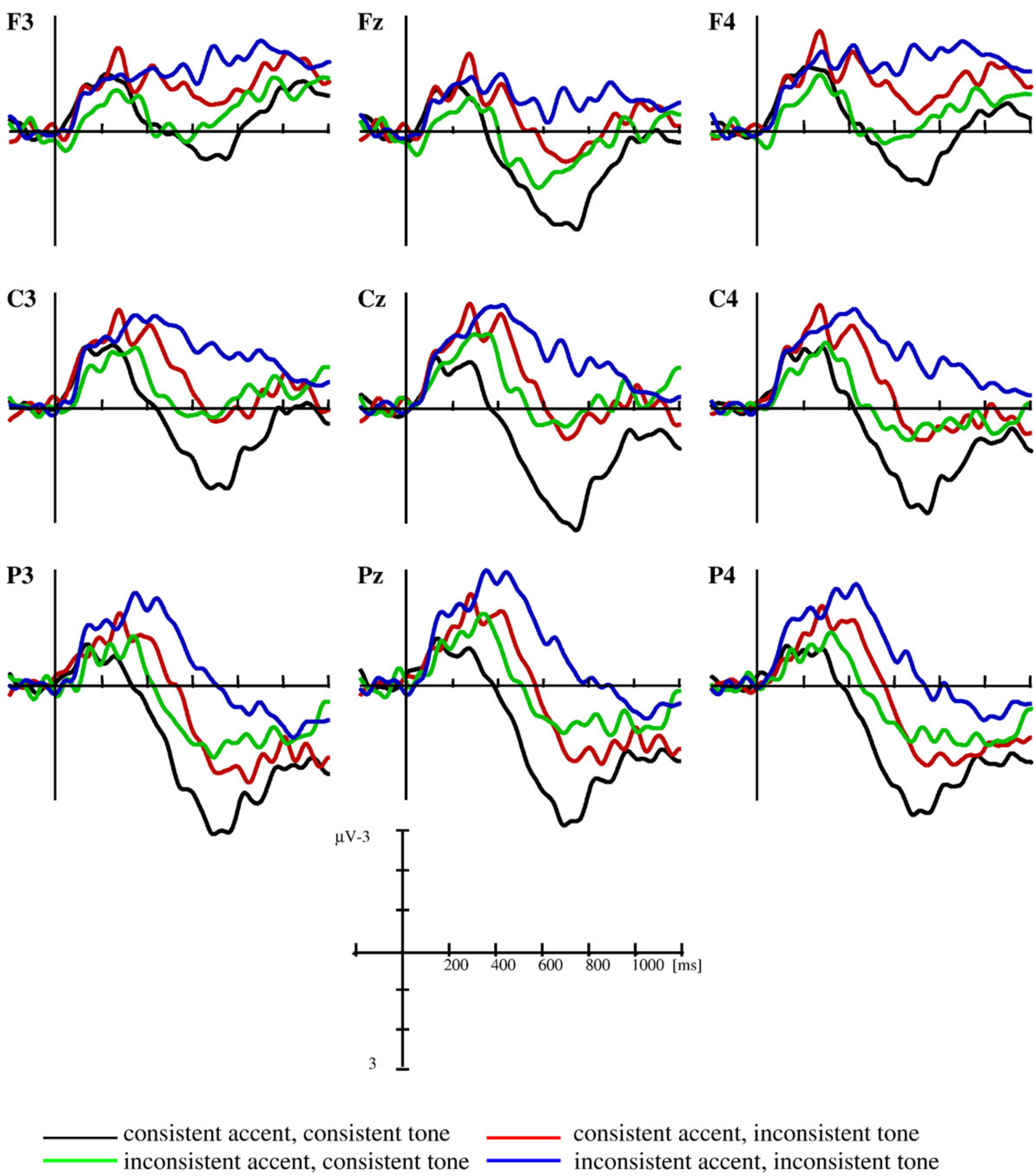

Fig. 1 – ERP effects of Pitch accent and Tone in spoken discourse comprehension.

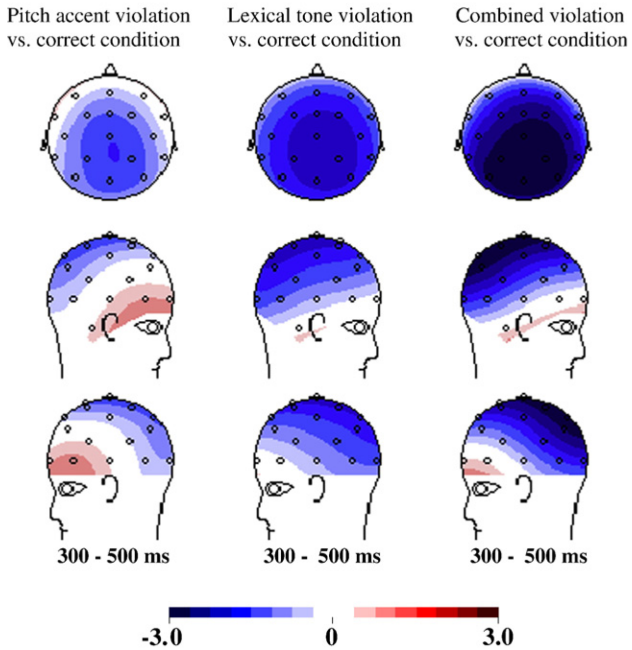

Fig. 2 – Topography of the N400 effects for pitch, tone and the combination of pitch and tone (Pitch accent violation– Correct, Lexical tone violation– Correct, and Combined violation– Correct) in the 300–500 ms latency window.

violation). When the degrees of freedom in the numerator were larger than one, the Greenhouse–Geisser correction was applied.

2.1. *Results in the 300–500 ms latency range*

The overall ANOVA for the time window of 300–500 ms revealed a significant main effect of Pitch accent ($F(1,15)=7.73$, $MSe=8.77$, $p<.05$), indicating that inconsistent pitch accent evoked a larger negative-going deflection (N400) than its consistent counterpart (effect magnitude: −0.69 μV). In addition, there was a significant main effect of Tone ($F(1, 15)=25.04$, $MSe=13.17$, $p<.00001$), due to the fact that inconsistent tone elicited a larger N400 than consistent tone (effect magnitude: −1.51 μV). However, the Pitch accent by Tone interaction failed to reach significance ($F(1, 15)=0.24$, $MSe=6.76$, $p=.63$). There was a significant interaction between Pitch accent and Anteriority ($F(2, 30)=9.10$, $MSe=1.52$, $p<.005$). Additional analyses showed that this interaction was due to the fact that the Pitch accent effect only reached significance on central and posterior sites ($F(1, 15)=6.11$, $MSe=4.57$, $p<.05$; $F(1, 15)=24.28$, $MSe=2.32$, $p<.00001$, respectively).

To establish the exact onsets of the Pitch accent effect and the Tone effect, we conducted a series of onset analyses in consecutive mean amplitude latency bins of 10 ms (e.g. 100–110 ms, 110–120 ms, etc.). Significance ($P<0.05$) on 5 consecutive bins was taken as evidence for the onset of a certain effect. According to this criterion, the main effect of Pitch accent was significant from 340–350 ms onwards. For the main effect of Tone, this criterion was reached in the latency range of 250–260 ms. Thus, the effect of Tone occurred approximately 90 ms earlier than the effect of Pitch accent during on-line spoken discourse comprehension.

The three planned comparisons resulted in main effects of Pitch accent ($F(1, 15)=6.42$, $MSe=7.02$, $p<.05$), Tone ($F(1, 15)=15.23$, $MSe=12.39$, $p<.001$), and the Combined accent–tone violation ($F(1, 15)=19.42$, $MSe=17.93$, $p<.001$). All three violations elicited a larger negative deflection than the Correct condition (effect magnitudes: −0.79 μV,−1.62 μV, −2.20 μV respectively).

Next, we analyzed the onset of the Pitch accent violation, Tone violation, and Combined violation effects in the three separate onset analyses. From 340–350 ms onwards, the effect of the pitch accent violation was found to be significant ($F(1, 15)=7.32$, $MSe=11.25$, $p<.05$). The effect of the tone violation reached significance in the 250–260 ms latency range ($F(1, 15)=5.90$, $MSe=7.73$, $p<.05$). The effect of the combined violation became significant at 320–330 ms ($F(1, 15)=5.12$, $MSe=27.04$, $p<.05$). All these effects remained significant for a long and continuous series of 10 ms latency bins.

Finally, we tested whether the effect of the combined violation was additive. To test this, we calculated the summed effect magnitudes of the Pitch accent and Tone violations. As Table 1 shows, in all of the three latency windows (150–300 ms, 300–500 ms, 600–900 ms), the sum of the effect magnitudes in the Pitch accent and Tone condition was approximately the same as the effect magnitude in the combined violation condition. This conclusion is supported by the absence of a statistically significant difference between the summed effect magnitudes and the combined violation effect. Similar analyses in the 150–300 ms and the 600–900 ms latency windows also resulted in the absence of statistical difference between these effect magnitudes.

2.2. *Results in the 600–900 ms latency range*

The overall ANOVA for the 600–900 ms latency window resulted in a main effect of Pitch accent ($F(1, 15)=23.12$, $MSe=15.81$, $p<.0001$; $F(1, 15)=28.15$), and a main effect of Tone ($F(1, 15)=19.29$, $MSe=19.07$, $p<.001$). Both consistent pitch accent and consistent tone evoked a larger positive deflection than their inconsistent counterparts (see Fig. 3; effect magnitudes: 1.59 μV, and 1.60 μV respectively). The interaction of Pitch accent and Anteriority also reached significance ($F(2, 30)=7.58$, $MSe=1.91$, $p<.01$), due to the posterior distribution. Pitch accent only reached significance over central and posterior sites.

Table 1 – ERP effects of Pitch accent violation (P), Tone violation (T) and combined violation (PT) in spoken language processing

Time window	P		T		Sum of P and T		PT	
	M	SE	M	SE	M	SE	M	SE
150–300 ms	0.04	0.30	−0.64	0.31	−0.60	0.53	−0.59	0.31
300–500 ms	−0.79	0.31	−1.62	0.42	−2.41	0.70	−2.20	0.50
600–900 ms	−1.66	0.36	−1.67	0.41	−3.33	0.67	−3.19	0.59

Notice: P refers the grand average effect for the pitch accent violation. T refers to the grand average effect for the tone violation. PT refers to the combined violation. Sum of the P and T refers to the summation of the pitch and tone violation effects.

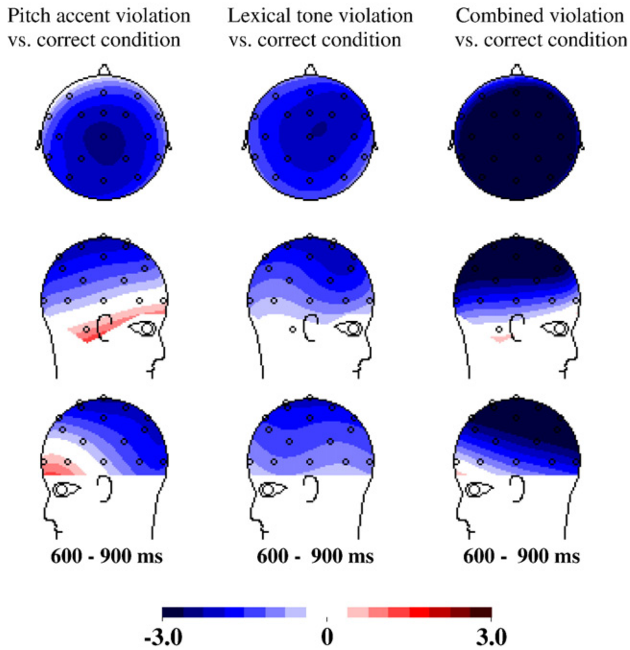

Fig. 3 – Topography of the effects for pitch, tone, and the combination of pitch and tone (Pitch accent violation– Correct, Lexical tone violation– Correct, and Combined violation– Correct) in the 600–900 ms latency window.

Planned comparisons in this latency range resulted in main effects of Pitch accent, Tone, and the combined violation ($F(1, 15)=21.57$, $MSe=9.24$, $p<.0001$; $F(1, 15)=16.70$, $MSe=12.01$, $p<.001$; $F(1, 15)=29.26$, $MSe=25.06$, $p<.0001$, respectively). This was due to the fact that the Correct condition evoked a larger positive deflection than the other three conditions.

2.3. *Results in the 150–300 ms latency range*

In the 150–300 ms window latency, the overall ANOVA only revealed a significant main effect of Tone ($F(1, 15)=9.84$, $MSe=5.93$, $p<.01$); the inconsistent tone elicited a larger negative deflection than the consistent tone (effect magnitude: −0.64 μV). In the three planned comparisons, no significant differences were obtained.

3. Discussion

In this experiment we investigated the role of pitch accent in spoken language comprehension, and the potential difference in the nature and time course of the ERP response to sentence meaning indicated by pitch accent, and lexical meaning indicated by tone. The results of this experiment can be summarized as follows. First, not only variation of tone but also variation of pitch accent influenced the amplitude of the N400. Listeners could rapidly extract the information structure aspect indicated by pitch accent and relate it to ongoing discourse context. Second, there was no significant interaction between pitch accent and tone. Moreover, the effect of the combined violation is additive, in that it equals the sum of the separate effects of pitch accent and tone. Finally, during spoken language comprehension, the tone effect appeared approximately 90 ms earlier than the pitch accent effect. These results will be discussed in more detail below.

3.1. *The role of pitch accent in spoken language comprehension*

The current experiment revealed that relative to the presence of pitch accent, the absence of such accent on focused information evoked a larger N400. It demonstrated that pitch accent can be used on-line by listeners to understand the information state of the message, and can be integrated with ongoing discourse context immediately.

As mentioned in the introduction, results of previous behavioral experiments have already shown that speech comprehension is facilitated when new information is accented (e.g., Cutler, 1976; Bock and Mazzella, 1983; Terken and Noteboom, 1987; Nooteboom and Kruyt, 1987; Birch and Clifton, 1995; Dahan et al., 2002). This experiment extends those results by showing that listeners can interpret the meaning indicated by pitch accent very quickly during spoken language comprehension. This on-line pitch accent effect is in line with other ERP experiments investigating the correspondence between pitch accent and IS in spoken language processing. Hruska et al. (2001) and Johnson et al. (2003) reported an earlier negativity to a focused word when it lacked a pitch accent. Magne et al. (2005) also found an ERP effect of a missing pitch accent on new information and a superfluous pitch accent on given information. Those off-line and on-line measures all point to the conclusion that pitch accent is used by listeners to process the information state of the message in a discourse context.

However, there were also some differences between the present result and the results of a previous ERP study exploring pitch accent processing (Li et al., in press). In this earlier study we found that the presence of a pitch accent on focused information elicited a larger N400 than when such an accent is missing (Li et al., in press). This finding was remarkable, since the N400 was larger to accented information that is in agreement with the information structure than to de-accented information which violates the requirement that focused information carries stress. We interpreted this finding as an indication that more integration resources are devoted to words whose information value is marked by pitch accent. In the current study we observed the reversed effect: The N400 was larger to focused information that was unstressed than to focused information carrying a pitch accent. The reason for the difference in the results of these two experiments might have something to do with the type of materials that we used. In contrast to the previous experiment, in the current experiment we used questions as the context for the subsequent target sentence. The results of previous psycholinguistic experiments have shown that question contexts generate stronger contextual predictions than non-question contexts (Altmann et al., 1998). The structure of the question–answer pairs in the current experiment induced strong expectations or constraints on the lexical item that is in focus. In this situation, a mismatch with the strong contextual constraint might override the saliency effect of the pitch accent in less constrained contexts.

3.2. The role of lexical tone in spoken language comprehension

The current results also showed that a violation of lexical tone elicited an N400 effect. In this experiment, the lexical meaning marked by tone was consistent with the local sentence context, and only inconsistent with respect to the wider discourse context. Therefore, the results established that during spoken language comprehension, listeners could quickly identify the lexical meaning indicated by tone, and integrate it into the discourse context on-line.

The present results were consistent with the nature of tone in Chinese. The pronunciation of virtually every syllable of every word in Mandarin Chinese is defined not only by the vowels and consonants but also by its tone. Tone is the pitch contour realized on a syllable, which provides information about the meaning of a word and distinguishes different words in the lexicon from each other (Ho and Bryant, 1997; Wang, 1973). Thus, a violation of tone on the critical words reflects a violation of lexical meaning, hence eliciting an N400 effect in discourse context.

Previously, there have been ERP studies investigating the nature of tone processing in Chinese sentences. In both Brown-Schmidt and Canseco-Gonzalez (2004) and Schirmer et al. (2005), three types of semantic anomalies were created by manipulating the appropriateness of tone, the syllable, or both tone and syllable of the critical word relative to a sentence context. The results showed a robust N400 effect for all three types of anomalies. Most importantly, the semantic violations induced by tone and syllable were comparable. The N400 effect elicited by a tone violation in the current experiment not only was in line with but also extended those two ERP studies, by showing that the lexical meaning indicated by tone is very quickly integrated not only in a sentence context but also into a wider discourse context.

3.3. Similarities between pitch accent processing and lexical tone processing

A very important motivation for the current study was to investigate potential differences in the nature of brain responses to sentence meaning indicated by accentuation and lexical meaning indicated by tone. Previous behavioral and ERP studies already established that sentence processing is speeded up when the accent placement is appropriate for the information structure requirements of the sentence in its context (e.g., Bock and Mazzella, 1983; Terken and Noteboom, 1987; Van Donselaar and Lentz, 1994); and that pitch accent information can be used on-line during spoken language comprehension (Hruska et al., 2001; Johnson et al., 2003; Magne et al., 2005). However the mechanisms behind these effects are unclear. One possible reason might be that the effects may simply reflect heightened attention to a word with a pitch accent. Early work by Cutler (1976) showed heightened attention (as indicated by faster phoneme monitoring responses) to a word that answered a preceding question (and hence was focused) and received a pitch accent. The other possible reason is that pitch accent not only mediates attention allocation but also conveys linguistic meaning directly. The distribution of accents has often been explained by appealing to the information status of the expressions in the utterance (e.g., Bolinger, 1985; for a review, see Ladd, 1996). For example, a word that is accented is often assumed to introduce new information into the discourse, as opposed to information that is already established or given. Therefore, it is expected that pitch accent will be processed like other types of linguistic information.

The results of the current experiment are consistent with the latter point of view. In this experiment, we compared the ERP responses to pitch accent violation and lexical tone violation relative to an ongoing discourse context. It was shown that both inconsistent Pitch accent and inconsistent Tone elicited N400 effects and there was no significant interaction between them; the negativity for Pitch accent violation and Tone violation had a similar central-posterior distribution. Our results indicate that there might be a correspondence between the neural mechanism underlying pitch accent and lexical meaning processing in context.

The semantic processing of pitch accent is consistent with previous linguistic and psycholinguistics studies. In spoken language, information is not only carried by words, but also by the speech melody (prosody). Accents can be seen as morphemes that convey the focus distribution of the sentence (Gussenhoven, 1983; Selkirk, 1995). Changing the accent pattern of an utterance, by accenting some words and failing to accent others, can change the meaning of an utterance dramatically. For example, Most and Saltz (1979) asked listeners to pair "wh" questions with their appropriate answers. They found that the position of the accentuation (focal accent) in the answer largely determined the listener's pairing choices. Reference resolution studies also showed that listeners interpret an accented word as introducing a new discourse entity and a de-accented word as anaphoric (e.g., Dahan et al., 2002). The current and previous studies together establish that in spoken language, pitch accent not only influences syntactic and semantic processing indirectly but also conveys specific meaning aspects directly.

3.4. The time course of pitch accent processing and tone processing

An important aspect of the current study is that it provides information about the speed with which discourse information is integrated with incoming information, especially sentence meaning indicated by pitch accent and lexical meaning indicated by tone. The results showed that both pitch accent violation and tone violation elicited an N400 effect, but the processing of inappropriate information varied according to the type of violation. During spoken language comprehension, lexical tone processing was approximately 90 ms earlier than pitch accent processing.

One possible explanation for the earliness of the tone effect relative to the pitch accent effect might be that lexical tone becomes psychoacoustically salient before pitch accent. Although in Chinese pitch is both the acoustic cue of accent and of tone, their acoustic correlates need not be identical in their temporal profile. This could then be an alternative explanation for the 90 ms difference between the tone and accent effect. However, we consider this alternative explanation unlikely, for the following reason. The pitch correlate of every syllable in Chinese is not a single point but a pitch contour, which is called register. The pitch contour has a top and a bottom line and only the former has to do with accentuation (Wang et al., 2002). Likewise, the perceptual center of tones in Chinese is also mainly located in the higher part of pitch contour (Liang, 1963). Therefore, in the current study the prosodic information distinguishing the correct and modified words occurred at about the same time for both pitch accent violation condition and tone violation condition. The finding that the effect of tone violation appeared earlier than that of pitch accent violation is thus most likely not due to acoustic differences of tone and pitch accent. However, this explanation cannot be proven wrong completely. This would require a gating study in which the moments in time at which perception of tone and accentuation takes place is established empirically. Another, more likely explanation for the earliness of the tone effect has to do with the semantic nature of tone and pitch accent in Chinese. In Chinese, tone information is a salient aspect of a word as it regularly determines which meaning, among a variety of possible meanings, is intended by the speaker. Lexical tone influences lexical access directly (Brown-Schmidt and Canseco-Gonzalez, 2004; Schirmer et al., 2005). However, in the present experiment pitch accent influenced meaning at a higher level, namely the information structure of the sentence in connection to previous discourse.

In addition, the current result also showed that the N400 effect evoked by a tone violation occurred at about 250 ms after the acoustic onset of critical words. This slightly delayed onset of the N400 effect in the current study has to do with the characteristics of lexical tone. In previous researches on spoken language processing, the critical words in the semantic mismatch condition generally differed from those in the correct condition (e.g., slow/quick) already at the initial consonant. However, in the current study, the critical words in the semantic mismatch condition (lexical tone violation condition) did not differ from those in the correct condition until the tone appeared. Lexical tone differences arise at the rhyme of the syllable, but not at its onset (Ho and Bryant, 1997). Therefore, tone differences will be available slightly later than a difference in syllable onset (cf. Brown-Schmidt and Canseco-Gonzalez, 2004).

4. Conclusion

In conclusion, results of the current study are consistent with findings in previous research, by establishing that pitch accent and lexical tone are processed on-line during spoken language comprehension. Listeners can quickly identify the information state indicated by pitch accent and lexical meaning indicated by tone. They immediately interpret them relative to the ongoing discourse context. The present results also extended previous studies, showing that there was a correspondence between the N400 effects evoked by a pitch accent violation and a lexical tone violation. They both reflect semantic integration into the discourse context (cf Hagoort and van Berkum, 2007). It is suggested that pitch accent conveys information state semantics and is processed linguistically. Moreover, during on-line spoken language comprehension, word meaning was processed approximately 90 ms earlier than the sentence-level meaning conveyed by pitch accent.

5. Experimental procedures

5.1. *Subjects*

Sixteen right-handed subjects (9 females) participated in the experiment; all of them were native speakers of Mandarin Chinese. None of them had any neurological impairment, had experienced any neurological trauma, or used neuroleptics.

5.2. *Stimuli*

The experiment materials consisted of 200 discourses which were spoken by a female speaker and recorded at a sampling rate of 20 kHz. Every discourse consisted of three sentences. The first sentence introduced some background information related to the critical sentence. The second sentence is a question which sets a narrow focus for the third sentence. The focused word in the third sentence was the critical word. This was a one-character noun. As Table 2 shows, in the third sentence, pitch accent is either appropriately placed on the critical word (focus), or inappropriately on another word in front of it[1]. Moreover, the lexical meaning of the critical word was either consistent or inconsistent with the previous dialogue context. This was realized by varying the lexical tone of the critical word. This resulted in a full factorial design with four experimental conditions (see Fig. 4): Correct (appropriate pitch accent, appropriate lexical tone), Tone violation (appropriate pitch accent, inappropriate lexical tone), Pitch accent violation (inappropriate pitch accent, appropriate lexical tone), Combined violation (inappropriate pitch accent, inappropriate lexical tone).

When constructing the materials, the following constraints applied. First, the critical words are all one-character words. Second, the critical words do not appear in the previous discourse, in order to avoid repetition effects. Third, the critical words are all in sentence-medial position in order to avoid a mixture of local effects and global sentence wrap-up effects (Hagoort, 2003). Fourth, the word frequency is controlled for the two critical words with different tones (mean 1815.27 vs. mean 1830.47).

Experimental materials were grouped into 4 lists of 200 dialogues according to the Latin square procedure based on the

[1] In Chinese, the acoustic correlate of accentuation is mainly the high point of the pitch contour of the critical word (Wang, Lü, Yang, 2002; Zhong, Wang, Yang, 2001). To establish that our speaker had accented the relevant words as intended, we performed a T-test on the spoken sentence materials to verify that the accented and deacccented words differed in the high point of the pitch contour. The results showed that for the focused critical words, the high point of pitch contour was significantly higher when the focused words received pitch accent than when they did not ($T_{(1, 199)}=24.47$, $MSe=2.19$, $p<.0001$; $T_{(1, 199)}=23.32$, $MSe=2.34$, $p<.0001$ for the words with appropriate tone and inappropriate tone respectively). Accordingly, for the particular words preceding the focused words, the high point of pitch contour was significantly lower when the focused words were accented than when they were deaccented ($T_{(1, 199)}=24.19$, $MSe=2.20$, $p<.0001$; $T_{(1, 199)}=18.48$, $MSe=2.55$, $p=.0001$ for appropriate tone and inappropriate tone condition respectively). Therefore, the acoustic measurements confirmed that target sentences were spoken with the intended pitch pattern.

Table 2 – Example of the four experimental conditions

For example, critical word: 花 huā (flower) 画 huà (picture)

(1) appropriate accentuation and appropriate tone	
现在正是玫瑰盛开的季节。	Now the roses are in full bloom.
明天小明去买什么把房间装饰一下？	Tomorrow what is Xiaoqin going to buy to decorate the house?
明天小秦去买(*花*)装饰房间。	Tomorrow *Xiaoqin* will buy (*flowers*) to decorate the house.
(2) appropriate accentuation and inappropriate tone	
现在正是玫瑰盛开的季节。	Now the roses are in full bloom.
明天小明去买什么把房间装饰一下？	Tomorrow what is Xiaoqin going to buy to decorate the house?
明天小秦去买(*画*)装饰房间。	Tomorrow *Xiaoqin* will buy (*pictures*) to decorate the house.
(3) inappropriate accentuation and appropriate tone	
现在正是玫瑰盛开的季节。	Now the roses are in full bloom.
明天小明去买什么把房间装饰一下？	Tomorrow what is Xiaoqin going to buy to decorate the house?
明天(小秦)去买*花*装饰房间。	Tomorrow (*Xiaoqin*) will buy *flowers* to decorate the house.
(4) inappropriate accentuation and inappropriate tone	
现在正是玫瑰盛开的季节。	Now the roses are in full bloom.
明天小明去买什么把房间装饰一下？	Tomorrow what is Xiaoqin going to buy to decorate the house?
明天(小秦)去买*画*装饰房间。	Tomorrow (*Xiaoqin*) will buy *pictures* to decorate the house.

Note. The underlined words are the two critical words. ERPs are aligned to the underlined and italic words. Brackets indicate accentuation.

four experimental conditions. In every list, there was an equal number of discourses for every experimental condition, and there were no discourses repeated across the four conditions. In addition, in every list there were also 60 filler dialogues (30 correct discourses, 15 with a standard semantic violation, 15 with a syntactic violation). Subjects were divided into 4 groups, with each group listening to only one list of materials.

5.3. *Experimental protocol*

After the electrodes were positioned, subjects were asked to listen to each discourse for comprehension. Meanwhile, their event-related brain potentials (ERP) were registered. The subjects were told that EEG recording would only occur while they listened to the last sentence of a discourse, and during that time they should avoid making (eye) movements.

Each trial consisted of a 300 ms auditory warning tone, followed by 700 ms of silence, context sentences (introduction and question), 250 ms silence, 500 ms asterisk (*), 1000 ms silence, and the target sentence. The subjects were told that when they saw the asterisk on the screen they should fixate and sit still for EEG recording until the beginning of next discourse. After a short practice session, consisting of eight discourses, the trials were presented in five blocks of about 14 min each, separated by brief resting periods.

5.4. *Evoked potential recording*

EEG was recorded from 32 Ag/AgCl sintered electrodes mounted in an elastic cap. Twenty eight electrodes (Fz, FCz, Cz, Pz, Fp1, Fp2, F7, F8, F3, F4, FC5, FC6, FC1, FC2, T7, T8, C3, C4, CP5, CP6, CP1, CP2, P7, P8, P3, P4, O1, and O2) were placed according to the standard system of the American Electroencephalographic Society (1994). All electrodes were referenced to the left mastoid on-line. The EEG electrodes were re-referenced off-line to linked mastoids. Vertical eye movements were monitored via a supra- to sub-orbital bipolar montage. A right to left canthal bipolar montage was used to monitor for horizontal eye movements.

EEG and EOG data were amplified with a BrainAmp DC EEG amplifier, using a high cutoff of 30 Hz and a time constant of 10 s. Electrode impedances were kept below 3 kΩ for the EEG recording and below 5 kΩ for the EOG recording. The EEG and EOG signals were digitized on-line with a sampling frequency of 500 Hz.

5.5. *Preprocessing of ERP data*

The EEG data were screened for eye movements, muscle artifact, electrode drifting, and amplifier blocking in a critical window ranging from 200 ms before to 1200 ms after the acoustic onset of the critical word. Trials containing such artifacts were rejected (5.4% overall). Rejected trials were evenly distributed among conditions. For each subject, average waveforms were computed across all remaining trials per condition. This was done after normalizing the waveforms of the individual trials relative to the mean amplitude of a 200 ms baseline before the onset of the critical word.

5.6. *Statistical analysis*

Analysis of variance (ANOVAs) used mean amplitude values computed for each subject, condition, and electrode in the

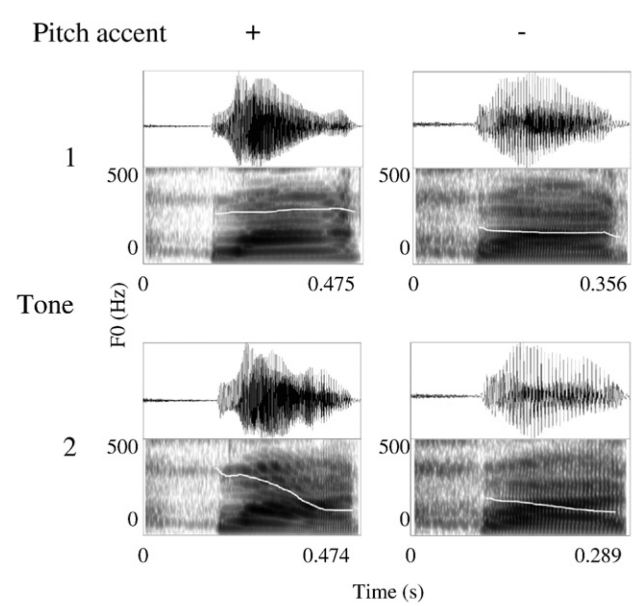

Fig. 4 – Four paired stimuli illustrating the critical word (hua) in the four different experimental conditions. The data set consists of voice spectrographs with uncorrected fundamental frequency (pitch) contours superimposed as a white line. Upper, "hua" with high level tone; Lower, "hua" with falling tone; Left, "hua" with pitch accent; Right, "hua" without pitch accent. (Figure created using PRAAT software.)

standard N400 latency range of 300–500 ms after acoustic onset of the CW. Two supplementary latency ranges (150–300 ms and 600–900 ms) were used to complement the standard latency range analysis. Analyses of variance were conducted on a selection of three midline electrodes (Fz, Cz, Pz) and six lateral electrodes (F3/F4; C3/C4; P3/P4).

REFERENCES

Altmann, G.T.M., Garnham, A., van Nice, K., Henstra, J.A., 1998. Late closure in context. J. Mem. Lang. 38, 459–484.

Birch, S., Clifton Jr., C., 1995. Focus, accent, and argument structure: effects on language comprehension. Lang. Speech 38, 365–391.

Bock, J.K., Mazzella, J.R., 1983. Intonational marking of given and new information: some consequences for comprehension. Mem. Cogn. 11, 64–76.

Bolinger, D., 1961. Contrastive accent and contrastive stress. Language 37, 83–96.

Bolinger, D., 1985. Two views of accent. J. Linguist. 21, 79–123.

Brown-Schmidt, S., Canseco-Gonzalez, E., 2004. Who do you love, your mother or your horse? An event-related brain potential analysis of tone processing in Mandarin Chinese J. Psycholinguist. Res. 33 (2), 103–135.

Cutler, A., 1976. Phonome-monitoring reaction time as a function of preceding intonation contour. Percept. Psychophys. 20, 55–60.

Dahan, D., Tanenhaus, M.K., Chambers, C.G., 2002. Accent and reference resolution in spoken-language comprehension. J. Mem. Lang. 47, 292–314.

Gandour, J., 1998. Aphasia in tone languages. In: Coppens, P., Lebrun, Y., Basso, A. (Eds.), Aphasia in Atypical Populations. Lawrence Erlbaum, London.

Gussenhoven, C., 1983. Focus, mode and the nucleus. J. linguist. 19, 377–417.

Hagoort, P., 2003. Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. J. Cogn. Neurosci. 15, 1–17.

Hagoort, P., Brown, C.M., 2000. ERP effects of listening to speech: semantic ERP effects. Neuropsychologia 38, 1518–1530.

Hagoort, P., Van Berkum, J.J.A., 2007. Beyond the sentence given. Philos. Trans. R. Soc. B 362, 801–811.

Halliday, M.A.K., 1967. Notes on transitivity and theme: Part 2. J. Linguist. 3, 177–244.

Ho, C.S., Bryant, P., 1997. Development of phonological awareness of Chinese children in Hong Kong. J. Psycholinguist. Res. 26 (1), 109–126.

Hruska, C., Alter, K., Steinhauer, K., Steube, A., 2001. Misleading dialogues: human brain's reaction to prosodic information. Paper presented at the Oralite et Gestualite, Aix en Provence, France, June, 2001.

Johnson, S.M., Breen, M., Clifton, C.J.R., Morris Florak, J., 2003. ERP investigation of prosodic and semantic focus. Poster presented at Cognitive Neuroscience, New York City.

Kadmon, N., 2001. Formal pragmatics. Blackwell Publishers, Malden, MA.

Kutas, M., Van Petten, C.K., 1994. Psycholinguistics electrified: event-related brain potential investigations. Gernsbacher MA editor, Handbook of psycholinguistics. Academic Press, San Diego, pp. 83–143.

Ladd, D.R., 1996. Intonational phonology. Cambridge University Press, Cambridge.

Lambrecht, K., 1996. Information structure and sentence form: Topic, focus, and the mental representation of discourse reference. Cambridge University Press.

Li, X.Q., Hagoort, P., Yang, Y.F., in press. ERP Evidence on the Influence of Accentuation in Spoken Discourse Comprehension in Chinese. J. Cogn. Neurosci. 20, 906–915.

Liang, Z.A., 1963. Auditory identification evidence for tone perception in Chinese. Acta Physiol. Sin. 2, 85–91.

Magne, C., Astésano, C., Lacheret-Dujour, A., Morel, M., Alter, K., Besson, M., 2005. On-line processing of "pop-out" words in spoken French dialogues. J. Cogn. Neurosci. 17, 740–756.

Most, R.B., Saltz, E., 1979. Information structure in sentences: New information.

Nooteboom, S.G., Kruyt, J.G., 1987. Accent, focus distribution, and the perceived distribution of given and new information: an experiment. J. Acoust. Soc. Am. 82, 1512–1524.

Prince, E.F., 1981. Taxonomy on the given and new status. In: Cole, P. (Ed.), Radical Pragmatics. Academic Press, New York, pp. 223–255.

Schirmer, A., Tang, S.L., Penney, T.B., Gunter, T.C., Chen, H.C., 2005. Brain Responses to Segmentally and tonally induced semantic violations in cantonese. J. Cogn. Neurosci. 17, 1–12.

Selkirk, E.O., 1995. Sentence prosody: intonation, stress, and phasing. In: Goldsmith, J. (Ed.), Handbook of phonological theory. Blackwell, Oxford, pp. 550–569.

Terken, J., Noteboom, S.D., 1987. Opposite effects of accentuation and deaccentuation on verification. Latencies for Given and New Information. Lang. Cogn. Processes 2 (3/4), 145–163.

Van Berkum, J.J.A., Brown, C.M., Hagoort, P., 1999. Early referential context effects in sentence processing:evidence from event-related brain potentials. J Mem. Lang. 41, 147–182.

Van Berkum, J.J.A., Zwitserlood, P., Hagoort, P., Brown, C.M., 2003. When and how do listeners relate a sentence to the wider discourse? Cogn Brain Res. 17, 701–718.

Van Donselaar, W., Lentz, J., 1994. The Function of sentence accents and given/new information in speech processing: different strategies for normal-hearing and hearing-impaired listeners? Lang. Speech 37, 375–391.

Van Lanker, D., Fromkin, V.A., 1973. Hemispheric specialization for pitch and "tone": Evidence from Thai. J. Phon. 1, 101–109.

Wang, B., Lü, S.N., Yang, Y.F., 2002. The pitch movement of stressed syllable in Chinese sentences. Acta Acustica 27 (3), 234–240.

Wang, W.S.Y., 1973. The Chinese Language. Sci. Am. 228, 50–60.

Zhong, X.B., Wang, B., Yang, Y.F., 2001. The perception of prosodic word stress standard Chinese. Acta Psychol. Sin. 33 (6), 481–488.