

Mental Representation of Verb Meaning: Behavioral and Electrophysiological Evidence

Xuesong Li^{1*}, Hua Shu¹, Youyi Liu¹, and Ping Li²

Abstract

■ Previous psycholinguistic research has debated the nature of the mental representation of verbs and the access of relevant verb information in sentence processing. In this study, we used behavioral and electrophysiological methods to examine the representation of verbs in and out of sentence contexts. In five experiments, word naming and event-related potential (ERP) components were used to measure the speed and the amplitude, respectively, associated with different verb–object combinations that result in different degrees of fit between the verb and its object. Both naming speed and ERP

amplitudes (N400) are proven to be sensitive indices of the degree of fit, varying as a function of how well the object fits the verb in terms of selectional restrictions. The results suggest that the semantic features of the verb's arguments are an integral part of the mental representation of verbs, and such information of the verb is accessed and used on-line during sentence processing. Implications of these results are discussed in light of recent computational semantic models that view the lexicon through high-order lexical co-occurrences in language use. ■

INTRODUCTION

The mental representation of the lexicon has been a central focus of psycholinguistic inquiry. A significant amount of work in the last decades has examined the nature and processes in the mental lexicon in a variety of languages. More recent approaches have used neuroimaging tools to tap into brain processes associated with the mental lexicon. Although a wide range of issues has been investigated in both behavioral and neuroimaging studies, researchers have not come to a consensus regarding the nature of the mental lexicon, in particular, on what information is represented in a lexical entry and how that information is retrieved and used during sentence comprehension. This study is designed to examine lexical representation during sentence processing. We present behavioral and electrophysiological evidence from Chinese, a language with lexical and grammatical properties distinct from Indo-European languages that have been the focus in past research (see Li, Tan, Bates, & Tzeng, 2006, for an overview).

The Role of Verbs in Language Processing

Verbs represent the core elements of linguistic constituents in a sentence, and they often stand at the center of linguistic theories; for example, in Fillmore's (1968) Case Grammar, verbs serve as the pivot to other nominal ar-

guments such as agents, patients, instruments, and so on (the so-called thematic roles of the verb arguments). It is thus no wonder that in many studies in this domain verbs play a central role in the investigation. Apart from the heavily investigated thematic roles that verbs impose on a sentence, verbs also place selectional restrictions (SRs) on the nominal arguments in the sentence (e.g., subject and object). In particular, morphologically impoverished languages such as Chinese and English emphasize the syntactic and semantic restrictions on the kinds of object that can occur after a verb. Some verbs require the presence of a direct object, often the patient or the receiver of the action, and the patient can be animate, inanimate, or both. Other verbs require two objects, one direct and one indirect, such as *give*. Still other verbs require not an object but a sentence complement to follow, such as *think* in English or *xiang* in Chinese.¹ Such SR information is part of the lexicon rather than syntax in many current psycholinguistic views and is represented in the mental lexicon in addition to the verb's general orthographic, phonological, and grammatical information (see review below). The inclusion of SR in the mental representation of verbs is also critical to the processing of sentences in which the verb occurs.

Previous Studies of Verb Representation

Behavioral Data

A number of studies have suggested that information on the verb's SR must be an integral part of the mental

*Current address: Southeastern University, Nanjing, China
¹Beijing Normal University, ²University of Richmond

representation of verbs,² and that such information will be automatically activated and used in processing whenever a verb is recognized in a sentence (e.g., Boland, 1993; Trueswell, Tanenhaus, & Kello, 1993; Shapiro, Zurif, & Grimshaw, 1987, 1989). Boland (1993) presented sentences that contain verbs such as *hit* (direct object required) and *insist* (sentence complement required). In a naming task, participants responded to probing words such as nominative *they* or accusative *them* after the verb (e.g., *the slow waitress hit them*). The results were that response times were significantly faster when the accusative case follows verbs like *hit* and when the nominative case follows verbs like *insist*, indicating that information about SR (specifically, the subcategorization frames) of the verb is activated when the verb is recognized. Trueswell et al. (1993) had similar findings from self-paced reading and eye-movement tasks. Participants showed slower reading times in both self-paced reading and eye movements when the complementizer *that* was absent from the sentence after a verb like *confirm*, which biases the following noun to be a direct object, for example, in *the waiter confirmed the reservation was made yesterday*. In contrast, participants showed no differences in reading times, irrespective of whether *that* was in the sentence complement after a verb like *insist*, which biases a sentence complement to follow (e.g., *the waiter insisted the reservation was made yesterday*). These results indicate that reading times for sentence complements with or without *that* were dependent on how frequently a sentence complement, relative to a direct object, occurs with the verb: The SR of a verb like *confirm* is biased more toward direct objects (e.g., *reservation*) than toward complements (e.g., *reservation was made*), whereas the reverse is true for a verb like *insist*.

In a similar vein, McRae, Spivey-Knowlton, and Tanenhaus (1998) examined the effects of SR in sentence processing. In a self-paced reading task, participants were presented with a subject noun (e.g., *crook*), followed by a verb in a reduced relative clause (e.g., *arrested by the detective*). The noun in the subject position varied in terms of the verb's SR (specifically thematic fit, e.g., *crook* is typically the patient of *arrest* and *cop* the agent of the action). The results indicated that when the noun took the patient role (e.g., *crook*), readers had no particular difficulty in reading the reduced relative clause. However, when the noun was biased to the agent role (e.g., *cop*), readers tended to interpret the first verb as the main verb of the sentence, leading to difficulty in interpreting the prepositional phrase following the verb (such as *by the detective*). These findings suggest that knowledge of thematic fit between the noun and the verb is automatically computed and used during on-line sentence processing.

Recently, Hare et al. (2003) examined the relationship between verb senses and the relative SR probabilities for subcategorization. They argue that not only the meaning

of a verb as a whole may be biased toward direct object versus sentence complement but different senses of the same verb also have different biases. For example, the verb *find*, when used to mean “locate” or “discover,” is more likely to occur with a direct object (e.g., *he found the book*), but when used to mean “realize” or “understand,” is more likely to occur with a sentence complement (e.g., *he found the book was written poorly*). Hare et al. presented participants with sentence contexts that biased different senses of the same verb, and measured self-paced reading of sentence complements that followed the verb. Their results indicate that when the sentence context cued a direct object sense of the verb, participants took more time in reading the main verb of the sentence complement (e.g., *was written* in the above example) if the complementizer *that* was absent; however, when the context cued a sentence complement sense of the verb, reading times were similar regardless of whether *that* was present. These results are highly compatible with those from McRae et al. (1998) and Trueswell et al. (1993).

Electrophysiological Data

Empirical studies have also provided electrophysiological evidence, particularly from event-related potentials (ERPs), for the representation of verb meanings in sentence processing. Because of the high temporal resolution provided by ERPs, these studies add additional insights into the time course of on-line processing as words unfold in a sentence. A major discovery in the ERP literature on language processing was the finding that the semantic violation or incongruence of a word in a sentence (e.g., *John put some sugar into the dog*) elicits the N400, a negative potential that peaks at about 400 msec into the processing of the word, compared to the absence of the N400 in semantically congruent sentences (e.g., *John put some sugar into the coffee*) (Kutas & Hillyard, 1980). Particularly relevant to our investigation here is the finding that the amplitude of the N400 can vary with the degree of semantic fit of the word to sentence context (the “cloze probability,” or degree of expectancy of a word in sentence; Kutas & Hillyard, 1984). Federmeier and Kutas (1999) tested participants with sentences such as the following: *They wanted to make the hotel look more like a tropical resort. So along the driveway, they planted rows of palms*. In three different testing conditions, the last word was either the expected exemplar (e.g., *palms*), an unexpected exemplar from the same category (e.g., *pinetrees*), or an unexpected exemplar from a different category (e.g., *tulips*), respectively. Compared to the absence of the N400 in the expected exemplar, the two unexpected exemplars both led to the N400; within the unexpected exemplars, the one from the same category elicited smaller N400 than the one from a different category. These patterns show that the ERP amplitudes may reflect the

the multiple weightings (e.g., frequency and distance) of all other co-occurring words (termed “global lexical co-occurrences”). Taking this perspective, researchers have argued that linguistic experiences with global lexical co-occurrences may have given the learner the necessary and sufficient conditions to acquire word meanings from sentence contexts (Li, Farkas, & MacWhinney, 2004; Li, Burgess, & Lund, 2000; Landauer & Dumais, 1997).

If what is important for the mental lexicon is the global lexical co-occurrence constraints, and if co-occurrence frequency plays a critical role in this process, then we need to reconsider semantic representation of the verb, the SRs, and the semantic congruency or violation at the phrasal or sentence level. Given the same grammatical roles assumed by the verb argument, we may still have incongruent or inconsistent verb-to-argument or verb-to-sentence semantic relationships. For example, the verb *feed* implies a situation in which an animate being nurses another animate being (the latter being typically immature or weak). Thus, *feed the baby* is semantically congruent, whereas *feed the table* would be incongruent. However, there could be cases that are less typical but acceptable, especially with support from sentence context, such as *feed the grass* (with fertilizers) or *feed the mother* (in the hospital). Such congruence–incongruence differences on a continuum may have arisen simply from the experiences of language use, in which specific semantic properties of the verb argument, as well as the other grammatical properties of it, are required as part of the verb’s SR and represented in the mental lexicon.

This new computational view, especially global lexical co-occurrence as a way of defining the lexicon, gives us a broader and more inclusive perspective on what we may call “lexical” or “lexicon” (more on this in the General Discussion): Rather than treating the lexicon as containing confined static lexical properties within lexical entries, we view the lexicon as a dynamic entity that is dependent on and defined by the context in which it occurs, and as such, it incorporates detailed and fine-grained grammatical and semantic information of not just the words themselves but also their contextual environment (the total linguistic environment that provides the usage history of the words). Although computational analyses have demonstrated the adequacy of this new perspective in capturing the meaning of words, the implications of this perspective for mental representation have not been independently verified with behavioral or neural data. Our study serves this purpose, to confirm or to disprove the computational account as it applies to the mental representation of verbs. In particular, our study makes the following hypotheses: (1) The mental representation of verbs includes not only grammatical information of the verb arguments as discussed previously but also specific semantic information such as animacy and humanness of the arguments,

defined properly by constraints imposed by lexical co-occurrences in natural language use; and (2) the SR fit varies as a result of the specific properties of the verb’s arguments, so that semantic congruency is a matter of degree, again a consequence of co-occurrence constraints in language use.

To the extent the SR “degrees of fit” may be reflected in the processes in which we understand words and sentences, we can test our hypotheses with behavioral and electrophysiological measures with normal adult speakers. In particular, we want to manipulate the degree of congruency between verb and object on animacy and other semantic dimensions. Effects of congruency at the word level (Experiments 1 and 2) would allow us to identify the role of lexical co-occurrence constraints in the mental representation of verbs, and effects of congruency at the sentence level (Experiments 3–5) would allow us to identify the ways in which such semantic constraint information is used and integrated during on-line sentence comprehension. Convergences between behavioral data and electrophysiological data would strengthen our proposal, whereas divergent patterns would lead to weakened arguments in support of the new lexical view.

METHODS

Participants

A total of 110 participants took part in our experiments: Experiments 1 and 2 (reaction time [RT] studies) each involved 30 participants (age range: 18–22 years, 35 women), Experiments 2 and 4 (ERP studies) each involved 20 participants (age range: 18–25 years, 13 women), and Experiment 5 involved 10 participants (age range: 18–25 years, 5 women). All participants were native speakers of Chinese with no reported hearing or neurological disorders. They were students from various universities in Beijing. They had normal or corrected-to-normal vision, and received monetary compensation for their participation.

Stimuli and Design

The stimuli and design of Experiments 1 and 2 were identical, as described below.

Two hundred verbs and their co-occurring objects were used in a single-factor within-subject design. They were all two-character words. In the congruent condition, the semantic features of the co-occurring object were [\pm animate], [\pm human], and [\pm specific],³ while in the three incongruent conditions, the semantic features of the object varied on one or a combination of the above features. An example given in Table 1 illustrates the four conditions: SR1, Appropriate Object; SR2, Inappropriate Human Object; SR3, Inappropriate Animal Object; and SR4, Inappropriate Inanimate Object.

Table 1. Example of Design and Materials Used in Experiments 1 and 2

| SR Conditions | Verb | Object | Semantic Features | | |
|---------------|-----------|-------------------------|-------------------|-------|----------|
| | | | Animate | Human | Specific |
| Appropriate | 雇佣 (hire) | 保镖 (<i>bodyguard</i>) | + | + | + |
| Inappropriate | | | | | |
| Human | | 婴儿 (<i>baby</i>) | + | + | – |
| Animal | | 母鸡 (<i>hen</i>) | + | – | – |
| Inanimate | | 电线 (<i>wire</i>) | – | – | – |

To obtain the degree of fit between the verbs and their co-occurring objects, a separate group of 60 college students was asked to rate each verb–object combination for 230 verb–object pairs on a scale from 1 (*highly inappropriate fit*) to 7 (*highly appropriate fit*). These verbs (along with their frequency information) were selected from the *Frequency Dictionary of Modern Chinese* (Beijing Language Institute, 1986). Based on the rating results, a set of 200 verbs and their co-occurring objects was selected as the experimental stimuli. In order to control for possible item-specific effects due to word frequency and number of character strokes, all critical object items were matched across conditions with respect to frequency and number of strokes, as shown in Table 2. The rating scores for the semantic fit between the verb and the object varied from 6.34 (for appropriate object) to 3.92 (inappropriate human), 3.03 (inappropriate animal), and 2.05 (inappropriate inanimate).

The stimuli for Experiment 3 were identical to those described above for Experiment 1, and the stimuli for Experiments 4 and 5 were identical to those for Experiment 2, except that they were tested in a sentence context. Sentence contexts and their co-occurring verb–object items were used in a single-factor within-subject design. The materials consisted of sentence contexts (e.g., *For the sake of his safety, the millionaire decided to hire a . . .*) paired with verb–object items that formed

semantically congruent or incongruent endings. There were three critical conditions: SR1, the congruent condition, in which the terminal object is appropriate for the verb and congruent with the previous context (e.g., *bodyguard*); SR2, the incongruent object appropriate condition, in which the object is appropriate for the verb but is incongruent with the previous context (e.g., *salesman*); and SR3, the incongruent object inappropriate condition, in which the object is inappropriate for the verb and incongruent with the previous context (e.g., *hen*). We did not include the fourth condition in which the inappropriate object is an inanimate object, given that the findings in Experiments 1 and 2 showed no significant difference between animal and inanimate inappropriate object conditions. An example in Table 3 illustrates the three conditions.

To obtain sentence contexts with varying degrees of constraint for the object, another group of 45 college students was asked to rate the degree of congruence between sentence contexts and their co-occurring final objects on a scale from 1 (*highly inappropriate fit*) to 7 (*highly appropriate fit*). The rating scores for the semantic fit between the sentence context and the final verb varied from 6.36 (congruent appropriate object) to 3.47 (incongruent appropriate object) and 1.64 (incongruent inappropriate object). All critical objects were matched across conditions with respect to frequency and the number of character strokes, as presented in Table 4.

All the stimuli were Latin-square assigned into four equal groups for Experiments 1 and 2 (three groups for Experiments 3–5) according to the test conditions. The participants were randomly assigned to each group, and no participant heard or read the same verb (or sentence context) or named the same object more than once. To balance the number of appropriate and inappropriate objects for verbs in Experiments 1 and 2, an additional set of 80 appropriate filler objects was used, such that for each participant, the stimulus list included 200 critical verb–object pairs and 80 filler pairs, of which 130 were appropriate combinations. To balance the number of congruent and incongruent sentences in Experiments 3, 4, and 5, an additional set of 66 congruent filler

Table 2. Mean Frequencies, Number of Strokes, and Ratings for the Object Items Used in Experiments 1 and 2

| | Appropriate | Inappropriate | | |
|--------------------------|-------------|---------------|--------|-----------|
| | | Human | Animal | Inanimate |
| Frequency ^a | 19.36 | 17.61 | 16.93 | 18.54 |
| Number of strokes | 14.65 | 14.84 | 15.96 | 15.42 |
| Rating (verb–object fit) | 6.34 | 3.92 | 3.03 | 2.05 |

^aPer million, according to *Frequency Dictionary of Modern Chinese*.

Table 3. Example of Design and Materials Used in Experiments 3 and 4

| Sentence Context | Object | SR Conditions |
|--|-------------------------|----------------------------------|
| 为了确保人身安全，这位百万富翁决定雇佣... | 保镖 (<i>bodyguard</i>) | Congruent appropriate object |
| <i>For the sake of his safety, the millionaire decided to hire a ...</i> | 店员 (<i>salesman</i>) | Incongruent appropriate object |
| | 母鸡 (<i>hen</i>) | Incongruent inappropriate object |

sentences was used, such that for each participant the stimulus list included 200 critical stimulus sentences and 66 filler sentences, of which 133 were congruent sentences. Finally, 20 items served as the practice trials for each of the four SR conditions in Experiments 1 and 2, and 10 items for each of the three conditions in Experiments 3, 4, and 5.

Apparatus and Procedure

The presentation of the stimuli, the recording of the RT (time between the onset of the visual object and the vocal response), and the vocal responses in Experiments 1 and 3 were controlled by the experimental software DMDX (Forster & Forster, 2003). The presentation of the stimuli and the recording of electroencephalographic (EEG) data in Experiments 2, 4, and 5 were controlled by the software Stim and Scan (NeuroSoft, 2003). All visual stimuli (in Chinese characters) were encoded in 48-point *songti* fonts, presented as pictures on a PC. A word was about 2.4×3.8 cm in size and participants were seated about 40 cm from the computer screen. All auditory stimuli were recorded using a DAT tape recorder and digitized into the PC.

Experiment 1

Each participant was tested individually in a dimly lit experimental room. The experimental session for each participant lasted about 20 min. Participants were given a short break during the experiment. A single trial con-

Table 4. Mean Frequencies, Number of Strokes, and Ratings for the Object Items Used in Experiments 3 and 4

| | Congruent Appropriate Object | Incongruent Appropriate Object | Incongruent Inappropriate Object |
|-------------------------|------------------------------------|--------------------------------------|--|
| Frequency ^a | 19.36 | 20.73 | 16.93 |
| Stroke | 14.65 | 14.76 | 15.96 |
| Rating (verb-object) | 6.34 | 6.34 | 3.03 |
| Rating (sentence) | 6.36 | 3.47 | 1.64 |

^aPer million, according to *Frequency Dictionary of Modern Chinese*.

sisted of a focal point (+) in the center of the screen for 250 msec, a blank screen for 200 msec, and an auditory verb (e.g., *guyong* “hire”), followed by a visual word (e.g., *baobiao* “bodyguard”; the object of the verb) until the participant named the word (within a 2-sec period). The intertrial interval was 1 sec. The task of the participant was to listen to the verb carefully and to name the visual word as quickly and accurately as possible. To ensure that the participant listened to the words carefully, the participant was also required to judge whether a randomly presented visual probe at the end of some trials was the verb they had just heard. Fifteen visual probes were among the verbs heard and 15 visual probes were not.

Experiment 2

The participants were tested in a dimly lit and sound-attenuated booth. They were instructed to sit comfortably and not to move their head. A single trial consisted of the following events: a focal point (+) in the center of the screen (300 msec), a blank screen (500 msec), a visual word (target verb, 800 msec), a blank screen (200 msec), a visual word (object of the verb, 200 msec), a blank screen (1500 msec), a visual probe (1000 msec), and a blank screen (3000 msec). The intertrial interval was 1 sec. The task of the participant was to judge whether the visual probe had appeared as the verb or the object in the preceding stimuli by pressing either the “yes” or the “no” response key. Among all trials, 140 were “yes” trials and 140 were “no” trials. The judgment task was to encourage the participants to read the critical stimuli (verb and object) carefully (the results of which were not analyzed). The experimental session lasted about 40 min. Participants were given two short breaks during the experiment.

Experiment 3

Testing environment and basic instructions to participants were identical to those in Experiment 1. A single trial consisted of a focal point (+) in the center of the screen for 250 msec, a blank screen for 200 msec, and an auditory sentence context, followed by a visual word (object of the verb) until the participant named the word (within a 2-sec period). The intertrial interval was

1 sec. The task of the participant was to listen to the sentence carefully and to name the visual object word as quickly and accurately as possible. To ensure that the participants listened to the sentences carefully, we required them to judge whether a randomly presented visual probe at the end of some trials had appeared in the sentence they just heard. Fifteen sentences had visual probes and 15 did not. Participants were given two short breaks during the experiment.

Experiment 4

Testing environment and basic instructions to participants were identical to those in Experiment 2. A single trial consisted of the following events: a focal point (+) in the center of the screen (500 msec), a blank screen (100 msec), a sentence context (presented word-by-word with 400 msec for each word and an interstimulus interval [ISI] of 200 msec), a blank screen (600 msec), a visual word (object of the verb, 1000 msec), a blank screen (1500 msec), a visual probe (2000 msec), and a blank screen (2500 msec). The task of the participant was to judge whether the probe occurred in the sentence that they had just seen by pressing either the “yes” or the “no” response key. Among all trials, 133 were “yes” trials and 133 were “no” trials. The judgment task was to encourage the participants to read the sentences carefully (the results of which were not analyzed). The experimental session lasted about 60 min. Participants were given three short breaks during the experiment.

Experiment 5

Testing environment, basic instructions, and procedures were identical to Experiment 4, except that the interval between the verb (the last word of the sentence context) and the object was not 600 msec but 200 msec. This additional experiment with the short ISI condition was carried out to safeguard against the possibility that patterns obtained in Experiment 4 were due to a long ISI which led to strategic, unnatural processing during comprehension.⁴

ERP Recording

EEG data were recorded using a 64-channel Quick-cap with Ag/AgCl electrodes. EEG electrodes were placed according to the extended 10–20 system. The average of the left mastoid and right mastoid was the reference. Horizontal eye movements were monitored with electrodes at the left and right outer canthi, and vertical movements were monitored with the electrodes above and below each participant’s left eye. All impedances were kept below 5 k Ω . Two sets of 32-channel amplifier (SYNAMPS, NeuroScan) were used for electrooculogram

and EEG recording. Band pass was 0.05–100 Hz and the sampling rate was 500 Hz.

ERP Data Analysis

A blink-correction algorithm (SpatialSVD) was applied off-line. Epochs of 1200 msec length were cut out from the continuously recorded data. The epochs started 200 msec before onset of the object words. After baseline correction and filtering (band pass was 0.05 Hz–30 Hz), trials with potentials greater than 70 μ V were rejected. The average of ERPs was computed for every participant, electrode, and condition.

On the basis of visual inspection and previous studies, we chose the time window of 300–450 msec for the N400 effects. For the statistical analysis of the EEG data, analyses of variance (ANOVAs) were computed for the midline and lateral electrodes separately for Experiments 2, 4, and 5. For the midline electrodes, the potentials from the electrodes Fz, Cz, Pz, and Oz were analyzed; for lateral electrodes, six regions of interest (ROIs) were defined by crossing the two factors, region (frontal vs. central vs. posterior) and hemisphere (left vs. right). Each ROI included three electrodes: left anterior (F3, F5, and F7); right anterior (F4, F6, and F8); left central (C3, C5, and T7); right central (C4, C6, and T8); left posterior (P3, P5, and P7); and right posterior (P4, P6, and P8). For Experiment 2, we conducted a 4 (electrode: Fz/Cz/Pz/Oz) \times 4 (SR conditions) ANOVA, and a 2 (hemisphere: left/right) \times 3 (region: anterior/central/posterior) \times 4 (SR conditions) ANOVA. Similarly, for Experiments 4 and 5, we conducted a 4 (electrode) \times 3 (SR) ANOVA, and a 2 (hemisphere) \times 3 (region) \times 3 (SR) ANOVA. Greenhouse–Geisser corrections were applied to all repeated measures with greater than 1 *df*.

RESULTS

Experiment 1: RT Study of Verb–Object Congruency in Isolation

One participant’s data (long RTs mostly over 1000 msec) were excluded from further analyses (mean RTs for other participants: 655 msec). RTs that were 3 *SDs* away were also excluded from our analyses. After these adjustments, mean RTs on the basis of correct responses are shown in Figure 1.

An ANOVA was conducted on the RT data with SR conditions as a within-participant factor. The main effect of SR was significant by participants [$F(3,84) = 9.92, p < .01$]. Post hoc LSD tests indicated that the difference between SR1 (appropriate) and SR2 (human) was marginally significant ($p = .076$), the differences between SR1 and SR3–4 (animal and inanimate) or SR2 and SR3–4 were significant ($p < .01$), and the difference between SR3 and SR4 was not significant ($p > .05$). These results show that the inappropriate objects elicited slower nam-

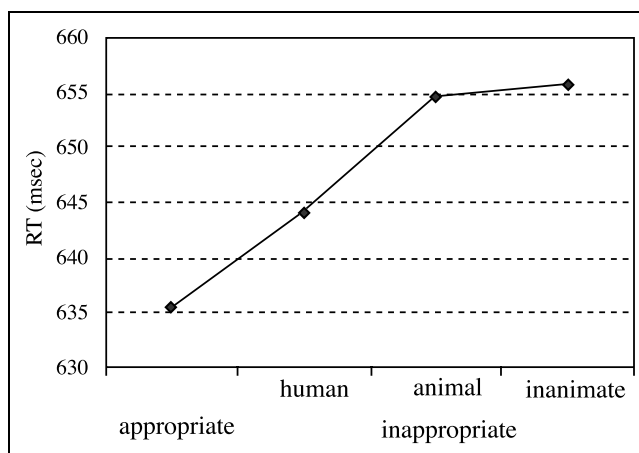


Figure 1. Mean RTs (msec) for the four SR conditions in Experiment 1.

ing responses than the appropriate objects, the inappropriate nonhuman objects elicited slower responses than the inappropriate human objects, and the two inappropriate nonhuman objects showed no significant difference. These patterns of results support the hypothesis that the degree of semantic fit between the verb and the object is reflected in the speed of processing during lexical access, suggesting that lexical representation of verbs

may include not only grammatical but also semantic details of the arguments of verbs.

Experiment 2: ERP Study of Verb–Object Congruency in Isolation

Figure 2 presents the ERPs for a subset of nine electrodes. Negative amplitudes are plotted above the horizontal midline.

As shown in Figure 2, the ERPs in the four conditions are very similar in the first 300 msec after the onset of the object word: A negative component (N100) is visible followed by a positive component (P200). This pattern is consistent with previous studies presenting language stimuli visually (see Kutas & Van Petten, 1994, for a review). In this experiment, our primary interest is in the subsequent negativity component and its sensitivity to verb–object congruence. The N400 component in this experiment was the slow negative shift, which peaked at around 400 msec after the onset of the word that had broader scalp distribution with the strongest effect observable at temporal and parietal sites, consistent with other reports.

ANOVA results for the time window 300–450 msec (the peak latency was around 375 msec) for midline and lateral electrodes are shown in Table 5.

Figure 2. ERP grand averages (nine electrodes) for the four SR conditions in Experiment 2.

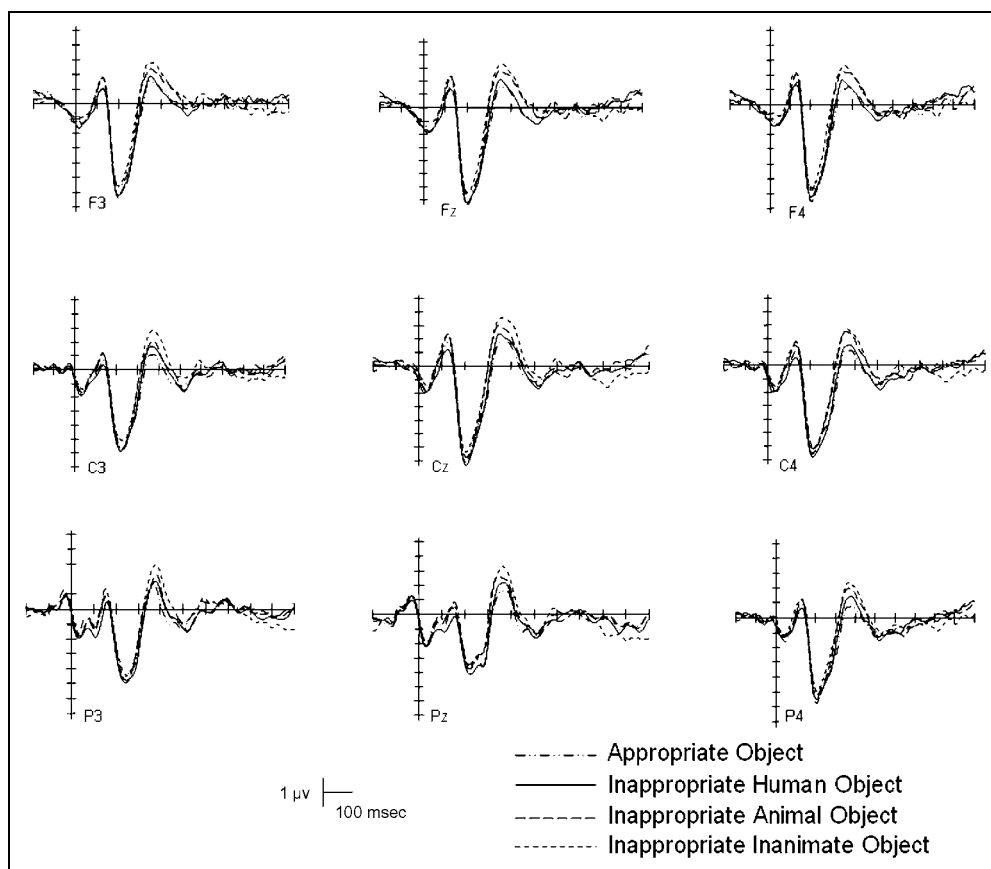


Table 5. ANOVA Results for N400 at Midline and Lateral Electrodes in Experiment 2

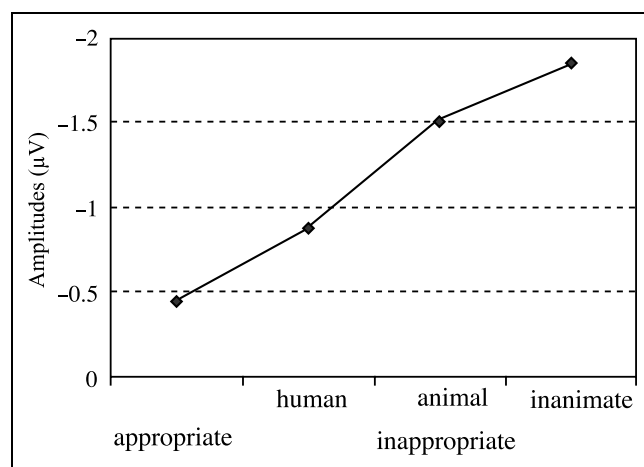
| ANOVA | <i>F</i> | <i>p</i> | ϵ |
|--|--------------------|----------|------------|
| Midline electrodes | | | |
| SR | $F(3,57) = 14.530$ | .000 | .948 |
| SR \times Electrode | $F(9,171) = 2.501$ | .069 | .486 |
| Lateral electrodes | | | |
| SR | $F(3,57) = 8.468$ | .000 | .778 |
| SR \times Hemisphere | $F(3,57) = 3.419$ | .096 | .425 |
| SR \times Region | $F(6,114) = 0.863$ | .409 | .265 |
| SR \times Hemisphere \times Region | $F(6,114) = 1.072$ | .321 | .188 |

Midline Electrodes

The main effect of the SR condition was significant. The interaction of SR by Electrodes failed to reach significance ($p > .05$) (Table 5). Multiple comparisons showed that there were significant differences between the appropriate object condition and all inappropriate conditions ($p < .05$). Within the inappropriate object conditions, the mean amplitudes of the N400 did not show a significant difference for animal and inanimate objects ($p > .1$), but both conditions were significantly more negative than the inappropriate human object condition ($p < .05$). The amplitudes of the N400 are shown in Figure 3.

Lateral Electrodes

The main effect of SR was significant. There were no significant effects of SR \times Hemisphere, SR \times Region, or the three-way SR \times Hemisphere \times Region interactions (Table 5). The results indicate clearly that the effect of SR was similarly observed over different hemispheres (left/right) and different cerebral regions (anterior/central/posterior). The N400 difference pattern among the var-

**Figure 3.** N400 amplitudes at midline electrodes in Experiment 2.

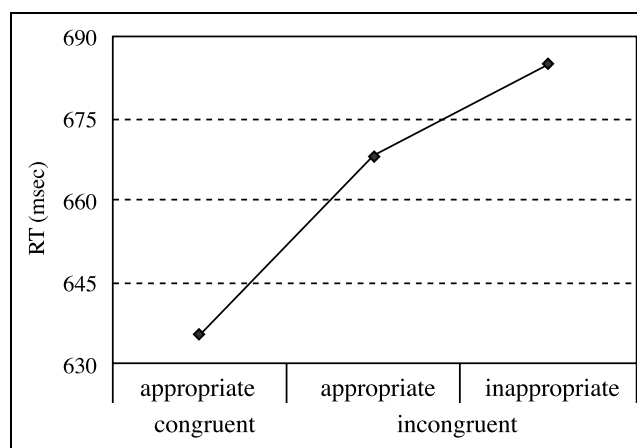
ious SR conditions was the same as that for the midline electrodes: inappropriate animate and inanimate object conditions elicited the largest N400, followed by the inappropriate human object condition, and then by the appropriate object condition. The classic N400 effect has been reported to have its maximal strength over centroparietal areas. A few studies also found the effect distributed in the anterior areas (Holcomb, 1993; Van Petten & Kutas, 1991). Our results are generally consistent with previous studies in terms of the scalp distribution of the N400.

To summarize Experiment 2, the graded ERP patterns from both the midline and lateral electrodes support the hypothesis that the amplitude of the N400 can reflect the degree of semantic fit between the verb and the object during lexical access. These results are highly consistent with the behavioral data from Experiment 1, showing that SRs in terms of semantic details of the verb arguments significantly impact the processing of the mental lexicon.

Experiment 3: RT Study of Verb–Object Congruency in Sentence Context

Exclusionary criteria were identical to those used in Experiment 1, which removed two participants' data (long RTs mostly over 1000 msec) as well as RTs more than 3 SDs from further analyses. Mean RTs on the basis of correct responses are shown in Figure 4.

An ANOVA was conducted on the RT data with SR conditions as a within-participant factor. The main effect of SR was highly significant by participants [$F(2,54) = 43.74, p < .01$]. No significant differences were found between SR conditions for error rates in naming. It is clear from Figure 4 that the naming speed for objects in both of the incongruent conditions was significantly slower than that for objects in the congruent condition. Within the incongruent conditions, naming speed for inappropriate objects was significantly slower than that for objects appropriate to their verbs. Post hoc LSD tests

**Figure 4.** Mean RTs (msec) for the three SR conditions in Experiment 3.

also indicated that the differences between SR1 (object appropriate, congruent) and SR2 (object appropriate, incongruent) and SR1 and SR3 (object inappropriate, incongruent) were significant ($p < .01$), and the difference between SR2 and SR3 was also significant ($p < .01$). These results show that naming speed is a function of both the fit of the object to the verb and the fit of the verb plus object to the entire sentence context. The patterns of results are consistent with those from previous studies on lexical and sentence processing in both English and Chinese, in which local information at the word level is continuously evaluated against and integrated with sentence contexts (Zhang, Wu, & Yip, 2006; Marslen-Wilson, 1987).

Experiment 4: ERP Study of Verb–Object Congruency in Sentence Context (Long ISI)

Figure 5 presents the ERPs for a subset of nine electrodes. Negative amplitudes are plotted above the horizontal midline.

As shown in Figure 5, the ERPs in the three SR conditions are very similar in the first 300 msec after the onset of the object word. Our primary interest is in the subsequent negativity component and its sensitivity to the SR fit between the verb and the object, and the congruency of verb–object and sentence context. In our statistical analyses, we examined the mean ERP amplitudes in the three experimental conditions, within a time window of 300 to 450 msec for negativity effects to unfold.

Figure 5. ERP grand averages (nine electrodes) for the three SR conditions in Experiment 4.

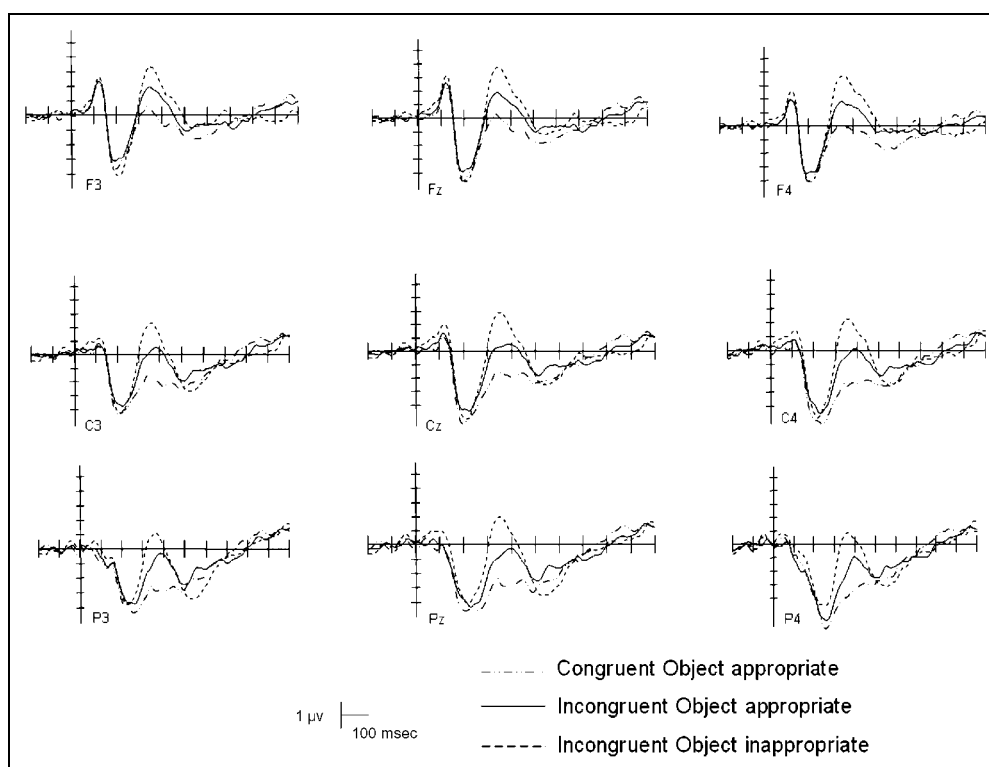


Table 6. ANOVA Results for N400 at Midline and Lateral Electrodes in Experiment 4

| ANOVA | <i>F</i> | <i>p</i> | ϵ |
|--------------------------|--------------------|----------|------------|
| Midline electrodes | | | |
| SR | $F(2,38) = 48.264$ | .000 | .888 |
| SR × Electrode | $F(6,114) = 5.208$ | .007 | .387 |
| Lateral electrodes | | | |
| SR | $F(2,38) = 19.889$ | .000 | .723 |
| SR × Hemisphere | $F(2,38) = 2.564$ | .123 | .541 |
| SR × Region | $F(4,76) = 0.895$ | .363 | .270 |
| SR × Hemisphere × Region | $F(4,76) = 1.150$ | .298 | .254 |

We also analyzed ERPs separately for midline and lateral electrode sites. From the measured electrodes, we selected four midline electrodes and 18 lateral electrodes, which were identical to those used in Experiment 2.

ANOVA results for the time window of 300–450 msec (the peak latency was around 360 msec) for the midline and the lateral electrodes are shown in Table 6.

Midline Electrodes

The main effect of SR was significant. There was a significant interaction of SR by Electrode (Table 6). Multiple comparisons revealed that there was a significant

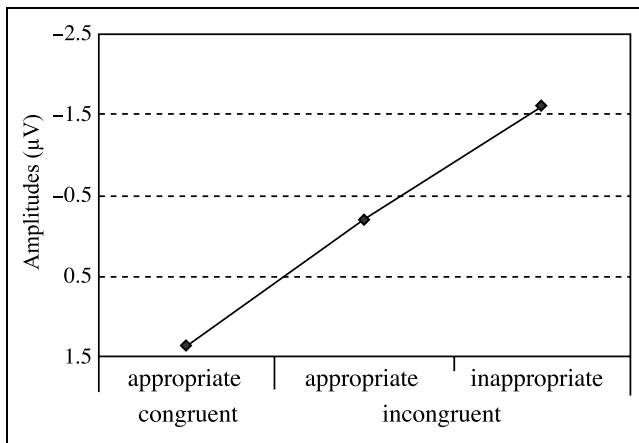


Figure 6. N400 amplitudes at midline electrodes in Experiment 4.

difference between the SR1 (object appropriate, congruent) and SR2–SR3, the two incongruent conditions ($p < .01$). Within the incongruent conditions, the amplitudes of the N400 for SR3 (object inappropriate, incongruent) were significantly larger than those for SR2 (object appropriate, incongruent) ($p < .01$). The amplitudes of the N400 are shown in Figure 6.

Lateral Electrodes

The main effect of SR was significant. There were no significant effects of SR \times Hemisphere, SR \times Region, or

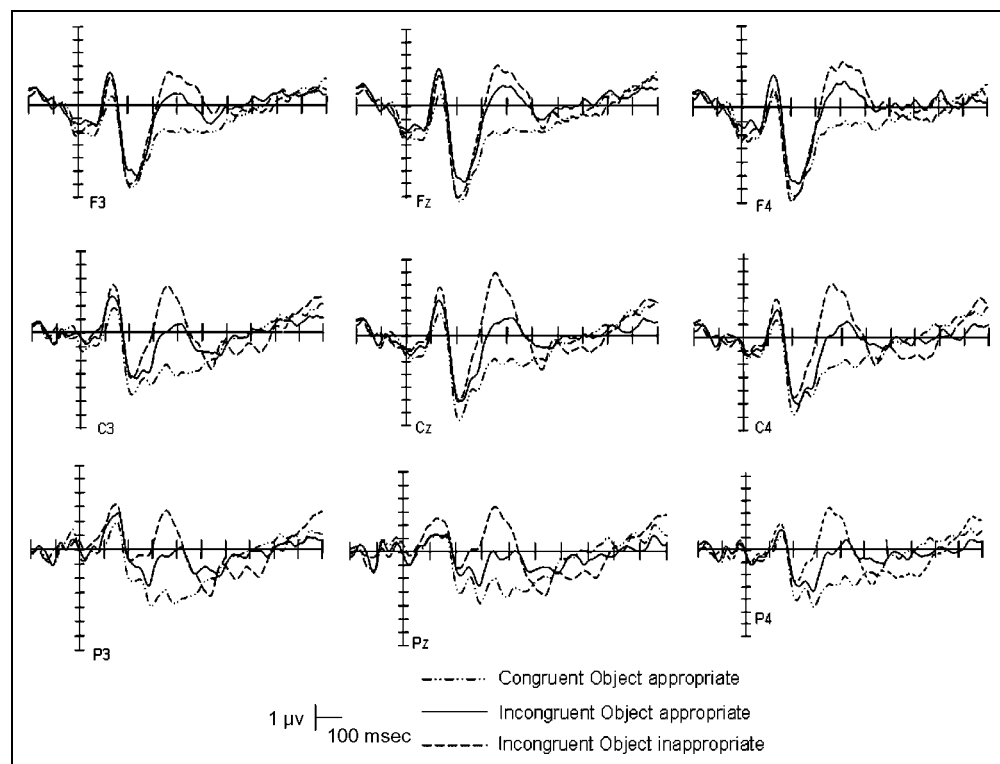
the three-way SR \times Hemisphere \times Region interactions (Table 6). Consistent with the results found in Experiment 2, the results here also indicate that the effect of SR was similarly observed over different hemispheres (left/right) and different cerebral regions (anterior/central/posterior). The N400 difference pattern among the SR conditions was the same as that for the midline electrodes as a result of the semantic and sentence context fit: SR3 elicited the largest N400, followed by SR2, and then by SR1.

To summarize, the graded N400 differences between congruent object, appropriate context-incongruent object, and inappropriate context-incongruent object provided further evidence that semantic information about the verb's argument as well as the verb itself is continuously evaluated and integrated with sentence contexts during processing.

Experiment 5: ERP Study of Verb–Object Congruency in Sentence Context (Short ISI)

Figure 7 presents the ERPs for a subset of nine electrodes, Figure 8 the amplitudes of the N400 for midline electrodes, and Table 7 the ANOVA results for the midline and the lateral electrodes within a time window of 300–450 msec. These results show very similar patterns as those found in Experiment 4 with the long ISI. The main effect of SR for the midline electrodes was significant (see Table 7), and multiple comparisons showed significant differences between SR1 and SR2

Figure 7. ERP grand averages (nine electrodes) for the three SR conditions in Experiment 5.



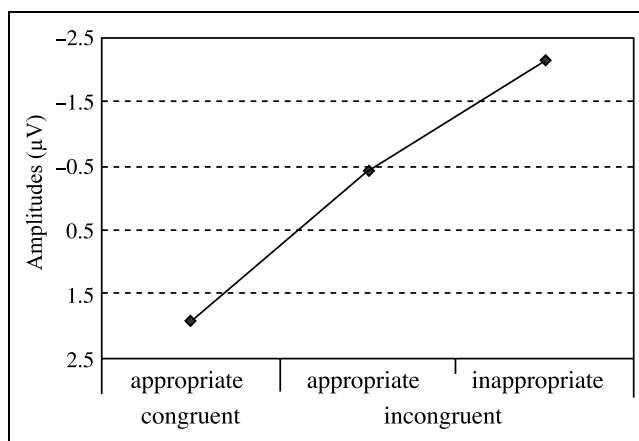


Figure 8. N400 amplitudes at midline electrodes in Experiment 5.

and between SR2 and SR3 ($p < .01$). The main effect of SR for the lateral electrodes was also significant, and as in Experiment 4, there were no significant effects of SR \times Hemisphere, SR \times Region, or the three-way SR \times Hemisphere \times Region interactions (Table 7). The similar patterns obtained from Experiments 4 and 5 indicate that the timing in our presentation of the visual stimuli did not adversely affect participants' on-line processing, and that the long (600 msec) versus short (200 msec) ISI between the verb and the object had a negligible effect on the ERP patterns, consistent with findings from other studies (e.g., Ainsworth-Darnell et al., 1998).

DISCUSSION

In this study, we set out to examine the representation of verb meanings in the mental lexicon by using both a psycholinguistic naming method and an ERP method to investigate different degrees of semantic fit in terms of SRs between the verb and its arguments. The behavioral results and the ERP results are highly consistent, in that both suggest that the semantic features of the verb argument may be an integral part of the verb meaning, and such information of the verb is accessed and used on-line during sentence processing. Both naming speed and ERP amplitude reflect the degree of fit. Naming times vary as a function of how well the verb argument fits the verb, with the fastest response for the most appropriate arguments (e.g., *hire a bodyguard*), in isolation or in sentence context, and the slowest response for the most inappropriate arguments (e.g., *hire a table*). ERP amplitudes vary as a function of how well the verb argument is congruent with the preceding verb, in and out of sentence contexts, with the largest N400 amplitude observed for the most inappropriate arguments and the smallest or no N400 amplitude for the most appropriate arguments.

These results match strongly with previous behavioral and ERP studies conducted in other languages, and at

the same time they provide new insights into the nature of mental representation of the lexicon. Earlier we had hypothesized that (1) the mental representation of verbs should include not only grammatical information of the verb arguments but also specific semantic features such as animacy and humanness of the arguments and (2) the specific features of the arguments would fit the verb at different levels so that semantic congruency is a matter of degree. Our results provide positive evidence on both of these hypotheses: naming speeds and ERP amplitudes (N400) both seem to index the degree of fit, such that slower RTs and larger N400 reflect stronger violations in fitting the verb's SRs. The ERP data also indicate that the congruency of verb-object may be additionally dependent on the sentence contexts which could specifically bias a given verb-object combination to be appropriate or inappropriate. These data are consistent with the role of sentence context in lexical processing, and with probabilistic models of language processing in general (Zhang et al., 2006; Hare et al., 2003; Federmeier & Kutas, 1999; MacDonald, 1993).

Confirmation of our hypotheses on the basis of empirical RT and ERP evidence in the present study leads us to further consider the nature of the mental lexicon in general. That the mental representation of a word should include grammatical and semantic information of not just the word per se but also information of words related to the target word becomes intuitively clear if we consider the role of global lexical co-occurrences in defining meanings of words. In the tradition of structural linguistics (Bloomfield, 1933; Saussure, 1916), the function of a word is determined entirely by the distributional properties of the word relative to other words in a sentence (e.g., words in the same category tend to occur in similar morphosyntactic contexts). Harris (1982) further elaborated these ideas in terms of co-occurrence constraints that can define not just large grammatical categories such as nouns or verbs but also subclasses within these categories (e.g., animate nouns that tend to occur before certain verbs, such as

Table 7. ANOVA Results for N400 at Midline and Lateral Electrodes in Experiment 5

| ANOVA | <i>F</i> | <i>p</i> | ϵ |
|--|--------------------|----------|------------|
| Midline electrodes | | | |
| SR | $F(2,18) = 26.794$ | .000 | .642 |
| SR \times Electrode | $F(6,54) = 10.819$ | .009 | .397 |
| Lateral electrodes | | | |
| SR | $F(2,18) = 28.645$ | .000 | .699 |
| SR \times Hemisphere | $F(2,18) = .803$ | .444 | .817 |
| SR \times Region | $F(4,36) = 1.474$ | .258 | .390 |
| SR \times Hemisphere \times Region | $F(4,36) = .332$ | .725 | .221 |

- de Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot. [English translation: *A course in general linguistics*. New York: Philosophical Library; Chinese translation: *Putong Yuyanxue Gangyao*. Beijing: Commercial Press].
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.
- Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, *8*, 301–306.
- Federmeier, K. D., & Kutas, M. (1999). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, *8*, 373–392.
- Fillmore, C. J. (1968). The case for case: Universals in linguistic theory. In Emmom Bach & Robert Harms (Eds.) *Universals of linguistic theory* (pp. 1–88). New York: Holt, Rinehart and Winston.
- Forster, K. L., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124.
- Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and Language*, *43*, 476–507.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*, 183–192.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The Syntactic Positive Shift (Sps) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*, 439–483.
- Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*, 438–441.
- Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language*, *48*, 281–303.
- Harris, Z. (1982). *A grammar of English on mathematical principles*. New York: Wiley.
- Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*, 220–225.
- Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*, 47–61.
- Jackendoff, R. S. (2002). *Foundations of language: Brain, meaning, grammar, and evolution*. Oxford: Oxford University Press.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.
- Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, *11*, 539–550.
- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161–163.
- 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, CA: Academic Press.
- Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of induction and representation of knowledge. *Psychological Review*, *104*, 211–240.
- Li, P. (2006). In search of meaning: The acquisition of semantic structures and morphological systems. In J. Luchjenbroers (Ed.), *Cognitive linguistics investigations: Across languages, fields and philosophical boundaries*. Amsterdam, Holland: John Benjamins.
- Li, P., Burgess, C., & Lund, K. (2000). The acquisition of word meaning through global lexical co-occurrences. In E. Clark (Ed.), *Proceedings of the Thirtieth Stanford Child Language Research Forum*. Cambridge: Cambridge University Press.
- Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks*, *17*, 1345–1362.
- Li, P., Tan, L. H., Bates, E., & Tzeng, O. (2006). New frontiers in Chinese psycholinguistics. In P. Li, L. H. Tan, E. Bates, & O. Tzeng (Eds.), *Handbook of East Asian psycholinguistics* (Vol. 1: Chinese). Cambridge: Cambridge University Press.
- MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, *32*, 692–715.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*, 71–102.
- McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312.
- Mintz, T., Newport, E., & Bever, T. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, *26*, 393–424.
- Münte, T. F., Heinze, H. J., & Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience*, *5*, 335–344.
- Nakagome, K., Takazawa, S., Kanno, O., Hagiwara, H., Nakajima, H., Itoh, K., & Koshida, I. (2001). A topographical study of ERP correlates of semantic and syntactic violations in the Japanese language using the multichannel EEG system. *Psychophysiology*, *38*, 304–315.
- Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*, 151–165.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.
- Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*, 425–470.
- Shapiro, L. P., Zurif, E., & Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. *Cognition*, *27*, 219–246.
- Shapiro, L. P., Zurif, E., & Grimshaw, J. (1989). Verb processing during sentence comprehension: Contextual impenetrability. *Journal of Psycholinguist Research*, *18*, 223–243.
- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 528–553.
- Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition*, *19*, 95–112.
- Zhang, Y., Wu, N., & Yip, M. (2006). Lexical ambiguity resolution in Chinese sentence processing. In P. Li, L. H. Tan, E. Bates, & O. Tzeng (Eds.), *Handbook of East Asian psycholinguistics* (Vol. 1: Chinese). Cambridge: Cambridge University Press.