Semantic composition engenders an N400: evidence from Chinese compounds

Chen Bai[a,b], Ina Bornkessel-Schlesewsky[a], Luming Wang[a], Yu-Chen Hung[b], Matthias Schlesewsky[b,c] and Petra Burkhardt[b]

[a]Independent Junior Research Group Neurotypology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, [b]Department of Germanic Linguistics, University of Marburg, Marburg and [c]Department of General and Comparative Linguistics, Johannes Gutenberg University Mainz, Mainz, Germany

Correspondence to Dr. Petra Burkhardt, Department of Germanic Linguistics, University of Marburg, Wilhelm-Röpke-Strasse 6a, D-35032 Marburg, Germany
Tel: + 49 6421 2824645; fax: + 49 6421 2824558; e-mail: petra.burkhardt@staff.uni-marburg.de

This study provides evidence for the role of semantic composition in compound word processing. We examined the online processing of isolated two meaning unit compounds in Chinese, a language that uses compounding to 'disambiguate' meaning. Using auditory presentation, we manipulated the semantic meaning and syntactic category of the two meaning units forming a compound. Event-related brain potential-recordings revealed a significant influence of semantic information, which was reflected in an N400 signature for compounds whose meaning differed from the constituent meanings. This finding suggests that the combination of distinct constituent meanings to form an overall compound meaning consumes processing resources. By contrast, no comparable difference was observed based on syntactic category information. Our findings indicate that combinatory semantic processing at the word level correlates with N400 effects. *NeuroReport* 19:695–699 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: auditory, Chinese, compounds, language processing, left-anterior negativity, N400, P600, prediction

Introduction

Combinations of words are very common in human languages and reflect the need to express a large number of concepts during everyday conversation. Recent neurophysiological findings from German suggest that the online composition (here understood in terms of the 'construction of meaning') of 'compound' words consumes processing resources [1]. Specifically, when the overall compound meaning can be derived from the combination of its individual constituent meanings (e.g. *wineglass*), the construction of this meaning engenders an increased N400 event-related brain potential (ERP) in comparison with compound words with nonderivable meanings that must be stored in the mental lexicon (e.g. *butterfly*). These existing results, however, can be interpreted in several ways. On the one hand, the N400 effect could reflect increased costs of semantic composition because the individual constituent meanings must be integrated with one another to derive the overall compound meaning. Alternatively, it might be the case that the combination of constituent units *per se* is costly, without a specific contribution of constituent meaning.

This study aimed to provide evidence for a semantic composition-based account. To this end, we capitalized upon the properties of compound word formation in Chinese. In contrast to European languages, like German or English, in which compounding serves to derive new word meanings, Chinese uses compounding to *disambiguate* meaning. Here, the characters that are the basic units of meaning often have many homophones (i.e. meaning units

sound identical, but have different meanings). For example, the following characters – 房 ('room'), 防 ('defend'), 妨 ('disturb') – are pronounced identically as /fang1/, which results in ambiguity during auditory processing. (Note that numerals in phonetic representations are used to indicate lexical tones.) Accordingly, the meaning of 房 ('room') can be disambiguated and identified immediately when this meaning unit combines with another /jian1/ 间 ('space') to form the compound 房 房间 ('room') (pronounced as /fang1jian1/). As a result, approximately 70% of Chinese words are compounds consisting of two meaning units [2].

If speakers of Chinese indeed use a strategy in which the second meaning unit of a two-constituent compound is expected to confirm the meaning of the first, the processing of the first meaning unit should engender predictions as to the form and meaning of the compound as a whole. (Note that the presence of a second meaning unit can be anticipated via prosodic information [3]). From a semantic perspective, the preactivation of lexical-semantic features by the first meaning unit should favor compounds whose constituents share a significant number of lexical-semantic features. Thus, compounds whose constituents are semantically distinct should evoke an increased N400, as the amplitude of the N400 is generally larger when the parser faces a semantic incongruity (see [4,5] for evidence from the visual and auditory modality). In addition, semantic similarity is also predicted to decrease the N400 amplitude [6].

The hypothesized effects of semantic prediction were contrasted with possible effects of syntactic prediction. In this regard, we examined the processing of compounds whose overall category corresponded to those of its constituent parts (e.g. two nouns combining to form a noun compound) or not (e.g. a verb and a noun combining to form a noun). This type of syntactic incongruity may be expected to engender a qualitatively different electrophysiological response to that for semantic incongruity. In analogy to findings for syntactic mismatches at the sentence level, this type of word-level syntactic incongruity might engender a more pronounced P600 [7,8].

In sum, an enhanced N400 was hypothesized for the violation of semantic expectations and a P600 effect for the violation of syntactic category expectations. To test these predictions for auditory compound processing in Chinese, we manipulated the semantic meaning (same vs. different) and the syntactic category (same vs. different) of Chinese compounds made up of two meaning units. This yielded four conditions as exemplified in Table 1. For condition 1, the meaning of the Chinese compound obtains a meaning similar to that of its constituents and it is of the same word category as its constituent parts. Thus, condition 1 is perfectly in tune with the two different predictions and should not result in increased computational effort. For condition 2, the meaning of the Chinese compound is unrelated or less related to the meaning of its first constituent. This represents a violation of the semantic expectation and should therefore evoke a more pronounced N400. The syntactic predictions are met, as all items represent combinations of two noun constituents. The two meaning units in condition 3 share similar meanings, but at least the first or both constituents represent units commonly used as verb or adjective. The identification of the compound thus involves a change in syntactic category, here hypothesized to engender a P600. Finally in condition 4, the meaning of the compound differs from the individual constituent meanings and different syntactic categories are combined to form a noun compound. In this case, both the semantic and the syntactic expectations are violated,

which are predicted to surface as enhanced N400 and P600 signatures, respectively.

Methods

Twenty-two native Chinese speakers (10 men; 19–30 years old; mean age=24.5 years) from the University of Marburg participated in the study after giving informed consent. All were right-handed, had normal or corrected-to-normal vision and good auditory acuity. After 20 training trials, each participant listened to 240 test words (120 real words and 120 nonsense words) distributed across four blocks of 60 trials each. The items in each block were pseudorandomized such that (i) each block contained an equal number of real and nonsense words, (ii) the target items in each block were evenly selected from the four critical conditions and (iii) not more than two items from the same condition were presented consecutively. The real word target items were assigned to the four test conditions (Table 1) that were manipulated by two experimental factors (semantic change, syntactic category change). All critical items represented commonly used nouns in modern Chinese and were high frequency compounds consisting of two meaning units (the overall frequency was 68 per million, 238 and 262 per million for the first and second meaning unit, respectively [2]). In addition, 120 nonsense compounds were constructed by randomly combining the second constituent with the first one from the target words (i.e. these items shared identical meaning units with the target words but their combination makes no sense in Chinese). If a nonsense word sounded similar to a real word, it was rejected and replaced by a new combination. All test items were presented auditorily. Participants were informed that after a fixation cross appeared in the middle of the computer monitor for 1 s, they would hear a Chinese compound. Their task was to indicate whether the compound represented a real word of Chinese or not by pressing a button on a response box as quickly and accurately as possible. They were asked to refrain from blinking from the presentation of the fixation cross until the presentation of three asterisks. After an intertrial interval of 3 s, the next trial started.

The electroencephalogram was recorded by means of 24 Ag/AgCl electrodes with a sampling rate of 250 Hz (impedances $< 4 k\Omega$) and was referenced to the left mastoid (rereferenced to linked mastoids offline). The horizontal electrooculogram was monitored with two electrodes placed at the outer canthus of each eye and the vertical electrooculogram with two electrodes above and below the left eye. The ground electrode was placed at C2. Only trials with correct answers in the lexical decision task and without artifacts entered the ERP analysis (90% of all trials). Single subject averages were computed relative to the onset of the first and second meaning unit of each compound, respectively. Grand average ERPs were computed per condition over all participants. A repeated-measures analysis of variance (ANOVA) with the factor COMPOUND first compared all real compounds to nonsense compounds time-locked to the onset of the first and second meaning unit (mean length of compounds: 804 ms, mean length of first constituent: 369 ms). This comparison should reflect the time-course of compound identification and show no differences at the first meaning unit. Second, an ANOVA involving the factors SEM (semantic change: yes/no) and SYN (syntactic category change: yes/no) were calculated for

Table 1 Examples of the four experimental conditions, which varied with respect to semantic change [yes(+)/no(−)] and syntactic category change [yes(+)/no(−)]

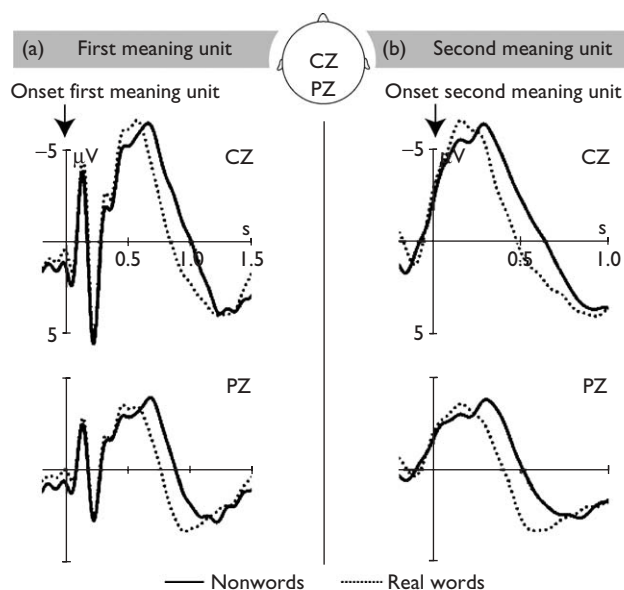| Conditions	Examples	Semantic change	Syntactic category change
1	声音 (sound + tone → sound) Noun + noun → noun	−	−
2	眼镜 (eye + mirror → glasses) Noun + noun → noun	+	−
3	开关 (fight + contest → war, fighting) Verb + verb → noun	−	+
4	战争 (open + close → switch) Verb + verb → noun	+	+
Nonword	声争 (sound + contest → nonsense)	NA	NA

For instance, in condition 1 the compounds and their constituents do not differ in meaning [no semantic change (−)], that is, 'sound' and 'tone' are semantically related, and they are both nouns that combine to form a noun [no syntactic category change (−)]. The last line presents an example of the nonword condition, where the two experimental manipulations are not applicable (NA).

mean amplitude values per time window per condition. The analyses involved the topographical factor 'region of interest' (ROI). Lateral regions of interest were defined as follows: left-anterior (F3, F7, FC1, FC5); left-posterior (CP1, CP5, P3, P7); right-anterior (F4, F8, FC2, FC6) and right-posterior (CP2, CP6, P4, P8). Midline analyses are not reported here for reasons of space, but confirmed the lateral analyses. Analyses were carried out in predetermined time ranges, 300–600 ms for the semantic prediction and 700–1000 ms for the syntactic prediction.

Results

For the behavioral data, participants performed at ceiling level in the lexical decision task (96.13% accuracy over all items). Figure 1 presents the grand average ERPs for real and nonsense compounds at the first and second meaning unit (left and right panel, respectively). As is evident from Fig. 1, there were no differences between the two conditions at the first meaning unit (which ended at approximately 369 ms). Time-locked to the onset of the second meaning unit, the ERPs for nonsense words showed a significant negativity compared with the real compounds in the time window between 300–600 ms [F(1,21)=94.4, $p < 0.01$] and an interaction of COMPOUND × ROI [F(3,63)=11.41, $p < 0.01$], which was highly significant over all four ROIs. In addition, there was a significant negativity between 100 and 200 ms for real compounds [F(1,21)=5.74, $p < 0.03$] and an interaction of COMPOUND × ROI [F(3,63)=3.53, $p < 0.02$], which was reliable over the right-anterior [F(1,21)=8.59, $p < 0.01$] and right-posterior ROIs [F(1,21)=5.49, $p < 0.03$].

The 2 × 2 (SEM × SYN) ANOVA was performed in the time windows from 300 to 600 ms and from 700 to 1000 ms of the second meaning unit. A main effect of SEM [F(1,21)=27.33, $p < 0.01$] and an interaction of SEM × ROI [F(3,63)=14.13, $p < .01$] were shown in the time window

from 300 to 600 ms. This interaction reached significance in all four ROIs {left-anterior [F(1,21)=6.93, $p < 0.01$], left-posterior [F(1,21)=50.31, $p < 0.01$], right-anterior [F(1,21)= 4.70, $p < 0.04$] and right-posterior [F(1,21)=43.49, $p < 0.01$]}. From 700 to 1000 ms, there was an interaction of SEM × ROI [F(3,63)=12.58, $p < 0.01$] that was resolved over the left-posterior [F(1,21)=9.70, $p < 0.01$] and right-posterior ROIs [F(1,21)=6.25, $p < 0.03$]. Figure 2 presents the grand average ERPs for the manipulation of semantic meaning change (Fig. 2a) and syntactic category change (Fig. 2b). Comparisons by the factor SEM revealed an enhanced negativity for words with meaning change compared with words with no meaning change. As all analyses were performed hierarchically, comparison by the factor SYN could not be resolved.

Discussion

To our knowledge, this study constitutes the first ERP investigation of the online auditory processing of isolated compounds comprised of two meaning units in Chinese. The comparison of real and nonsense compounds (Fig. 1) revealed that the identification of a real compound is accomplished more quickly than that of a nonsense compound. The early negativity (100–200 ms) for real compounds may reflect the access to a lexical representation. This assumption is in line with findings from ERP measures and eye fixations that illustrate that lexical access takes place within this temporal window [9]: on the basis of a number of empirical findings, the authors demonstrate that different factors that influence lexical access (e.g. frequency of occurrence, predictability) are observable in eye movement and ERP measures within the first 250 ms. Furthermore, differences in the length of the first meaning unit cannot account for this effect (words: 362 ms; nonsense words: 376 ms). The lexical search associated with the nonsense compounds is in turn reflected in a later negativity (300–600 ms). The primary concern of this study was the effect of predictive parsing, which we investigated through the window of semantic and syntactic expectations for the compound word (Fig. 2). The ERP data showed a significant influence of semantic – but not syntactic – information on the processing of compounds in Chinese. The more enhanced N400 for compounds that consist of semantically distinct meaning units indicates that the identification and composition of such a compound consumes processing resources.

This finding provides converging support for a semantic composition-based account of the costs of compound processing (e.g. as previously observed in [1]). Although all of our critical stimuli required the combination of two meaning units, only those compounds that necessitated a semantic reinterpretation at the position of the second constituent engendered an increased N400. This suggests that the N400 effect observed in the processing of compound words cannot be attributed to combinatory costs of compounding *per se*. Rather, this effect seems to reflect the semantic/interpretive aspects of this composition. An interpretation of our results along these lines is further strengthened by the absence of an effect for syntactic (category-based) composition costs.

We, however, cannot rule out at present that the absence of a syntactic prediction effect in Chinese compounds may constitute a language-specific effect, as the status of syntactic categories in Chinese is highly controversial. Most

Fig. I ERPs for compounds time-locked to the onset of the first (a) and second meaning unit (b) at two selected electrode sites CZ and PZ; real compounds (dotted) are compared with nonsense compounds (solid). Time span is charted horizontally starting at 200 ms before the onset (onset at vertical bar). Negative voltage is plotted upwards. Real compounds show negativity with an earlier onset latency at the second meaning unit.
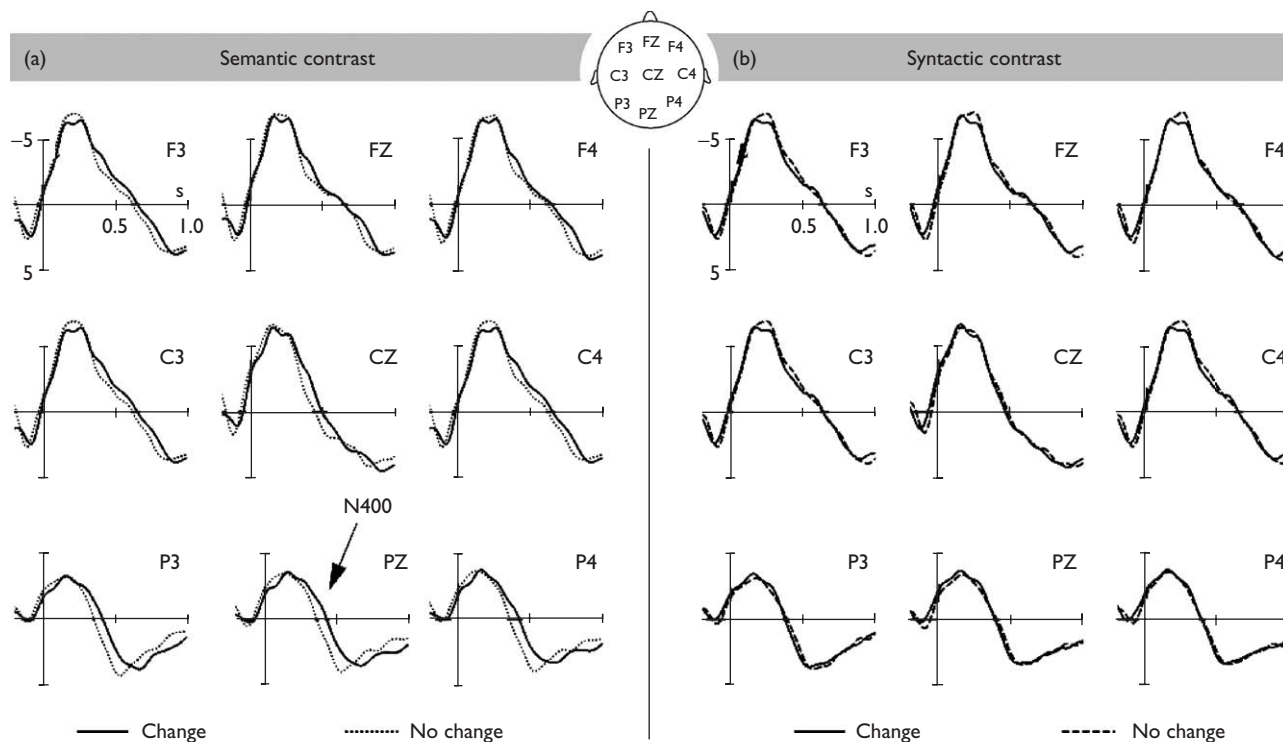
Fig. 2 Grand-average ERPs for semantic (a) and syntactic contrast (b) time-locked to the onset of the second meaning unit at nine selected electrodes. Conditions with change are plotted in solid lines. Time window ranges from 200 ms before to 1000 ms after the onset of the second meaning unit (onset at vertical bar). Negativity is plotted upwards.

Chinese grammarians suggest that Chinese words are not obligatorily linked to a predefined syntactic category (i.e. categorical ambiguity is much more widespread than in languages like English) [10,11]. Thus, this language-specific characteristic of words in syntactic categorization provides an alternative explanation for why no cost of syntactic category change occurred in this study. Future research must therefore determine whether this finding crucially hinges upon the specific properties of Chinese or whether it is in fact more general.

Finally, the observation that N400 effects can index processes of semantic composition sheds new light on the functional interpretation of the N400 component during language processing. Traditionally, combinatory (or supposedly 'rule-based') linguistic processes are associated with left-anterior negativities (e.g. [12,13]). The N400, by contrast, is typically interpreted as a correlate of semantic memory use [14]. From the perspective that the N400 in this study is modulated by the semantic complexity of the combinatory process, the functional distinction between this component and the left-anterior negativities no longer seems so clear-cut. This claim is not entirely without precedent in the neuro-cognitive literature on word processing: at least two previous studies observed N400-like effects for the processing of productive or (sub)regular morphological structure [15,16].

Conclusion

These data indicate that semantic composition effects in the processing of Chinese compound words engender an N400. By offering a unique baseline condition in which the second constituent of the compound substantiates the meaning of the first constituent, the specific properties of Chinese

allowed us to disentangle these combinatory semantic effects from possible costs of compound formation *per se*. In addition, the absence of a difference during syntactic composition implies that word category information does not play a critical role in Chinese compounding. On a more general level, the data suggest that the N400 component may index combinatory processes in language comprehension – a functional interpretation that is traditionally ascribed to left-anterior negativities.

References

1. Koester D, Gunter TC, Wagner S. The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPs. *Brain Lang* 2007; **102**:64–79.
2. Institute of Language Teaching and Research. *Modern Chinese frequency dictionary.* Beijing: Beijing Language Institute Press; 1986.
3. Isel F, Gunter TC, Friederici AD. Prosody-assisted head-driven access to spoken German compounds. *J Exp Psychol: Learn Mem Cogn* 2003; **29**: 277–288.
4. Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 1980; **207**:203–205.
5. Van Petten C, Coulson S, Rubin S, Plante E, Parks M. Time course of word identification and semantic integration in spoken language. *J Exp Psychol: Learn Mem Cogn* 1999; **25**:394–417.
6. Rugg MD. The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology* 1985; **22**:642–647.

7. Neville HJ, Nicol J, Barss A, Forster KI, Garrett MF. Syntactically based sentence processing classes: evidence from event-related brain potentials. *J Cogn Neurosci* 1991; **3**:151–165.

8. Osterhout L, Holcomb PJ, Swinney DA. Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *J Exp Psychol: Learn Mem Cogn* 1994; **20**:786–803.

9. Sereno SC, Rayner K. Measuring word recognition in reading: eye movements and event-related potentials. *Trends Cogn Sci* 2003; **7**:489–493.

10. Zhu D. *Xiandai Hanyu Yufa Yanjiu (Study on syntax of modern Chinese)*. Beijing: Commercial Press; 2001.

11. Hu M. *Cilei Wenti Kaocha (Research on Chinese lexical category)*. Beijing: Beijing Language Institute Press; 1996.

12. Friederici AD. Towards a neural basis of auditory sentence processing. *Trends Cogn Sci* 2002; **6**:78–84.

13. Ullman MT. A neurocognitive perspective on language: the declarative/procedural model. *Nat Rev Neurosci* 2001; **10**:717–726.

14. Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn Sci* 2000; **4**:463–470.

15. Janssen U, Wiese R, Schlesewsky M. Electrophysiological responses to violations of morphosyntactic and prosodic features in derived German nouns. *J Neuroling* 2006; **19**:466–482.

16. Bartke S, Rösler F, Streb J, Wiese R. An ERP-study of German 'irregular' morphology. *J Neuroling* 2005; **18**:29–55.