CROSSLINGUISTIC VARIATION AND SENTENCE PROCESSING: THE CASE OF CHINESE

PING LI

Department of Psychology
University of Richmond
Richmond, Virginia

1. INTRODUCTION

Theories of sentence processing have for the last two decades focused on issues such as parallel versus serial, interactive versus modular, and top-down versus bottom-up processes. One classic example is the study of lexical ambiguity (Small, Cottrell, and Tanenhaus, 1988; see other chapters in this volume), out of which two contrastive hypotheses have emerged. The first is the exhaustive access hypothesis, which argues that all meanings of an ambiguous word will be accessed momentarily following the occurrence of the word, and that semantic context can only help to select the appropriate meaning at a post-access stage. This hypothesis assumes that sentence processing is a modular, bottom-up process in which nonlexical, sentential information does not penetrate lexical access (Onifer and Swinney, 1981; Swinney, 1979). The second hypothesis is the context-dependency hypothesis, which argues that the contextually appropriate meaning of an ambiguous word can be selectively accessed early on if sentence context provides a strong bias to the appropriate meaning. This hypothesis assumes that sentence processing is an interactive process in which information can flow both bottom-up and top-down, and that lexical access and sentential context can mutually influence one another at a very early stage (Simpson, 1981; Simpson and Krueger, 1991; Tabossi, 1988).

0092-4563/98 $25.00

Most sentence-processing theories have been built on facts from Western languages, in particular, English. They emphasize the extent to which languages resemble each other in sentence processing, for example, on the use of a universal set of parsing principles (such as "late closure" and "minimal attachment," see Frazier, 1987, 1990). In recent years, however, we begin to see growing interests in crosslinguistic research in sentence processing. In particular, the competition model of Bates and MacWhinney (1982, 1987, 1989; see also Bates, McNew, MacWhinney, Devescovi, and Smith, 1982) has attracted attention from researchers with a crosslinguistic perspective on sentence processing. In contrast to other models, the competition model emphasizes the extent to which languages differ, and the extent to which these differences affect sentence processing.

Chinese, because of its many unique linguistic properties, plays an important role in crosslinguistic studies of sentence processing. Spoken by one-fifth of the world's population, Chinese differs significantly from most Indo-European languages and offers unique features in its phonological, lexical, and syntactic structures. On a phonological level, Chinese involves a tonal system that distinguishes lexical items, with each item carrying a particular tone. Tonal information can differentiate lexical items, but it does not eliminate lexical ambiguities associated with homophones: on a lexical-morphemic level, Chinese has a massive number of homophones. Moreover, on a syntactic level, first, Chinese does not have devices that indicate differences in tense, number, gender, or case; in other words, grammatical functions and relations for sentence constituents are not linked by morphological associations. Chinese has some grammatical morphemes such as the object-patient marker *ba* and the passive-agent marker *bei*, but these are often optional in natural speech. Second, Chinese involves a high degree of ellipsis. A sentence can be subjectless or objectless, sounding telegraphic in a richly inflected language when literally translated. Third, word order is relatively free in Chinese. In addition to the canonical SVO order, Chinese permits several other word orders in daily spoken language: OSV, in which the object is topicalized; SOV, in which the speaker provides information counter to the expectation of the listener (Li and Thompson, 1981); and VOS, in which the subject is usually an after-thought (Lu, 1980).

These language-specific properties of Chinese raise interesting questions about mechanisms of sentence processing. For example, the lack of inflections, the high degree of ellipsis, and the variability of word order together offer sentence status to many constructions that would be incomplete and/or ungrammatical in English or other Indo-European languages (e.g., V, OV, VO). Given that many daily sentences consist of fragments with omissions, we may expect partial information and sentence fragments to play a more important role in Chinese processing than they do in the processing of Indo-European languages. One of the goals of this chapter is to examine the differences between Chinese and other languages with respect to the use of different types of information in sentence processing.

The basic principles of the competition model seem to give rise to a natural explanation of the language-specific properties of Chinese and how these properties influence Chinese sentence processing. The competition model is a cue-based interactive model. It is concerned with how speakers integrate various types of information or cues in a sentence (e.g., word order, subject-verb agreement, and lexical semantics) to determine sentence roles (e.g., who does what to whom). A central tenet of the model is that the same cues vary in their validity across languages, with validity being determined by how often the cue is available, and how reliably the cue leads to the correct identification of linguistic functions. A typical finding in this research is that the strongest cue in one language can be one of the weakest cues in another. For example, in Italian, the strongest cue to the identification of sentence roles is verb morphology, whereas in English, morphological marking is extremely weak as a cue (Bates et al., 1982). In the last few years, my colleagues and I have used the basic constructs of the competition model to study how Chinese speakers assign sentence roles in the absence of overt grammatical devices and in the presence of a relatively free word order (Li, Bates, Liu, and MacWhinney, 1992; Li, Bates, and MacWhinney, 1993; Li, 1996a). This chapter provides an overview of some basic results from these studies.

The assignment of sentence roles is concerned with global functional (i.e., syntactic and semantic) relationships among sentence constituents.[1] In addition to the understanding of sentence role assignment in Chinese, I have also studied how Chinese speakers resolve local ambiguities associated with homophones within a sentence (Li, 1996b; Li and Yip, 1996; in press). Previous research in lexical ambiguity has focused on English and several Indo-European languages (e.g., Dutch and Italian). In those languages, homophony is a relatively low-frequency event. In Chinese, by contrast, homophony is extensive. According to the Modern Chinese Dictionary (Institute of Linguistics, 1985), 80% of the monosyllables in Chinese are ambiguous, and half of them have five or more homophones. The single syllable *yi* with the dipping tone has up to 90 homophones (e.g., skill, justice, benefit, discuss, intention, translate, hundred-million, etc.), and this number would increase to 171 if identical syllables with different tones were considered as homophones. Upon hearing *yi* in a sentence, do Chinese speakers activate all 90 or more meanings of the syllable? The exhaustive access hypothesis should predict they do, because lexical access is an autonomous and capacity-free process. The context-dependency hypothesis should predict that Chinese listeners activate only the contextually appropriate meaning with aid from sentence context. This chapter reviews relevant experimental evidence from my studies of the processing of Chinese homophones.

The chapter is organized as follows. First, I present results from experiments on how Chinese speakers determine sentence roles in a sentence-interpretation task (Experiment 1) and a sentence-gating task (Experiment 2). Second, I present results from experiments on Chinese speakers' processing of homophones, in a cross-modal naming task (Experiment 3) and a word-gating task (Experiment 4). Finally, I conclude with a discussion on the importance of crosslinguistic variations in our understanding of the mechanisms of sentence processing.

2. EXPERIMENT 1: SENTENCE INTERPRETATION

The sentence-interpretation task, according to Bates and MacWhinney (1982), is a task in which speakers choose one of the two nouns in a sentence as the agent or subject of the sentence. In the experiment, adult participants respond verbally for a choice and children act out the sentence with toys, after listening to the sentence (such as *The elephant kicked the dog*). A number of previous studies in Chinese have employed this task, including Miao (1981), Chen, Tzeng, and Bates (1990), and Li, Bates, Liu, and MacWhinney (1992). In the study reported here, a reaction time technique was used, in which listeners make their choices by pressing one of the two buttons upon a computerized presentation of a spoken sentence (Li, Bates, and MacWhinney, 1993).

2.1. Method

2.1.1. Participants

Eighteen native Mandarin Chinese speakers from mainland China participated in this experiment.

2.1.2. Materials and Design

Three independent variables were tested in this experiment:

1. *Noun animacy.* A noun in a sentence was either animate or inanimate. Because all test sentences contained two nouns, the first and the second nouns varied in three ways: Animate-Inanimate (AI), Inanimate-Animate (IA), and Animate-Animate (AA).
2. *Word order.* The sentences varied in three orders: the canonical Noun-Verb-Noun (NVN), usually associated with a SVO interpretation; the noncanonical Noun-Noun-Verb (NNV), associated with either a SOV or a OSV interpretation, and the noncanonical Verb-Noun-Noun (VNN), associated with a VOS interpretation.
3. *The presence or absence of a grammatical marker.* The marker *ba,* which marks the following noun as the patient of the sentence, or the marker *bei,* which marks the following noun as the agent of the sentence.

A total of 162 test sentences were used in this study: 54 each for (a) simple transitive sentences, (b) sentences with the marker *ba*, and (c) sentences with the marker *bei*.[2] In each set of sentences, there were six instances for each of the following nine types: (a) AVA (animate noun-verb-animate noun), AAV, VAA, AVI (animate noun-verb-inanimate noun), AIV, VAI, IVA (inanimate noun-verb-animate noun), IAV, and VIA; (b) AV*ba*A, A*ba*AV, VA*ba*A, AV*ba*I, A*ba*IV, VA*ba*I, I*ba*VA, I*ba*AV, and VI*ba*A; and (c) AV*bei*A, A*bei*AV, VA*bei*A, AV*bei*I, A*bei*IV, VA*bei*I, I*bei*VA, I*bei*AV, and VI*bei*A. The markers *ba* and *bei* always occurred before the second noun, to match up with their position of occurrence in natural speech.

2.1.3. Procedure

The test sentences were read by a native Mandarin speaker and were digitized into a Macintosh computer with a sampling rate of 22 kHz. During the experiment, participants received each sentence auditorially, and an Experimental Control System (Clynes and MacWhinney, 1990) controlled the presentation of test sentences. At the onset of the presentation of each sentence, participants also received on the computer screen a pair of pictures that were described by the two nouns of the sentence. They were instructed (in Chinese) to determine, as quickly and as accurately as possible, which of the two pictures represented the agent or the actor in the sentence they heard. They indicated this choice by pressing one of the two buttons mounted on a CMU button box (Cohen, MacWhinney, Flatt, and Provost, 1993).

2.1.4. Data Analysis

Although data from both choice responses and reaction times were collected and analyzed, only the choice response data are discussed here (see Li et al., 1993, for analyses of the reaction time data). For the scoring of choice responses, a participant's response was given a score of 1 if it was a choice of the first noun, and 0 if it was a choice of the second noun. The percent of first-noun choice is thus inversely related to that of second-noun choice: a score of 100% represents an exclusive choice of the first noun, while a score of 0% represents an exclusive choice of the second noun.

2.2. Results and Discussion

Figures 1a to 1c present the percent first-noun choice for the simple transitive sentences, the sentences with the marker *ba*, and the sentences with the marker *bei*, respectively. A 3×3 (animacy by word order) analysis of variance (ANOVA) was conducted separately on the results of each set of sentences, and a $3 \times 3 \times 2$

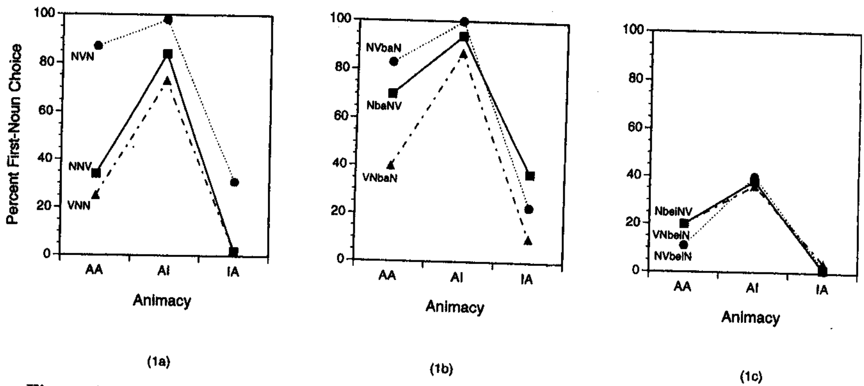

Figure 1. Percent first-noun choices for the simple sentences (1a), the *ba* sentences (1b), and the *bei* sentences (1c). (From Li, Bates, and MacWhinney, 1993. Reprinted by permission of Academic Press.)

(animacy by word order by the presence or absence of the marker) was conducted separately on the combined data of the simple and the *ba* sentences and those of the simple and the *bei* sentences. The latter procedure was used to understand the effects of the grammatical marker more clearly.

For the simple sentences, ANOVA revealed that there were significant main effects of animacy ($F(2,34) = 75.54, p < .01$) and word order ($F(2,34) = 21.61, p < .01$), and an interaction between the two ($F(4,68) = 11.46, p < .01$). An analysis of the magnitude of effect (7 score) shows that animacy accounted for 72% of the experimental variance, whereas word order accounted for 23%. This analysis shows that animacy is a more important cue than word order in Chinese, in contrast to the predominant role of word order in English. The interaction between animacy and word order shows that when the two cues agreed (i.e., when they both pointed to the same noun as agent), the sentence led to consistent responses; when they competed (i.e., when they pointed to different nouns as agent), the sentence led to inconsistent responses. For example, listeners chose 98% first noun for the AVI sentences, in which both animacy and word order favor the first noun. In contrast, listeners chose only 31% first noun for the IVA sentences, in which word order favors the first noun but animacy favors the second. These effects clearly demonstrate how different cues compete and collaborate in determining the assignment of sentence roles.

For the sentences with the marker *ba*, a 3 × 3 ANOVA revealed that there were main effects of animacy ($F(2,34) = 66.64, p < .01$) and word order ($F(2,34) = 10.32, p < .01$), and an interaction between the two ($F(4,68) = 11.28, p < .01$). A 3 × 3 × 2 ANOVA incorporating the simple sentences revealed that there was a significant main effect of the marker *ba* ($F(1,17) = 15.75, p < .01$), suggesting that the presence of *ba* contributed to the identification of sentence roles. However, a comparison of Figure 1b with 1a shows that the basic patterns for the simple sentences and for the *ba* sentences are similar, except that the A*ba*AV sentences elicited a higher percentage of first-noun choice than the simple AAV sentences. The presence of *ba* had its effect most clearly on the NNV word order, since this is the order in which *ba* naturally occurs in speech.

For the sentences with the marker *bei*, a different pattern of response emerged, in contrast to the patterns for the simple and the *ba* sentences. A 3 × 3 ANOVA revealed that there was a main effect of animacy ($F(2,34) = 12.80$, $p < .01$) and a significant interaction between animacy and word order ($F(4,68) = 2.89$, $p < .05$), but there was no main effect of word order ($F(2,34) = 1.51$, ns). Figure 1c shows that the difference between the three word-order types within the *bei* sentences was minimal. A 3 × 3 × 2 ANOVA incorporating the simple sentences revealed that there was a significant main effect of the marker *bei* ($F(1,17) = 27.65$, $p < .01$). A comparison of Figure 1c with 1a indicates that in all cases, the presence of *bei* suppressed the first-noun choice and promoted the second-noun choice. The 3 × 3 × 2 ANOVA also revealed a significant interaction between animacy and *bei* ($F(2,34) = 9.97$, $p < .01$). This interaction indicates effects of competition and convergence between cues: when *bei* (which marks the second noun as agent) agreed with animacy (when the second noun is animate, i.e., IA), the sentence led to almost exclusive second-noun choice; when *bei* competed with animacy (i.e., AI), the sentence had less consistent responses (about 40% first-noun choices). But it was clear that the effect of *bei* overwhelmed both animacy and word order in all cases of competition. This effect differed from that of *ba*, which was reflected only in the NNV orders.

The above results suggest that the strength of cues in Chinese sentence processing can be rank ordered as follows: the marker *bei* → animacy → word order → the marker *ba*. The results show that Chinese speakers use a weighted array of linguistic cues to help solve the task of sentence role assignment. Speakers of Indo-European languages usually rely on one primary type of cue in sentence interpretation: morphological, syntactic, or semantic, since in these languages the primary type of information is both highly available and highly reliable in sentence processing. Chinese lacks such primary types of cues (e.g., *bei* is highly reliable but often not available), so Chinese speakers have to rely on a multitude of cues, including noun animacy, word order, and the grammatical markers.

The difference between the effects of *ba* and those of *bei* raises another important issue. Although both *ba* and *bei* are considered as grammatical markers (or "semimorphological markers," see discussion in Li et al., 1993), their influence on sentence processing differs. In the above cue-strength hierarchy, *ba* and *bei* (grammatical markers) are intercepted by animacy (semantic cues) and word order (syntactic cues), which suggests that cues that fall into the same cluster of linguistic type do not necessarily fall on the same hierarchy of cue importance to sentence processing. Such results are best described by an interactive process in which different types of linguistic cues work together in collaboration and in competition. The outcome of the interactive process depends on the strengths of cues and the directions in which they interact (e.g., word-order effect is significant in sentences with *ba* but not in sentences with *bei*).

Experiment 1 indicated the relative importance of several cues in Chinese sentence processing, and the ways in which these cues interact to determine the final outcome of sentence role assignment. Although a reaction time study, the experiment was, however, confined to the "one sentence one choice" decision process, and was thus relatively uninformative to the question of when a choice decision is reached (i.e., the time course of sentence processing). To understand how cues interact with one another at different temporal locations of the sentence, in Experiment 2 I adapted the gating method from spoken-word recognition to study sentence role assignment (Li, 1994, 1996a).

3. EXPERIMENT 2: SENTENCE GATING

3.1. Method

3.1.1. Participants

Twenty native Mandarin Chinese speakers from mainland China participated in this experiment. None had participated in Experiment 1.

3.1.2. Materials and Design

To match up with Experiment 1, the test sentences included three sets of sentences: simple, *ba*, and *bei*. Each set had 36 sentences, 12 each for the three-word orders NVN, NNV, and VNN. The design included three levels of word order (NVN, NNV, and VNN), and either three or four levels of gate size (simple sentences had three levels, i.e., 1, 2, or 3 words; *ba* and *bei* sentences had four levels, i.e., 1, 2, 3, or 4 words). A third variable, position of the marker, was nested within the *ba* and *bei* sentences only, reflecting the position of *ba* or *bei* in the sentence (before the first word, before the second word, or before the third word). Thus, the design was 3×3 for the simple sentences, and $3 \times 4 \times 3$ for *ba* and *bei* sentences. Animacy was not varied in this experiment because the relative importance of animacy versus word order and *ba* and *bei* is clear from Experiment 1.

3.1.3. Procedure

The procedure was similar to that in Experiment 1, except the following. In Experiment 1, listeners heard the complete sentence and responded for their choice. In this experiment, listeners were presented auditorially with gated fragments of a sentence, one at a time, in increasing length, until they heard the complete sentence. At each presentation, they were required to decide on the agent of the sentence on the basis of the partial information provided up to that point (the last gate corresponded to the complete sentence). This procedure is an adaptation of the word-gating paradigm developed by Grosjean and his associates (Grosjean, 1980; Cotton and Grosjean, 1984; Tyler and Wessels, 1985), and it bears similarity to the self-paced sentence reading task.

3.1.4. Data Analysis

The gating study produced a large number of data points because each sentence involved at least three choice responses. The following discussion focuses on the simple sentences, the canonical NVN orders with the marker *ba*, and the canonical NVN orders with the marker *bei*. As in Experiment 1, only the choice responses are discussed here (see Li, 1996a, for a detailed discussion of the data).

3.2. Results and Discussion

Figures 2a to 2c present the percent first-noun choice for the simple sentences, the NVN sentences with the marker *ba*, and the NVN sentences with the marker *bei*, respectively.

For the simple sentences, a 3 × 3 (word order by gate size) ANOVA indicated a significant main effect of word order ($F(2, 38) = 85.75, p < .01$), showing that different orders led to different interpretations.[3] There was also a significant main effect of gate size ($F(2, 38) = 28.78, p < .01$), showing that listeners' choice responses tended to change along the course of the sentence. More important, the significant interaction between gate size and word order ($F(4, 76) = 26.21, p < .01$) showed that listeners built up their interpretations across the sentence in

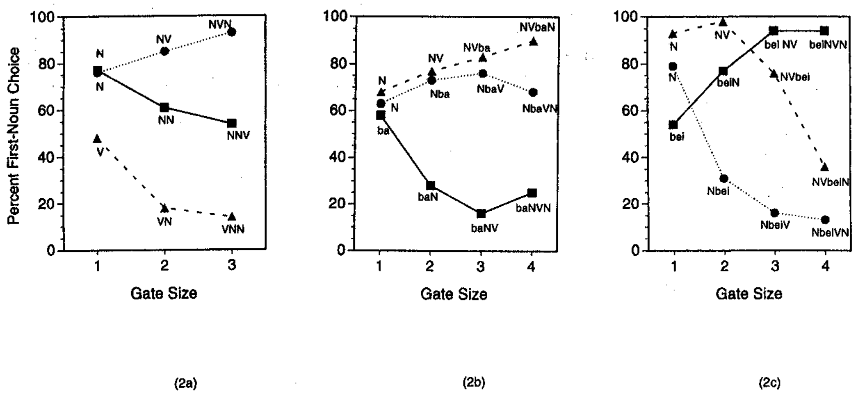

Figure 2. Percent first-noun choices for the simple sentences (2a), the NVN sentences with *ba* (2b), and the NVN sentences with *bei* (2c). (From Li, 1996a. Reprinted with permission of the Psychonomic Society Publications.)

different ways for different sentence types. For example, in Figure 1a, first-noun choices increased for NVN as gate size increased, but decreased for NNV and VNN. Listeners started out with either biased or random default interpretations, and then gradually adjusted (confirming or disconfirming) the interpretations as more information came in. An increase in the amount of sentence information may lead to more certain decisions (as in NVN or VNN), but it can also lead to more ambiguous decisions (as in NNV), depending on the interpretations that the construction is associated with in the language (e.g., NNV is ambiguous between SOV and OSV interpretations).

For the NVN sentences with *ba*, a 3 × 3 (position of marker by gate size) ANOVA indicated a significant main effect of position of the marker ($F(2, 38) = 41.92, p < .01$), no main effect of gate size ($F(3, 57) = .58$, ns), and a significant interaction between the two ($F(6,114) = 17.51, p < .01$). Figure 2b shows that the N*ba*VN and NV*ba*N sentences had mostly first-noun choices, similar to the simple NVN sentences without the marker *ba* (cf. Figure 2a). The similarity in the basic patterns of interpretation across the simple NVN and the N*ba*VN and NV*ba*N sentences suggests that the ungrammaticality of *ba* (i.e., occurring before and after a verb) did not disrupt listeners' comprehension; in Chinese, *ba* is always before a noun and it is ungrammatical before or after a verb. In contrast to N*ba*VN and NV*ba*N, for the *ba*NVN sentences, listeners quickly decided on the second noun as the agent (by the time they heard *ba*N), and this interpretation was further confirmed when they heard the verb (*ba*NV). Note that listeners' choice decisions were based on an interpretation of possible fragments (*ba*N and *ba*NV) from the complete grammatical N*ba*NV sentence: the N after *ba* was marked as the patient so that the alternative N could be assigned the agent role. Now if this interpretation were to hold, the sentence would end with a verb. But the sentence continued as *ba*NVN, with a final N after the verb. This continuation created a competition between two fragment interpretations: that of *ba*N or *ba*NV, in which the N before the verb should be interpreted as the patient, and that of VN, in which the N after the verb should be interpreted as the patient, leading to a final increase in first-noun choice.

For the NVN sentences with *bei*, a 3 × 3 (position of marker by gate size) ANOVA indicated significant main effects of both position of the marker ($F(2, 38) = 66.87, p < .01$) and gate size ($F(3, 57) = 30.44, p < .01$), and a significant interaction between the two ($F(6,114) = 42.73, p < .01$). Figure 2c shows that listeners had a first-noun bias when they heard only the first N alone in the N*bei*VN and NV*bei*N sentences (as they did with the simple NVN and NNV sentences; see Figure 2a and Note 3). First, for NV*bei*N, the bias was further confirmed when listeners heard NV, leading to a very high 98% first-noun choice with a SV interpretation. But when they heard *bei* after NV (i.e., NV*bei*), the first-noun choice lost weight, since *bei* indicates the following noun as the agent in the language. The weight went down further to favor a second-noun choice when the final N appeared, due to a competition between NV and *bei*N (with both nouns competing for agenthood). Second, for N*bei*VN, the decision moved quickly from a first-noun bias to a second-noun choice when *bei* occurred, and the second-noun choice continued through to the final gate. Note that the occurrence of *bei* in N*bei*VN, unlike *ba* in N*ba*VN, strongly affected the SVO interpretation associated with NVN. This is because the noun following *bei* can be legitimately omitted from N*bei*NV in the language (resulting in N*bei*V similar to the English truncated passives), but the noun after *ba* in N*ba*NV cannot be omitted. Thus, listeners could readily assign an object role to the first N in N*bei*V, using fragment interpretation based on the complete form N*bei*NV. Finally, for *bei*NVN, the choices started at random when only *bei* was heard, but quickly moved to a first-noun choice. This pattern was the opposite of that for *ba*NVN, because the fragment interpretations of *bei*N (N as agent) and VN (N as patient) were consistent, unlike the two nouns in *ba*N and VN, which competed for patienthood in *ba*NVN.

Experiment 2 points to the usefulness of the gating method in unraveling the process of sentence comprehension as the auditory stimulus unfolds in time. Gating allows one to see the temporal structure involved in this process: for example, it shows how choice decisions move along the course of the sentence. Gating also permits the evaluation of listeners' processing of a sentence on the basis of partial or fragment information for that sentence: for example, a given sentence such as *ba*NVN is associated with multiple fragment interpretations (as shown in parentheses below), including *ba*N(O), NV (SV), NVN (SVO), and VN (VO), and gating provides a means of probing into these interpretations.

Experiment 2 suggests that Chinese listeners rely strongly on the integration of fragment information in sentence processing. Daily spoken Chinese contains many sentences that are either incomplete or ungrammatical when literally translated into Indo-European languages, due to frequent omissions of subject and object and a flexible word order. For instance, both *ba*NV and VN can stand alone as complete sentences with omitted subject, and NV as a complete sentence with omitted object (SV) or subject (OV). Thus, Chinese listeners have to rapidly incorporate partial sentence information rather than to wait for "complete sentences" to occur, whenever fragment interpretations become available. For example, in this experiment, the fragments *ba*N and *ba*NV were interpreted as part of the complete N*ba*NV, and N*bei* and N*bei*V were interpreted as part of the complete N*bei*NV. Moreover, the results with *ba* also indicate that the grammatical violation produced by the wrong positioning of the marker does not disrupt comprehension, showing that Chinese listeners are less sensitive to grammaticality and can interpret ungrammatical sentences by reference to grammatical models. This contrasts with sentence processing in English, in which grammaticality plays a more important role (Bates, 1991; von Berger, Wulfeck, Bates, and Fink, 1996). These properties of processing in Chinese reflect language-specific properties of the language, in which morphological, grammatical, and syntactic constraints on sentence constituents are very weak.

The above experiments were concerned with how Chinese speakers determine the global relationships among different sentence constituents (nouns and verbs).

The next two experiments are concerned with how Chinese speakers identify locally ambiguous items, homophones, in a sentence.

4. EXPERIMENT 3: CROSS-MODAL NAMING

4.1. Method

4.1.1. Participants

Thirty native Cantonese Chinese speakers from Hong Kong participated in this experiment.

4.1.2. Materials and Design

Thirty spoken homophone nouns were selected, each with at least two different meanings that share the same syllable and the same tone. Each homophone was embedded in two different sentences, one biasing either of the two selected meanings. A separate group of 20 speakers was asked to judge the degree of constraint of the prior context on the target homophone. They were given the test sentences with the prior context but without the homophone, and were asked to fill in a word that naturally completes the sentence. Their responses were scored on a 1 to 4 scale, depending on how close the word that they chose matched the homophone (Marslen-Wilson and Welsh, 1978). A high-constraint score (mean score 1.6) was derived from these responses, indicating that the sentence contexts were highly constraining.

Three independent variables were manipulated in this experiment.

1. *Probe type.* The visual probe was (a) biased, which was related to the contextually biased meaning of a homophone, or (b) unbiased, which was related to the second meaning not biased by the context, or (c) unrelated control.
2. *Dominance.* Half of the prior contexts biased the dominant (more frequent) meaning of the homophone, and the other half the subordinate meaning (less frequent). The frequency information was based on Ho and Jiang (1994).
3. *Homophone density.* Half of the homophones had many competitors (four or more alternative meanings—high density) and the other half had few competitors (two to three alternative meanings—low density).

The 30 participants were randomly assigned to six groups of five. Each group received an equal number of sentences in the 3 (probe type) × 2 (dominance) × 2 (homophone density) design. The order of presentation was counterbalanced across participants.

4.1.3. Procedure

The test sentences were read by a native Cantonese speaker and were digitized into a Macintosh computer with a sampling rate of 22 kHz. The presentation of auditory and visual stimuli was controlled by the PsyScope program (Cohen et al., 1993). Naming latencies were recorded by the CMU button-box with a voice-activated relay (Cohen et al., 1993). A cross-modal naming task (cf. Seidenberg, Tanenhaus, Leiman, and Bienkowski, 1982; Swinney, 1979) was used as follows. Participants saw a fixation point, and immediately heard on a pair of headphones the sentence in which the homophone was embedded. A visual probe that was either related (i.e., biased or unbiased) or unrelated to the homophone occurred at the offset of the homophone.[4] The participant's task was to, as accurately and quickly as possible, name the visual probe aloud.

4.1.4. Data Analysis

The dependent variable was participants' response latencies to each visual probe. The latency was measured from the onset of the visual probe to the onset of the vocal response (see Li and Yip, 1996, for a detailed discussion).

4.2. Results and Discussion

Figure 3 presents the naming latencies as a function of probe type, dominance, and homophone density. A 3 × 2 × 2 (probe type × dominance × homophone density) ANOVA indicated a main effect of probe type ($F(2, 28) = 8.84, p < .01$) and an interaction between probe type and dominance ($F(2,58) = 4.62, p < .05$). No other effects were significant.

First, the main effect of probe type shows that the naming of the visual probes differed in speed as a function of whether the probe was related to a contextually biased, unbiased, or totally unrelated meaning of the homophone. This result pro-

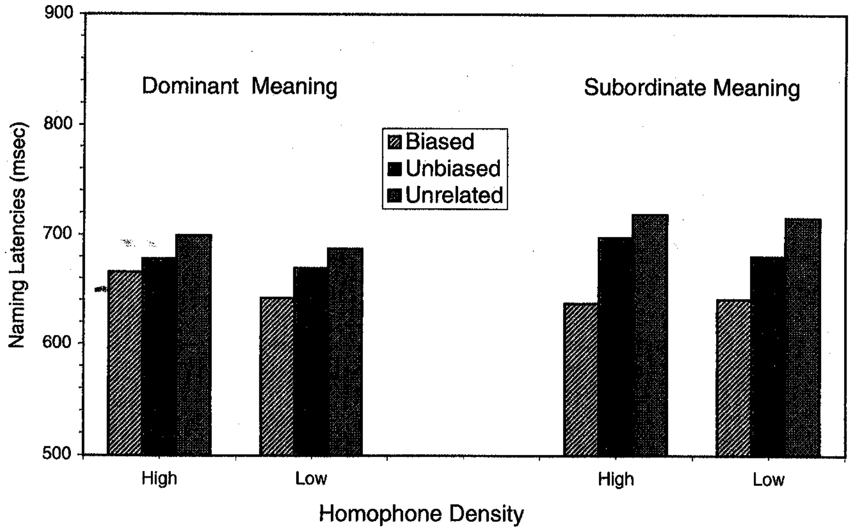

Figure 3. Naming latencies as a function of probe type, dominance, and homophone density. (From Li and Yip, 1996. Reprinted with permission of Lawrence Erlbaum Associates, Inc.)

vided evidence for the operation of context effects, since a contextually biased meaning elicited the fastest response to the visual probe, as compared with the unbiased and the unrelated meaning. Second, the interaction between probe type and dominance was due to the pattern that dominant (more frequent) meanings were in general accessed faster than subordinate (less frequent) meanings, but when the context biased the homophone meaning (i.e., in the biased probe condition), dominant and subordinate meanings did not differ ($F(1,59) = 2.65$, ns). This pattern indicates that context and frequency can mutually interact during the identification of homophone meanings, and that context may be sufficiently constraining to act early on to reduce the effect of the frequency of homophone meanings. Finally, the absence of a homophone density effect shows that the number of competing meanings in a homophone did not affect listeners' responses, suggesting that the context effect operated early enough, probably within the acoustic boundary of the homophone, to select the contextually appropriate meaning and suppress the inappropriate meanings.

Our results show that Chinese speakers are sensitive to the contextually biased meaning at an early stage, at least immediately following the occurrence of the homophone (i.e., the acoustic offset of the word). This context effect occurs much earlier than what has been previously assumed (e.g., about 1.5 sec following the occurrence of the ambiguous word, e.g., as shown in Onifer and Swinney, 1981). It seems that Chinese listeners, to cope with the extensive ambiguity created by massive homophones, can rapidly disambiguate alternative meanings as they hear a homophone in a sentence. Our results are consistent with the context-dependency hypothesis, which argues that ambiguous meanings of a word may be selectively accessed at an early stage according to prior sentential context (Simpson, 1981; Simpson and Krueger, 1991; Tabossi, 1988).

In a recent study Moss and Marslen-Wilson (1993) argued that the offset of an ambiguous word may not be the critical point for tapping into the locus of context effect, because many words in context could be recognized before the acoustic offset of the word. They suggested that the initial access and selection of contextually appropriate meaning may occur earlier than the word offset. Since the cross-modal method in Experiment 3 tapped the recognition process only at the offset of the homophone, in Experiment 4 I used a word gating-method (Grosjean, 1980) to examine how early context effects can be observed within the acoustic boundary of the homophone.

5. EXPERIMENT 4: WORD GATING

5.1. Method

5.1.1. Participants

Eighteen native Cantonese Chinese speakers from Hong Kong participated in this experiment. None had participated in Experiment 3.

vided evidence for the operation of context effects, since a contextually biased meaning elicited the fastest response to the visual probe, as compared with the unbiased and the unrelated meaning. Second, the interaction between probe type and dominance was due to the pattern that dominant (more frequent) meanings were in general accessed faster than subordinate (less frequent) meanings, but when the context biased the homophone meaning (i.e., in the biased probe condition), dominant and subordinate meanings did not differ ($F(1,59) = 2.65$, ns). This pattern indicates that context and frequency can mutually interact during the identification of homophone meanings, and that context may be sufficiently constraining to act early on to reduce the effect of the frequency of homophone meanings. Finally, the absence of a homophone density effect shows that the number of competing meanings in a homophone did not affect listeners' responses, suggesting that the context effect operated early enough, probably within the acoustic boundary of the homophone, to select the contextually appropriate meaning and suppress the inappropriate meanings.

Our results show that Chinese speakers are sensitive to the contextually biased meaning at an early stage, at least immediately following the occurrence of the homophone (i.e., the acoustic offset of the word). This context effect occurs much earlier than what has been previously assumed (e.g., about 1.5 sec following the occurrence of the ambiguous word, e.g., as shown in Onifer and Swinney, 1981). It seems that Chinese listeners, to cope with the extensive ambiguity created by massive homophones, can rapidly disambiguate alternative meanings as they hear a homophone in a sentence. Our results are consistent with the context-dependency hypothesis, which argues that ambiguous meanings of a word may be selectively accessed at an early stage according to prior sentential context (Simpson, 1981; Simpson and Krueger, 1991; Tabossi, 1988).

In a recent study Moss and Marslen-Wilson (1993) argued that the offset of an ambiguous word may not be the critical point for tapping into the locus of context effect, because many words in context could be recognized before the acoustic offset of the word. They suggested that the initial access and selection of contextually appropriate meaning may occur earlier than the word offset. Since the cross-modal method in Experiment 3 tapped the recognition process only at the offset of the homophone, in Experiment 4 I used a word gating-method (Grosjean, 1980) to examine how early context effects can be observed within the acoustic boundary of the homophone.

5. EXPERIMENT 4: WORD GATING

5.1. Method

5.1.1. Participants

Eighteen native Cantonese Chinese speakers from Hong Kong participated in this experiment. None had participated in Experiment 3.

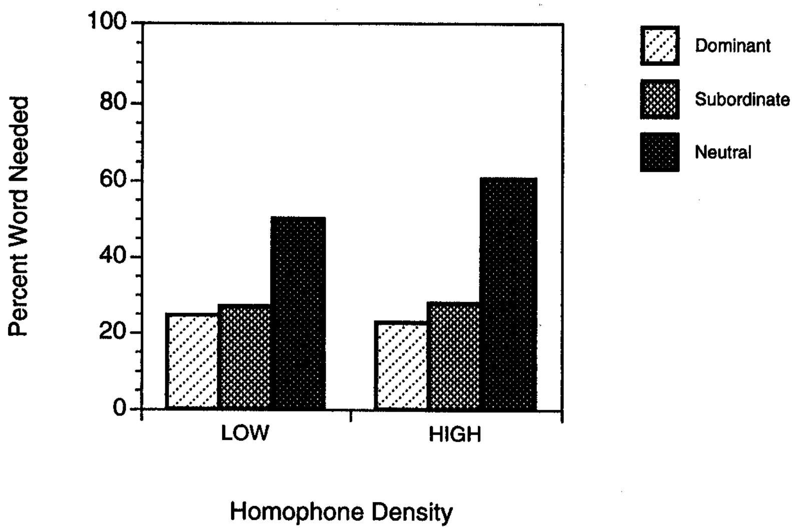

Figure 4. Percent word needed for the identification of the homophone, as a function of context and homophone density.

A 3 × 2 (context × homophone density) ANOVA revealed main effects of both context ($F(2,34) = 51.28$, $p < .01$) and homophone density ($F(1,17) = 13.15$, $p < .05$), and a significant interaction between the two ($F(2,34) = 14.29$, $p < .01$).

As seen in Figure 4, the main effect of context came from the difference between the semantically neutral context and the biased contexts. A post hoc test (Tukey HSD) revealed no difference between the two biased contexts, but differences between the neutral context and either of the biased contexts. The main effect of homophone density was due to the difference between the low- versus high-density homophones in the neutral context ($F(1,17) = 16.93$, $p < .01$); the other two contexts showed no difference in homophone density. The result that the homophone density effect existed only in the semantically neutral context but not in the semantically biased contexts also accounted for the interaction between context and homophone density. The interaction is consistent with the absence of homophone density effects in Experiment 3, in which I suggested that context effects can operate early to select only the contextually appropriate meanings and suppress the inappropriate meanings.

These results indicate that when the context is semantically biased toward a specific meaning of the homophone, Chinese speakers can identify the appropriate meaning with a minimal amount of acoustic-phonetic information. On the average (averaging over the two biasing conditions), they needed only 25% of the homophone's acoustic signal. Note that 25% does not provide sufficient acoustic information about the status of the word. The processing system thus must be contextually driven, relying on top-down information to identify the word on the basis of only a quarter of the word's information. The results match up well with established estimates of word recognition times in English: Grosjean (1980) and Marslen-Wilson (1987) showed that when spoken-word recognition takes place in context, only half or even less of the acoustic information of a word is needed for correct identification. Our results also indicate that, consistent with the context-dependency hypothesis, the access of the appropriate meaning of the homophone can take place well before the acoustic offset of the homophone.

6. GENERAL DISCUSSION

This chapter presents results from a number of experimental studies of sentence processing in Chinese. In Experiment 1, I examined several basic cues that are important to Chinese speakers' identification of sentence roles, and the ways in which these cues interact to determine the final outcome of the identification process. The results suggest that Chinese speakers use a weighted array of linguistic cues to guide sentence role assignment, including noun animacy, word order, and grammatical markers. In Experiment 2, I studied interpretations associated with sentence fragments at various temporal locations of the sentence, and showed that the rapid integration of fragment interpretations and the interactions between these interpretations characterize much of sentence processing in Chinese. These processing characteristics are described with reference to the weak morphological and syntactic constraints in Chinese. The results also suggest that Chinese speakers show less sensitivity to grammatical violations than do English speakers. In Experiments 3 and 4, I examined how Chinese speakers resolve extensive homophone ambiguities within a sentence. The results suggest that Chinese speakers rely on sentential context at an early stage to disambiguate homophone meanings, at least at the acoustic offset of the homophone (according to the cross-modal naming method), and possibly earlier (according to the gating method). The results are consistent with an interactive, context-dependency hypothesis but inconsistent with a modular, exhaustive-access hypothesis.

A seemingly mysterious question that initiated my inquiries into Chinese sentence processing is this: how can Chinese speakers ever understand each other, if the language does not have clear inflectional markings (in contrast to many Indo-European languages), lacks a stable word order (in contrast to English), involves a high degree of ellipsis (to the extent contrary to any languages known), and has a high degree of ambiguity with lexical items (again to the extent contrary to any languages known). The fact is, that Chinese speakers do not seem to have more trouble understanding each other than do speakers of English or other Indo-European languages. That mysterious question seems to be less mystical now, given that Chinese speakers do have in stock a number of grammatical devices (though not always available, such as *ba* and *bei*), strategies to cope with variable word orders (default SVO and pragmatic cues to other orders), strategies to rely on semantic (such as noun animacy) and pragmatic information (context in which the sentence is uttered), strategies to integrate incomplete or ungrammatical constructions on the basis of complete and grammatical models, and strategies to rely on prior sentential context to resolve lexical-morphemic ambiguities. The use of these strategies and interpretations in real-time sentence processing suggests interactive processes in which multiple sources of information (i.e., syntactic, semantic, and contextual) compete and collaborate in sentence processing (Marslen-Wilson, 1987; McClelland, Rumelhart, and the PDP Research Group, 1986; MacWhinney and Bates, 1989).

Crosslinguistic variations can reveal important aspects about sentence processing, clearly in unraveling the specific strategies used to cope with language-specific properties, such as those associated with grammatical morphology, word order, ellipsis, and homophone ambiguities in Chinese. For example, to deal with the high degree of homophony, Chinese speakers rely strongly on contextual information in disambiguating various meanings in a homophone. To deal with the high degree of syntactic variability, they rely strongly on using fragment interpretations and semantic and pragmatic information; they depend less on syntactic or grammatical relations and show less sensitivity to grammatical violations. These processing patterns run at odds with traditional emphasis: linguistic theories have assigned a central role to grammaticality and syntactic structure. Meaningless sentences like "colorless green ideas sleep furiously" are used to demonstrate that it is possible for the speaker to judge the grammaticality and understand the structure of a sentence without retrieving its meaning (Chomsky, 1957). Our studies seem to suggest the opposite: Chinese speakers can retrieve the meaning of a sentence without heavily relying on grammatical relations or being disturbed by grammatical violations, and partial structures or fragments lend themselves to probabilistic interpretations based on complete and grammatical models in the language.

ACKNOWLEDGMENTS

The preparation of this chapter was supported by a Faculty Research Initiation Grant from the School of Arts and Sciences, University of Richmond. Partial support for the preparation of Experiments 1 and 2 was provided by the Human Frontier Science Program to the author and by NIH/NIDCD (#R01-DC00216) to Liz Bates, and Experiments 3 and 4 by the Chinese University of Hong Kong to the author and Michael Yip. I would like to thank Liz Bates, Francois Grosjean, and Brian MacWhinney for insightful discussions on the materials presented here. I also thank Hua Liu, Yeung Lai Ping, Wong Wing, Edward Yang, and Michael Yip for their assistance to the experiments.

REFERENCES

Bates, E. (1991). *On-line studies of sentence processing and the competition model.* Unpublished manuscript, University of California, San Diego.

Bates, E., and MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner and L. Gleitman (Eds.), *Language acquisition: the state of the art.* New York: Cambridge University Press.

Bates, E., and MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum.

Bates, E., and MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney and E. Bates (Eds.), *The crosslinguistic study of sentence processing.* New York: Cambridge University Press.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., and Smith., S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition,* 11, 245–299.

von Berger, E., Wulfeck, B., Bates, E., and Fink, N. (1996). Developmental changes in real-time sentence processing. *First Language,* 16, 193–222.

Chen, S., Tzeng, O., and Bates, E. (1990). Sentence interpretation in Chinese aphasia. In *Proceedings of the Tenth Meeting of the Second Language Research Forum,* Department of Linguistics, University of Oregon, Eugene, Oregon.

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton and Co., Publishers.

Clynes, D., and MacWhinney, B. (1990). *The ECS software.* Department of Psychology, Carnegie Mellon University.

Cohen, J., MacWhinney, B., Flatt, M., and Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, and Computers,* 25, 257–271.

Cotton, S., and Grosjean, F. (1984). The gating paradigm: A comparison of successive and individual presentation formats. *Perception and Psychophysics,* 35, 41–48.

Frazier, L. (1987). Sentence processing: A tutorial overview. In M. Coltheart (Ed.), *Attention and performance* (Vol. XII). Hillsdale, NJ: Erlbaum.

Frazier, L. (1990). Exploring the architecture of the language-processing system. In G. Altmann (Ed.), *Cognitive models of speech processing.* Cambridge, MA: The MIT Press.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics,* 28, 267–283.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes,* 3, 233–274.

Ho, H.-H., and Jiang, Y.-H. (1994). *Word frequency in Hong Kong in the 90's.* Research Institute of Humanities, Chinese University of Hong Kong.

Institute of Linguistics, The Academy of Social Sciences. (1985). *Xianda Hanyu Cidian* [Modern Chinese Dictionary]. Beijing: Commercial Press.

Li, C., and Thompson, S. (1981). *Mandarin Chinese: A functional reference grammar.* Berkeley: University of California Press.

Li, P. (1994). Understanding the time course of sentence comprehension: A sentence gating study in Mandarin Chinese. In H.-W. Chang, J.-T. Huang, C.-W. Hue, and O. Tzeng (Eds.), *Advances in the study of Chinese language processing* (Vol. I). Taipei: National Taiwan University Press.

Li, P. (1996a). The temporal structure of spoken sentence comprehension in Chinese. *Perception and Psychophysics*, 58, 571–586.

Li, P. (1996b, November). *Spoken word recognition of homophones in Chinese.* Poster presented at the Thirty-seventh Annual Meeting of the Psychonomic Society, Chicago, 1996.

Li, P. (1996c). Spoken word recognition of code-switched words by Chinese-English bilinguals. *Journal of Memory and Language*, 35, 757–774.

Li, P., Bates, E., Liu, H., and MacWhinney, B. (1992). Cues as functional constraints on sentence processing in Chinese. In H. C. Chen and O. Tzeng (Eds.), *Language processing in Chinese.* Series of Advances in Psychology 90. Amsterdam: Elsevier Science Publishers.

Li, P., Bates, E., and MacWhinney, B. (1993). Processing a language without inflections: A reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language*, 32, 169–192.

Li, P., and Yip, M. (1996). Lexical ambiguity and context effects in spoken word recognition: Evidence from Chinese. In G. Cottrell (Ed.), *Proceedings of the 18th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Li, P., and Yip, M. (in press). Context effects and the processing of spoken homophones. *Reading and Writing: An Interdisciplinary Journal.*

Lu, J-M. (1980). Hanyu kouyu jufa-li -de yiwei xianxiang [On reversed sentence patterns in spoken Chinese]. *Zhongguo Yuwen* [The Chinese Language], 1, 28–41.

MacWhinney, B., and Bates, E. (Eds.). (1989). *The crosslinguistic study of sentence processing.* New York: Cambridge University Press.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71–102.

Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In G. T. Altmann (Ed.), *Cognitive models of speech processing.* Cambridge, Mass.: The MIT Press.

Marslen-Wilson, W. D., and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29–63.

McClelland, J., Rumelhart, D., and the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. II). Cambridge, Mass.: The MIT Press.

Miao, X-C. (1981). Word order and semantic strategies in Chinese sentence comprehension. *International Journal of Psycholinguistics*, 8, 109–122.

Moss, H., and Marslen-Wilson, W. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1254–1276.

Onifer, W., and Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, 9, 225–236.

Seidenberg, M. S., Tanenhaus, M. J., Leiman, J. M., and Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 14, 538–559.

Simpson, G. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior*, 20, 120–136.

Simpson, G. B., and Krueger, M. A.(1991). Selective access of homograph meanings in sentence context. *Journal of Memory and Language*, 30, 627–643.

Small, S., Cottrell, G., and Tanenhaus, M. (1988). *Lexical ambiguity resolution: perspectives from psycholinguistics, neuropsychology, and artificial intelligence.* San Mateo: Morgan Kaufmann Publishers.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645–659.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language*, 27, 324–340.

Tyler, L., and Wessels, J. (1985). Is gating an on-line task? Evidence from naming latency data. *Perception and Psychophysics*, 38, 217–222.

NOTES

[1] The sentence-role assignment task differs from sentence-processing tasks involved in several other models of sentence processing (e.g., Frazier, 1987) in which speakers read locally ambiguous structures in a sentence. For this reason, it was often difficult to directly compare our model with other models. Moreover, our studies have focused on the spoken modality whereas most other models focus on the visual modality; note that it is debatable whether visual and auditory processes are identical (Marslen-Wilson, 1990).

[2] A fourth type of sentence was also tested in this experiment: sentences with the indefinite marker *yi*. (See Li et al., 1993, for details.)

[3] For the NVN and NNV sentences, there was an initial bias toward first-noun choice. That is, upon hearing only the first noun, listeners found it easier to select the first noun as the agent than to select another noun that had not yet appeared. The bias was toward what has been heard, in the absence of an alternative form in the input.

[4] A condition in which the visual probe occurred at the onset of the homophone was also tested originally. The results are not reported here because we suspect that the auditory signal of the homophone may cause interferences to listeners' naming of the visual probe that occurs at the onset of the homophone (also Grosjean, personal communication).

[5] The sentence context variable here is comparable to the dominance variable in Experiment 3. Sentence context makes more sense than dominance in this experiment because of the addition of the neutral context that biased neither meaning of the homophone.