Event-related Potential Evidence on the Influence of Accentuation in Spoken Discourse Comprehension in Chinese

Xiaoqing Li[1], Peter Hagoort[2], and Yufang Yang[1]

Abstract

■ In an event-related potential experiment with Chinese discourses as material, we investigated how and when accentuation influences spoken discourse comprehension in relation to the different information states of the critical words. These words could either provide new or old information. It was shown that variation of accentuation influenced the amplitude of the N400, with a larger amplitude for accented than for deaccented words. In addition, there was an interaction between accentuation and information state. The N400 amplitude difference between accented and deaccented new information was smaller than that between accented and deaccented old information. The results demonstrate that, during spoken discourse comprehension, listeners rapidly extract the semantic consequences of accentuation in relation to the previous discourse context. Moreover, our results show that the N400 amplitude can be larger for correct (new, accented words) than incorrect (new, deaccented words) information. This, we argue, proves that the N400 does not react to semantic anomaly per se, but rather to semantic integration load, which is higher for new information. ■

INTRODUCTION

Spoken language is the archetypical form of language. An important difference with written language is that spoken language carries prosodic information. Prosody refers to one type of super-segmental information in the speech signal. Super-segmental properties cannot be derived from the underlying sequence of single phonetic segments because they usually involve two or more segments and occur simultaneously with those segments. In prosody, one distinguishes prosodic structure, accentuation, and intonation. Prosodic structure refers to the relative temporal groupings of words in speech. Accentuation refers to the relative prominence of a particular syllable, word, or phrase in a certain prosodic structure realized by greater intensity or by modulation of pitch. Intonation refers to the way in which the voice pitch rises and falls across (parts of) an utterance. During spoken language processing, a variety of syntactic, semantic, and phonological cues converge to achieve comprehension. However, the majority of research has focused on semantic or syntactic information during comprehension. How and when prosody influences spoken language comprehension is still a relatively underdeveloped area.

This study focused on semantic aspect of prosody, namely, accentuation. At the discourse level, one impor-

tant function of prosody, especially accentuation, is to signal the information state of discourse constituents. Based on perception experiments and corpus analysis, considerable research has shown that, in spoken language, there is some correspondence between accentuation and information structure. Speakers tend to place a pitch accent on new information, while leaving given information deaccented (Wang, 2002; Zhang, 2002; Zhu & Miao, 2000; Bock & Mazzella, 1983; Cutler, 1976).

Not only do speakers use accentuation to indicate information structure but it has also been found that accentuation plays an important role in spoken discourse processing. Listeners consider a sentence inappropriate when new information is unaccented and old information is accented, and they are more likely to indicate that a sentence is appropriate when new information is accented and old information is unaccented (Birch & Garnsey, 1995). Moreover, compared to inappropriate accentuation, appropriate accentuation speeds up sentence processing during discourse comprehension (Donselaar & Lentz, 1994; Terken & Noteboom, 1987; Bock & Mazzella, 1983; Cutler, 1976).

However, most of these studies involved off-line methods such as end-of-discourse comprehension or appropriateness judgments. The exceptions are Dahan, Tanenhaus, and Chambers' (2002) and Venditti et al.'s (2002), who, by using an eye-tracking technique, were able to investigate on-line spoken language processing as the discourse unfolded in real time. Dahan et al.

[1]Chinese Academy of Sciences, Beijing, People's Republic of China, [2]Radboud University, Nijmegen, Netherlands

(2002) showed that the discourse implications of an accented word could be inferred very quickly, well before the accented words had been fully heard. In their study, the visual scene contained various objects, two of which shared the same primary-stressed first syllable (e.g., *candle* and *candy*). The first part of the auditory instruction introduced either the candle or the candy into the discourse context (*Put the candle/candy below the triangle.*), establishing it as given. The second part of the instruction then accented or deaccented the candle (*Now put the CANDLE/candle above the square*). Fixations to the competitor (e.g., candy) demonstrated a bias to interpret deaccented nouns as given and accented nouns as new (Dahan et al., 2002).

Previously, electroencephalogram (EEG) has also been used to investigate the influence of prosody in spoken language processing. Holcomb and Neville (1991) found that the auditory N400 effect has an earlier onset in natural speech than is observed for visually presented words or sentences. They suggested that prosodic cues in spoken sentences could influence the time course of the classic N400 semantic anomaly effect (Holcomb & Neville, 1991). However, they did not directly investigate the role and time course of prosody effects (especially accentuation) in spoken language comprehension. In their ERP study, Hruska, Steinhauer, Alter, and Steube (2000) investigated the on-line processing of pitch accent more directly. They found that when a focused word lacked a pitch accent, a broadly distributed negativity was observed. No ERP response was observed to the presence of an extraneous pitch accent on a nonfocused word. However, the materials they used were question–answer pairs which, as shown by the results of several psycholinguistic experiments, generate stronger contextual predictions than nonquestion contexts (Altmann, Garnham, van Nice, & Henstra, 1998). Therefore, their results cannot be generalized to the role of prosody in spoken discourse comprehension.

The above studies contributed to the understanding of how the correspondence between accentuation and information structure influences spoken discourse processing. However, there are still questions that need to be explored. First, most of these studies used question–answer pairs or a series of spoken instructions which always required subjects to perform an additional task during discourse comprehension. Second, there is still no clear picture of the time course of accentuation effects in spoken language processing.

Given the issues mentioned above, the first aim of this study, then, was to establish the accentuation effect in natural spoken discourse comprehension without additional tasks. The second and main aim was to investigate the nature and time course of the interaction between accentuation and previous discourse context, which determines the information state of the accented word.

These research questions were investigated by using scalp-recorded event-related potentials (ERPs). The ERP component of interest is the N400, a negative-going potential following a semantically based stimulus event. The by-now classical N400 effect has a posterior distribution, with an onset around 200–250 msec after word presentation, and with a maximal amplitude between 380 and 440 msec (Hagoort & Brown, 2000; Kutas & Van Petten, 1994). Previous studies on the N400 have led to a fairly general consensus that, within the language domain, this ERP effect reflects some aspect of the processes that relate the meaning of a particular word to a higher-order semantic interpretation of the unfolding message (Hagoort & Brown, 2000; Van Berkum, Brown, & Hagoort, 1999; Kutas & Van Petten, 1994).

Although Kutas and Hillyard (1980) first described the N400 component using semantically anomalous words in sentence context, a series of studies showed that unrelated word pairs, incongruous sentence or discourse endings, and congruous but unexpected sentence endings all elicit larger N400 responses than, respectively, related word pairs, congruous sentence endings, and congruous but expected sentence endings. The N400 effect was not only found in reading but also in spoken language (Van Berkum, Zwitserlood, Hagoort, & Brown, 2003; Hagoort & Brown, 2000; Van Berkum et al., 1999). In addition, modulations of the N400 amplitude have been found for pictures and faces (Debruille, Pineda, & Renault, 1996; Nigam, Hoffman, & Simons, 1992). The N400 effect is thus not a simple reflection of semantic anomaly, but can be classified as an indicator of semantic processing in general. This study allowed us to establish whether the amplitude of N400 is increased when information is correct and highly salient compared to information that is incorrect while being less salient. This would establish that, indeed, the N400 effects are not triggered by semantic anomalies per se but by the semantic integration load that is usually, but not necessarily, higher for contextually inconsistent information. This was tested by manipulating the semantic consequences of prosody, in particular, accentuation.

If, during spoken language comprehension, accentuation influences semantic processing, then manipulating accentuation will result in N400 effects. This allows us to assess how the correspondence between accentuation and information structure affects spoken discourse comprehension and the N400. Because reliable ERP effects can be obtained in the absence of a potentially intrusive secondary task, this study investigates the role of accentuation during the natural process of spoken discourse comprehension. For this, we used discourses selected from journals, and the original texts were lightly edited.

The other noteworthy advantage of the ERP method is its high temporal resolution. This allows us to investigate the time course of the accentuation effect in spoken discourse comprehension, and the interaction between accentuation and information structure.

In summary, this study aimed to investigate the role of accentuation in natural spoken discourse comprehension

without a secondary task, and to investigate the nature and time course of accentuation in spoken discourse comprehension. In addition, the study aimed to further specify the processing nature of the N400 effect. To address those questions, native Chinese speakers were asked to listen to spoken discourses for comprehension. Meanwhile, their ERPs were registered. Every discourse consisted of three sentences. The critical word in the third sentence either conveyed new information or conveyed old information in relation to the preceding context. In addition, the critical word was either accented or deaccented.

As to the old information, we expected that it would elicit a larger N400 when it received pitch accent than when it did not, because the semantic information (i.e., information state) conveyed by pitch accent mismatches with that conveyed by preceding discourse context. As to the new information, there might be two possibilities. One possibility is that inconsistent accentuation (new information without pitch accent) elicits a larger N400 than consistent accentuation (new information with pitch accent) because of the mismatch between pitch accent and information structure. The other possibility is that the N400 amplitude to consistent accentuation is larger than the N400 evoked by inconsistent accentuation. The reason is that pitch accent indicates the presence of important information. For accented new information, a listener might allocate more cognitive resources to integrate the current word into discourse context, hence evoking a larger N400.

METHODS

Participants

Sixteen right-handed subjects (8 women) participated in the experiment; all of them were university students (Beijing Forestry University) and native speakers of Chinese. Mean age was 20 years (range 19–23). None of them had any neurological impairment, had experienced any neurological trauma, or used neuroleptics.

Stimuli

One hundred twenty discourses were selected from popular scientific journals and edited slightly to meet the requirements of this experiment. The discourses were spoken by a female speaker and recorded at a sampling rate of 22 kHz. Each discourse included three sentences. The first and the second sentences were ''context sentences.'' The third sentence was the target sentence, which included one critical word conveying new information and one critical word conveying old information. The first and the second critical words in each target sentence were separated by at least five syllables, with a total duration of 1000 msec or more. All critical words were double-character words. To avoid

critical words to be in the sentence-final position, there were at least four syllables between the second critical word and the end of sentence. For all critical words, the acoustic onset was determined by using speech software that allowed us to both visually inspect the sound spectrum and to listen to the sound repeatedly. Three researchers who were experienced in sound file analysis determined the acoustic onset together. Onset information was used to time-lock the ERP waveforms.

For each target sentence, two versions of the context sentences were constructed. In one version of the context, the first critical word was new and the second was old; in the other version, the first critical word was old and the second critical word was new. The target sentence had two kinds of accentuation pattern. In the first accentuation pattern, the first critical word in the target sentence was accented and the second critical word was deaccented. In the second accentuation pattern, the first critical word was deaccented and the second critical word was accented. Together, they realized a full factorial design with all combinations of the factors information state (new vs. old) and accentuation (accented vs. deaccented) (see Table 1 for an example of the materials).

In this study, we exploited the prosodic characteristics of the Chinese language. Chinese is a tone language. The pitch correlate of every syllable in Chinese is not a single point but a pitch contour which has a top and a bottom line and only the former has to do with the semantic accentuation. The acoustic correlates of accentuation are mainly the duration and the high point of the pitch contour of the critical word (CW) (Wang, Lü, & Yang, 2002; Zhong, Wang, & Yang, 2001).

To establish that our speaker had succeeded in correctly accenting the relevant critical words, we performed a t test on the spoken sentence materials to verify that the accented and deaccented critical words differed in the high point of the pitch contour and the duration of the critical words. The results showed that the duration of accented words was longer than that of deaccented words [$t(1, 239) = 42.21, p < .0001$], and the high point of the pitch contour of accented words was higher than that of deaccented words [$t(1, 239) = 12.87, p < .0001$]. Therefore, the acoustic measurements confirmed that target sentences were spoken with the intended accentuation pattern.

Discourses were grouped into four lists based on the four experimental conditions, resulting from the combination of the variables information state and accentuation. In each list, each spoken version of discourse and target sentence occurred only once and there were an equal number of discourses in each of the four conditions. Across the four lists, each discourse appeared once in all of the four conditions. In addition to 120 experimental discourses, in every list there were 120 filler discourses. These were also made up of three sentences, and were added to avoid that subjects discovered the

Table 1. Illustration for Four Conditions (Accentuation × Information State)

First word: new–accented; second word: old–deaccented

Context sentence: Now we all use toothbrushes to clean and protect our teeth.

It was recently found that primitive people also had their own tooth-cleaning tools.

Target sentence: They probably used (culms) as cleaning tools for their ''teeth''.

(___) ''   ''

First word: new–deaccented; second word: old–accented

Context sentence: Now we all use toothbrushes to clean and protect our teeth.

It was recently found that primitive people also had their own tooth-cleaning tools.

Target sentence: They probably used (culms) as cleaning tools for their ''teeth''.

(   ) ''___''

First word: old–accented; second word: new–deaccented

Context sentence: Now culms are the main materials to produce feedstuff for critter.

However, primitive people used culms as some kind of cleaning tools.

Target sentence: They probably used ''culms'' as cleaning tools for their (teeth).

''___'' (   )

First word: old–deaccented; second word: new–accented

Context sentence: Now culms are the main materials to produce feedstuff for critter.

However, primitive people used culms as some kind of cleaning tools.

Target sentence: They probably used ''culms'' as cleaning tools for their (teeth).

''   '' (___)

Parentheses indicate new information; quotes indicate old information; Underline indicates accentuation.

nature of the experimental manipulation and developed processing strategies.

Procedure

After the electrodes were positioned, subjects received instructions regarding the experimental task and then listened to the stimuli over headphones. They 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 movements and they should fixate on an asterisk displayed on the screen before them. Subjects were asked to listen to each discourse for comprehension. To ensure that the subjects indeed listened to the discourse for comprehension, at the end of 60 filler discourses, they were asked to answer a question regarding the content of the current discourse.

Each trial consisted of a 300-msec auditory warning tone, followed by 700 msec of silence, the spoken discourse context, 1000 msec of silence, and the spoken target sentence. To inform subjects when to fixate and sit still for EEG recording, an asterisk was displayed from 1000 msec before onset of the target sentence to 1000 msec after its offset. After a short practice session consisting of eight discourses, the trials were presented in six blocks (each block consisting of 40 discourses) of about 13 min each, separated by brief resting periods.

The EEG was recorded using DC amplifiers (Neuroscan). An elastic cap, equipped with 64 Ag/AgCl electrodes according to the International 10–20 system, was fixed to the participant's scalp. The average of left and right mastoid electrodes served as reference, and a forehead electrode served as the ground. Vertical eye movements and blinks were monitored via a supra- to suborbital bipolar montage. A right to left canthal bipolar montage was used to monitor horizontal eye movements. The EEG and electrooculogram (EOG) recordings were amplified using a high cutoff of 40 Hz. All electrode impedances were kept below 5 kΩ. The EEG and EOG signals were digitized on-line with a sampling frequency of 500 Hz, screened off-line for eye movements, muscle artifacts, electrode drifting, and amplifier blocking in a critical window that ranged from 200 msec before to 1200 msec after acoustic onset of the critical word. Averaging and artifact rejection were performed off-line.

RESULTS

For each subject, average waveforms for the critical words (CWs) were computed across all trials per condition after normalizing the waveforms of the individual trials on the basis of a 200-msec pre-CW baseline. As can be seen in Figure 1, relative to deaccented old CWs, the other three conditions (accented old CWs, deaccented new CWs, accented new CWs) all elicited a large negative deflection that emerged at approximately 200 msec after their acoustic onset, and lasted until approximately 700 msec, with the maximal amplitude occurring between 400 and 500 msec. This negativity was larger over posterior than anterior sites. We classified the negative deflections as N400 effects because their latency and topography fit the standard characteristics of an N400 effect (see Figure 2). Relative to the old-deaccented CWs, all other conditions showed an N400 effect with the same topography.
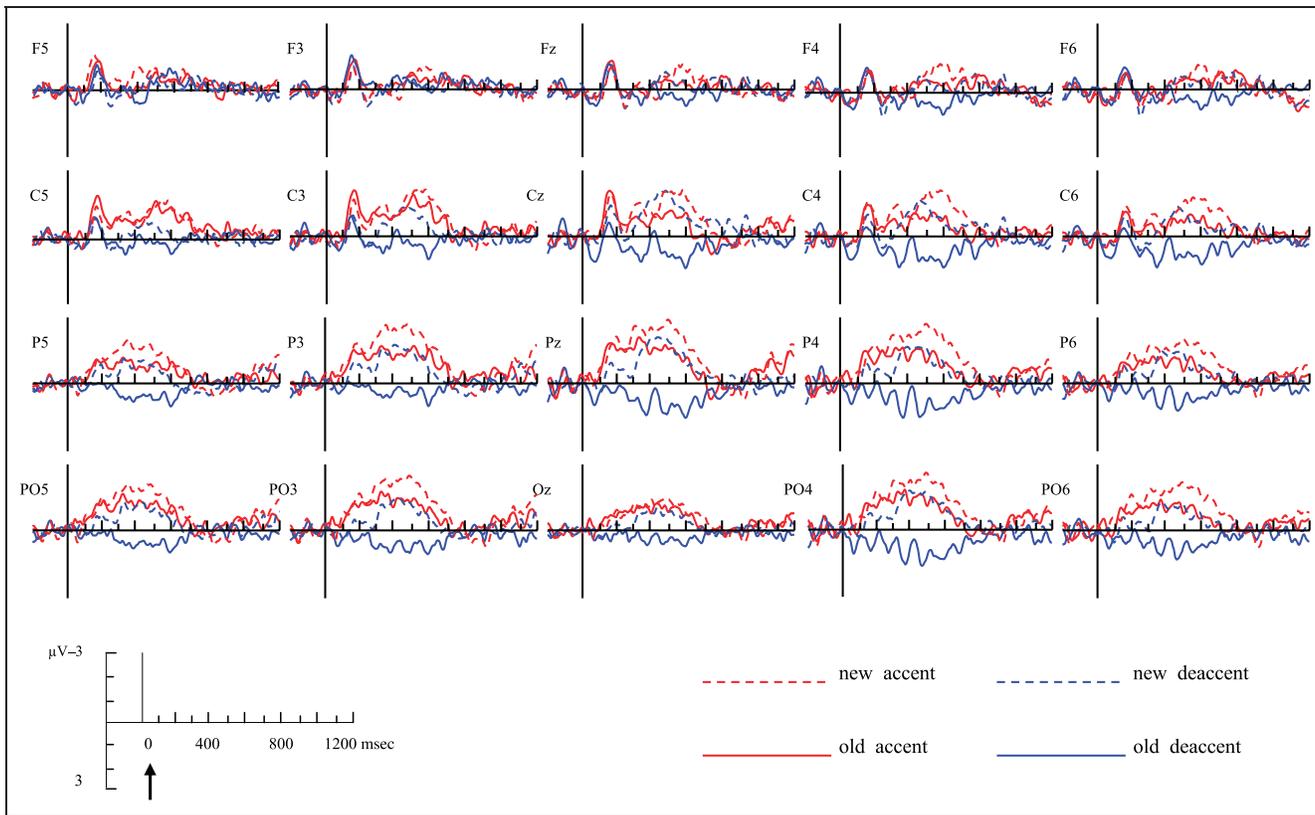
Figure 1. Accentuation and information state effects in spoken discourses.

Figure 2. Topography of the ERP effects of accentuation and information state in spoken discourses.
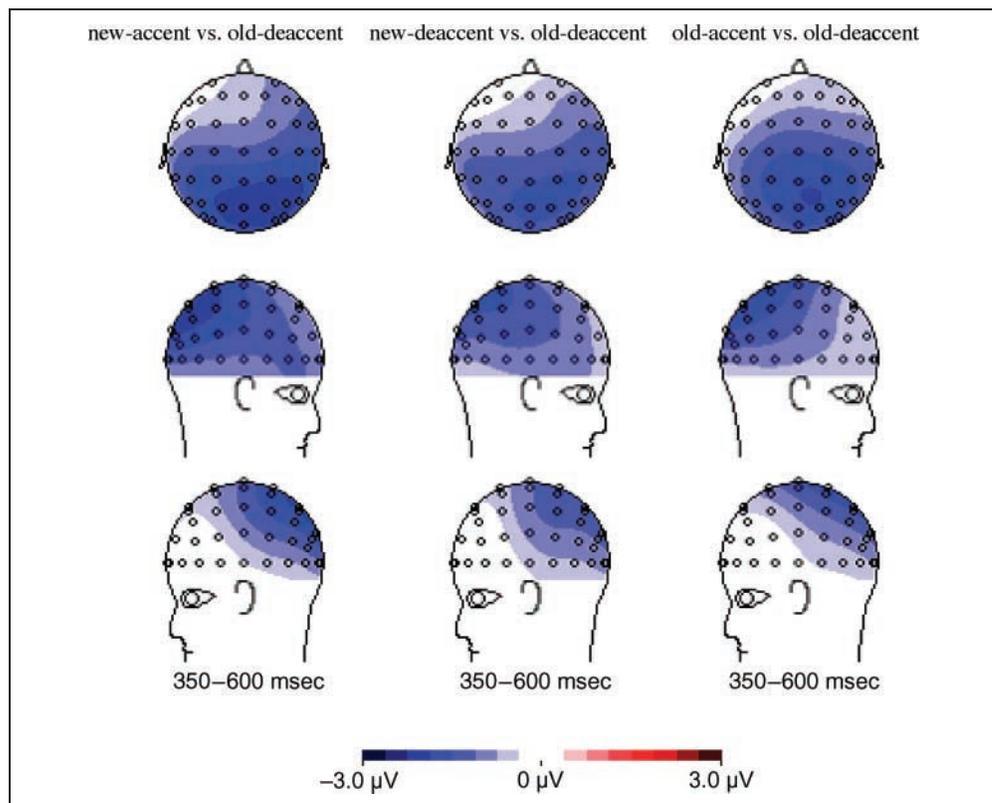
Statistical analyses were done on the mean amplitude in the 350–600 msec latency range following the acoustic onset of the CWs. Analyses of variance (ANOVAs) were conducted on a selection of midline electrodes and lateral electrodes, respectively. For the midline electrode sites, the mean amplitude values were entered into repeated-measures ANOVAs with accentuation (accented, deaccented), information state (old, new), and electrodes (Fz, Cz, Pz, Oz) as independent factors. For lateral electrodes, the mean amplitude values were entered into repeated measures ANOVAs with hemisphere (left, right) as an additional factor and lateral electrodes (F5/F6; F3/F4; C5/C6; C3/C4; P5/P6; P3/P4; PO5/PO6; PO3/PO4) nested under hemisphere. When the degree of freedom in the numerator was larger than one, the Greenhouse–Geisser correction was applied.

As can be seen in Table 2, the ANOVAs revealed a significant main effect of accentuation, indicating that the N400 was larger for accented than for deaccented words. In addition, a significant main effect of information state was found, due to the fact that new information elicited a larger N400 than old information. The ANOVAs also revealed a significant Accentuation by Electrodes interaction, due to the posterior distribution of the N400 effect. Importantly, the Accentuation by Information state interaction was found to be significant as well. This was caused by a different effect of accentuation for old and new information. A planned comparison showed that for old information, accentuation (inconsistent accentuation) elicited a larger N400 than deaccentuation (consistent accentuation). For new information, there was no significant

Table 2. Effects of Accentuation and Information State in Spoken Discourse Comprehension

Source	Midline			Lateral		
	df	F	p	df	F	p
A	1, 15	9.69	<.01	1, 15	16.62	<.001
I	1, 15	20.20	<.0001	1, 15	14.68	<.005
E	3, 45	10.22	<.0001	7, 105	9.08	<.0001
A × I	1, 15	14.48	<.005	1, 15	7.93	<.05
A (on O)	1, 15	22.10	<.0001	1, 15	29.09	<.0001
A (on N)	1, 15	0.72	.410	1, 15	3.64	.076
A × E	3, 45	10.80	<.001	7, 105	12.79	<.0001
A × I × E	3, 45	2.85	.076	7, 105	3.67	<.05
A × I × H				1, 15	2.27	.152
A × I × H × E				7, 105	0.87	.458

ANOVAs were based on the mean amplitude in 350–600 msec latency range. It included the following experimental variables: accentuation (A, accentuation vs. deaccentuation), information state (I, new vs. old), electrodes (E), and hemisphere (H). In addition, O indicates old information, and N indicates new information.

difference between accentuation and deaccentuation for the midline electrodes. However, for the lateral electrode sites, accentuation (consistent accentuation) elicited a larger N400 than deaccentuation (inconsistent accentuation). This difference was marginally significant. Note that, in this case, the N400 seems larger for the contextually appropriate prosodic marking of information as new compared to the absence of such a prosodic marking, which could be seen as incorrect.

Additional ANOVAs were conducted on the mean amplitude values in the 200–350 msec and 600–750 msec latency ranges. In the early 200–350 msec range, a main effect of accentuation was observed for both midline and lateral electrodes [$F(1, 15) = 9.37, p < .005; F(1, 15) = 15.95, p < .001$, respectively]. Moreover, a main effect of information state was observed for the midline electrodes [$F(1, 15) = 4.64, p < .05$]. A similar picture emerged for the 600–750 msec latency range, with a significant main effect of accentuation for midline and lateral electrodes [$F(1, 15) = 3.84, p < .069; F(1, 15) = 13.62, p < .005$, respectively], and a main effect of information state for the midline electrodes [$F(1, 15) = 7.47, p < .05$]. However, importantly in the early and late latency windows, no significant Accentuation by Information state interaction was observed.

To establish the onsets of the effects, we conducted a series of onset analyses in consecutive mean amplitude latency bins of 10 msec wide (e.g., 100–110 msec, 110–120 msec, etc.). Accentuation effects started in 230–240 msec latency bin for the midline electrodes and in the 120–130 msec latency bin for the lateral electrodes, followed by a long and uninterrupted series of bins with a main effect for accentuation. The main effects of information state started later than the accentuation effect. We observed a long and uninterrupted series of information state effects starting in the 290–300 msec latency bin for the midline electrodes, and in the 300–310 msec latency range for the lateral electrodes. Finally, the interaction between accentuation and information state became first significant in the 320–330 msec latency bin for the midline electrodes, and in the 350–360 msec latency range for the lateral electrodes.

Consistent and Inconsistent Accentuation

In the above section, we analyzed the effect of accentuation on new and old information, respectively. We found that for old information, accentuation (inappropriate) elicited a larger N400 relative to deaccentuation (appropriate); for new information, even though accentuation (appropriate) still elicited a larger N400 than deaccentuation (inappropriate), the difference between accentuation and deaccentuation became smaller. In order to examine the overall effect of correspondence between accentuation and information state, in the subsequent analysis, we combined accented new information and deaccented old information as appropriate
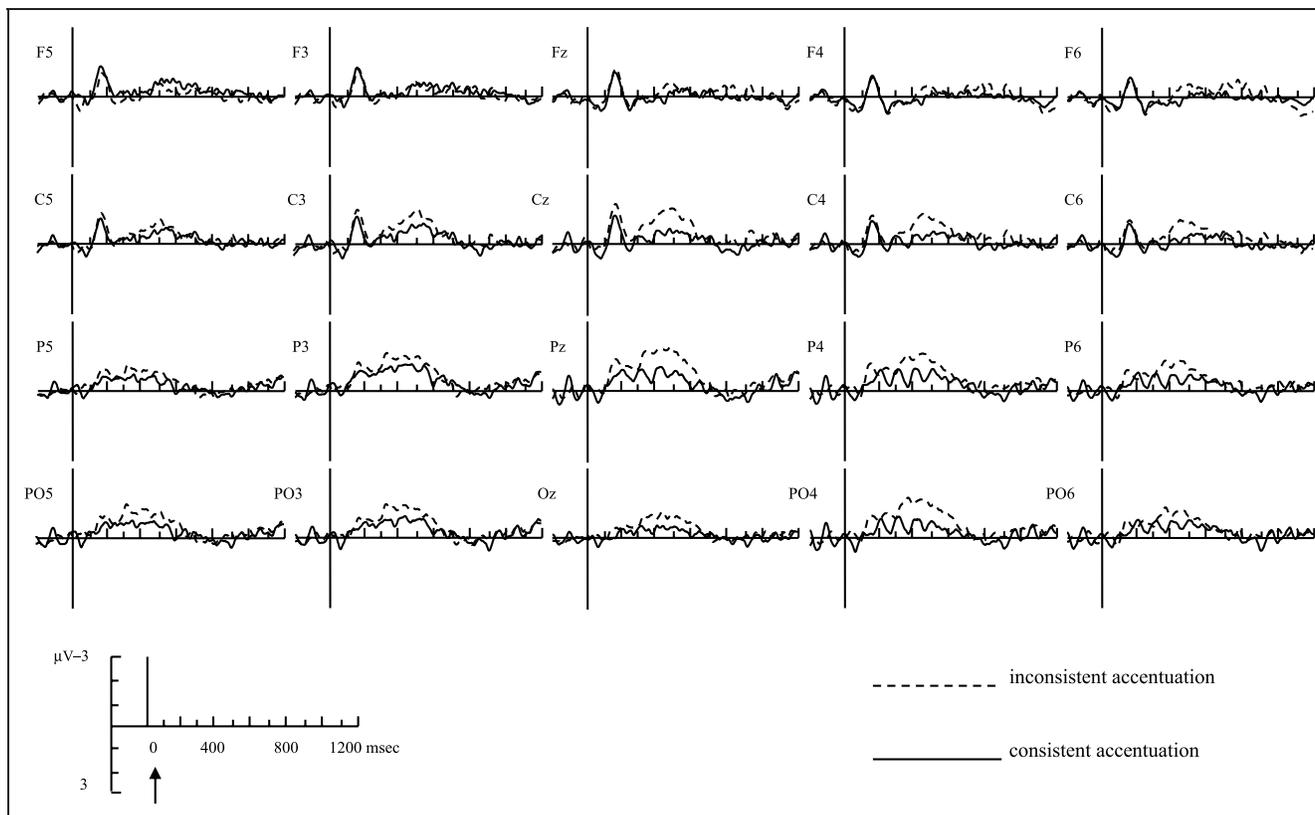
Figure 3. Correspondence between accentuation and information state effects in spoken discourses.

accentuation, and deaccented new information and accented old information as inappropriate accentuation.

As Figure 3 shows, relative to appropriate accentuation, inappropriate accentuation elicited a larger N400. As in the previous section, we entered the mean amplitude in 350–600 msec latency range into repeated measures ANOVAs with accentuation (appropriate; inappropriate) and midline electrodes, or lateral electrodes and hemisphere as independent factors. The ANOVAs revealed main effects of accentuation, indicating that inappropriate accentuation elicited a larger N400 than appropriate accentuation (see Table 3). We also observed a significant interaction between accentuation and electrodes due to the centro-posterior distribution of the effect. The ANOVAs revealed no differential left or right hemisphere contribution.

DISCUSSION

This experiment was designed to investigate the role of accentuation in spoken discourse comprehension, in particular, how and when accentuation interacts with information state. It was found that not only variation of information state but also variation of accentuation influenced the amplitude of the N400, with, in addition, a significant interaction between accentuation and information state. Accentuation had a different effect on the processing of old and new information.

These results indicate that listeners could rapidly extract the information from the acoustic parameters and relate it to the previous discourse context. Moreover, the accentuation effect came before the effect of information state. Shortly after the occurrence of the information state effect, accentuation and information state were integrated. Listener began to integrate accentuation and information structure when only about half of the accented double-character word was heard. In the subsequent section, we discuss those results in more detail.

Table 3. Effects of Accentuation Type (Relative to Information State) in Spoken Discourse Comprehension

Source	Midline			Lateral		
	df	F	p	df	F	p
A	1, 15	14.47	<.005	1, 15	7.70	<.05
E	3, 45	10.17	<.001	7, 105	9.05	<.0001
A × E	3, 45	2.90	<.05	7, 105	3.69	<.001
A × H				1, 15	2.41	.142
A × H × E				7, 105	0.89	.520

ANOVAs were based on the mean amplitude in 350–600 msec latency range. It included the following experimental variables: accentuation type (A, consistent vs. inconsistent accentuation), electrodes (E), and hemisphere (H).

Effects of Information State and Accentuation

The current experiment revealed that relative to old information, new information elicited a larger N400. The reason might be that listeners allocated more cognitive resources to the processing of new information for the sake of integrating this information into the context. Previous studies already found that information structure is one of the main factors controlling the listener's attention in language comprehension. For example, studies by Hornby (1974) and Just and Clark (1973) showed that it was more difficult to verify given information than new information. Donselaar and Lentz study also found that judgment latencies for new words were shorter than for given words (Donselaar & Lentz, 1994). Therefore, the results of this experiment are in line with those previous results, indicating that subjects allocated more attention to the new information. However, an alternative conclusion cannot be excluded, that is, the old–new effect is based on the repetition of the old words because they were presented in the first sentence of the discourse context.

The results of this experiment showed that, compared with deaccented information, accented information elicited a larger N400. Presumably, the accent signals that more semantic integration effort should be devoted to the accented word and its meaning.

Interaction between Accentuation and Information State

This experiment revealed that there was a significant interaction between accentuation and information state. For old information, accentuation (inappropriate) elicited larger N400 than deaccentuation (appropriate). For new information, accentuation still elicited a larger N400 than deaccentuation, even though accentuation was appropriate and deaccentuation was inappropriate. But the N400 amplitude difference between accented and deaccented old information was larger than that between accented and deaccented new information.

The results indicated that it was not simply a mismatch between accentuation and information structure that elicited the N400 effect. If this was true, for new information deaccentuation should elicited a larger N400 relative to accentuation, because deaccentuation was inappropriate. We think that it might be related to the role of accentuation in attention allocation. Because accentuation usually indicates the advent of important information, it can cause the listeners to allocate more integration resources to the processing of accented information. Accented information would thus evoke deeper semantic processing, which could result in the increase of the N400 amplitude. Therefore, for new information, accentuation elicited a larger N400 than deaccentuation. The current result is consistent with previous studies. There is ample psycholinguistic evidence indicating that pitch accent may function to draw attention to the information that is in focus (Frazier, 1999; Cutler, 1976). Numerous studies also showed that the N400 is sensitive to the amount of attention directed toward semantic processing (Hohlfeld & Sommer, 2005; Chwilla, Brown, & Hagoort, 1995; Holcomb, 1988). Nevertheless, in future experiments, it is still important to directly examine the role of pitch accent in conveying semantic meaning and in modulating attention allocation.

However, for old information, the deeper semantic processing was most likely not the only reason of the larger N400 in the accentuation condition. Accentuation on old information not only could attract more cognitive or integration resources but also mismatched with the information state. Accentuation inappropriately placed on old information was conflicting with its information state, hence, resulting in difficulty in integrating the current information into the discourse context. The larger N400 amplitude difference for old compared to new information might come from the combined effects of attention allocation and mismatch.

Recently, a study by Li and Yang (2004) indeed revealed that accentuation placed on new information enhanced the activation of new information in the discourse representation. Hence, the most important information gained a privileged status in listeners' working memory. Accentuation placed on old information enhanced the activation of old information, but suppressed the activation of new information. Those effects took place at about 1000 msec after the acoustic offset of the accented new information and 500 msec after the acoustic offset of the accented old information.

When examining the effect of appropriate accentuation relative to inappropriate accentuation, we found that relative to inappropriate accentuation, appropriate accentuation elicited a smaller N400. This suggested that accentuation that is appropriately placed can ease the semantic processing of the corresponding information and its integration within the discourse context. Previous studies had already established that prosody, especially accentuation, influenced language processing. We have now established that those results can be generalized to our materials (relatively large-scale discourse instead of question–answer pairs or spoken instructions) and task (just listening for comprehension rather than making a specific judgment).

Not only did accentuation play an important role in spoken discourse comprehension, the accentuation effect also happened very quickly. A striking aspect of the present results is what they reveal about the speed at which accentuation takes effect and interacts with information state. The critical words in this experiment were all double-character words which took about 600 msec to unfold. The results of the present experiment revealed that the accentuation effect occurred first, starting at no more than about 120–130 msec after the acoustic onset of critical words. After the accentuation effect, the effect of

information state became visible in the waveforms, which was followed by the interaction between accentuation and information state at about 300–310 msec after the acoustic onset of these words. Thus, during spoken discourse comprehension, accentuation not only comes into play very rapidly but also can do so before the first character of a double-character word has been fully heard and before the listener has fully identified the information state of the word on the basis of the acoustic information alone. When listening to spoken language, a variety of cues are extracted from the input, including syntactic, semantic, and phonological cues (Webera, Griceb, & Crocker, 2006; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Accentuation belongs to the level of phonological cues, which can influence the semantic processing of the current discourse information very quickly.

The Processing Nature of the N400

Finally, our results provide important new information about the nature of the N400 effect. Within the language domain, one way in which N400 effects are elicited is by a semantic mismatch between the meaning of a word and the semantic specification of its context (Van Berkum et al., 2003; Kutas, 1993). However, a semantic anomaly is not required to elicit an N400 effect. For example, Van Berkum, Kooijman, Brown, Zwitserlood, and Hagoort (2002) found that a coherent spoken word in sentence context, but relatively unpredictable in the discourse, elicited a larger N400 than a much more predictable coherent word. The results of this experiment go beyond establishing that the N400 ERP effect is not a simple reflection of a semantic anomaly. In fact, for the new information, accentuation is semantically correct, whereas deaccenting new information could be claimed to be anomalous. Nevertheless, the correct condition (the accented new information) elicited a larger N400 than the anomalous condition (deaccented new information). The accented and deaccented new information had the same lexical meaning and the same discourse context. The difference in amplitude could thus not be triggered by a difference in context or lexical meaning. More importantly, it could also not be triggered by a contextual anomaly per se. Instead, the slightly larger amplitude for the correctly accented new information suggests that the N400 is sensitive to the information value of the word in its context. The accent signals that the current word is relevant and new information, and this might recruit additional integration resources to integrate the current lexical information into the overall discourse model.

Acknowledgments

REFERENCES

Altmann, G. T. M., van Nice, K. Y., Garnham, A., & Henstra, J. A. (1998). Late closure in context. *Journal of Memory and Language, 38,* 459–484.

Birch, S., & Garnsey, S. M. (1995). The effect of focus on memory for words in sentences. *Journal of Memory and Language, 34,* 232–267.

Bock, J. K., & Mazzella, J. R. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory & Cognition, 11,* 64–76.

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32,* 274–285.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics, 20,* 55–60.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language, 47,* 292–314.

Debruille, J. B., Pineda, J., & Renault, B. (1996). N400-like potentials elicited by faces and knowledge inhibition. *Cognitive Brain Research, 4,* 133–144.

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? *Language and Speech, 37,* 375–391.

Frazier, L. (1999). *On sentence interpretation (Studies in theoretical psycholinguistics 22)*. Dordrecht: Kluwer Academic Publishers.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia, 38,* 1531–1549.

Hohlfeld, A., & Sommer, W. (2005). Semantic processing of unattended meaning is modulated by additional task load: Evidence from electrophysiology. *Cognitive Brain Research, 24,* 500–512.

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language, 35,* 66–85.

Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 194,* 286–300.

Hornby, P. A. (1974). Surface structure and presupposition. *Journal of Verbal Learning and Verbal Behavior, 13,* 530–538.

Hruska, C., Steinhauer, K., Alter, K., & Steube, A. (2000). *ERP effects of sentence accents and violations of the information structure*. In Poster presented at the 13th annual CUNY conference on human sentence processing, San Diego, CA.

Just, M. A., & Clark, H. H. (1973). Drawing inferences form the presuppositions and implications of affirmative and negative sentences. *Journal of Verbal Learning and Verbal Behavior, 12,* 21–31.

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes, 8,* 533–572.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–205.

Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego: Academic Press.

Li, X. Q., & Yang, Y. F. (2004). The role of accentuation in spoken discourse comprehension. *Acta Psychologica Sinica, 36,* 393–399.

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience, 4,* 15–22.

Terken, J., & Noteboom, S. D. (1987). Opposite effects of accentuation and deaccentuation on verification. Latencies for given and new information. *Language and Cognitive Processes, 2,* 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. *Journal of Memory and Language, 41,* 147–182.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 443–467.

Van Berkum, J. J. A., Kooijman, V., Brown, C. M., Zwitserlood, P., & Hagoort, P. (2002). *Do listeners use discourse level information to predict upcoming words in an unfolding sentence? An ERP study*. In 9th Annual Meeting of the Cognitive Neuroscience Society (CNS-2002), San Francisco, April 14–16.

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? *Cognitive Brain Research, 17,* 701–718.

Wang, B. (2002). *The research on perception of prosody in Mandarin*. Doctoral thesis, pp. 68–74.

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

Webera, A., Griceb, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition, 99,* 63–72.

Zhang, Y. F. (2002). Focus in Chinese and the de sentence. *Chinese Language Learning, 3,* 24–30.

Zhong, X. B., Wang, B., & Yang, Y. F. (2001). The perception of prosodic word stress in standard Chinese. *Acta Psychologica Sinica, 33,* 481–488.

Zhu, Y. S., & Miao, X. W. (2000). Discourse functions of pragmatic presupposition. *Journal of Foreign Language, 3,* 25–30.