ADVANCES IN THE STUDY OF CHINESE LANGUAGE PROCESSING, Volume 1

UNDERSTANDING THE TIME COURSE OF SENTENCE COMPREHENSION: A SENTENCE GATING STUDY IN MANDARIN CHINESE

Ping Li

Department of Psychology

The Chinese University of Hong Kong

This study uses a sentence gating task to investigate the time course of sentence comprehension in Chinese. Experiments with several sentence types (simple, *ba*, and *bei*) suggest that listeners can, like in word recognition, identify the meaning of a sentence before they have the complete acoustic-phonetic information. More importantly, the early or late comprehension of a sentence is constrained by the particular cue configurations across the course of the sentence. If the cues are consistent with the listener's prediction (based on patterns of specific linguistic knowledge) at each point in time, they lead to faster responses and early comprehension; inconsistent and competing cues lead to slower responses and late comprehension. Our study confirms the predictions of the Competition Model regarding the relationship between cues and response times in sentence comprehension. The similarities and differences between word recognition and sentence processing are also discussed in connection with our data.

One of the basic findings in research on spoken word recognition is that listeners can recognize a word before they have received all the acoustic-phonetic information that distinguishes that word from others in the lexicon (Grosjean, 1980; Marslen-Wilson, 1987). What about sentence recognition? Intuitively, one could speculate that a similar conclusion would hold for parsing and comprehension at the sentence level. That is, there may be circumstances which permit listeners to complete assignment of sentence roles before they have received all the acoustic-phonetic information that distinguishes that sentence from other possible sentences in the language. Unfortunately, there has been very little research that would allow us to support or reject this conjecture. This study is an attempt to provide some empirical evidence on this issue for Chinese, a language that has fewer options for inflectional marking but permits a large range of sentence types than are possible in English.

Not only are we interested in *whether* listeners can assign semantic roles before the end of the sentence, more importantly, we are also interested in the issues of *when* and *how* they do so. The issue of *when* has to do with the time course of sentence comprehension – specifically, at what point(s) in time across the course of the sentence can listeners assign sentence roles? The issue of *how* has to do with the factors that act and interact to shape sentence comprehension at different points in time.

Experiments relevant to these issues have been carried out within a framework for cross-linguistic research called the Competition Model (Bates & MacWhinney, 1982, 1987; MacWhinney and Bates, 1989). Based on results from studies of sentence processing in more than a dozen languages, this model emphasizes the extent to which languages can vary in the way that cues compete and converge to determine meaning. A cue, in this context, is a particular piece of information that the listener can use to determine the relationship between meaning and form. The Competition Model assumes an interactive process in which the mapping between surface forms and underlying meanings is mediated by competitions and collaborations among cues. Cues can be evaluated with respect to their validity, i.e., their information value for the identification of linguistic functions. In any given language, the overall validity of a cue is a joint product of its availability (how often the cue is present when a given interpretation has to be made) and its reliability (when the cue is available, how often it leads to the right answer). Having different cue validities, different cues cooperate and compete in the comprehension process, resulting in different interpretation patterns in different languages.

Most previous studies within the Competition Model have adopted a sentence comprehension task in which native speakers of different languages are presented with simple sentences in their own languages and are asked to identify the agent (the actor or doer of the action) of the sentence. For example, in the sentence *The pencil kicks the elephant*, native English speakers choose the *pencil* much more often than the *elephant*, while native Chinese speakers show the opposite strategy, choosing *elephant* as the agent regardless of word order (MacWhinney and Bates, 1989; Li, Bates, & MacWhinney, 1993). This finding is compatible with the cue validity principle, because Chinese permits far more word order variation than English, and because the sentence subject at preverbal position can often be omitted, thus reducing the reliability of word order cues. Hence word order information is a very strong cue to the agent role in English, but not as strong as animacy in Chinese.

The Competition Model also yields a number of relatively clear predictions regarding the relationship between cue configurations and decision time. Some of the predictions have strong implications for our understanding of the time course of sentence comprehension: (1) There should be a monotonic relationship between cue strength and speed of response (i.e., stronger cues will lead to faster decision times; weaker cues will be associated with slower decision times); (2) Converging cues should facilitate sentence comprehension and thus lead to faster decision times; (3) Competing cues should inhibit immediate comprehension and thus slow decision times down; (4)

Prediction 1 will interact with predictions 2 and 3, so that a very strong cue may still result in relatively fast decision times, despite competition from weaker sources of information. These predictions have been examined and confirmed in our recent study of sentence interpretation in Chinese (Li, Bates, and MacWhinney, 1993). That study investigated the cues that Chinese speakers employ in real-time sentence interpretation, in the absence of inflectional morphology. Complex patterns of interaction among cue types were observed in both the choice and the decision time data, reflecting the principles of competition and convergence proposed within the Competition Model. The results indicated that Chinese speakers rely heavily on semantic cues (e.g., animacy of nouns), which is consistent with previous untimed studies of sentence interpretation in Chinese (e.g. Miao, 1981; Miao, Chen, and Ying, 1986; Li, Bates, Liu, & MacWhinney, 1992). Chinese speakers also make important use of the grammatical markers, especially the passive marker *bei*. Word order appeared to be a somewhat volatile piece of information, compared with word order cues in other languages. The absolute effect of the strengths of word order varies with context: it can be either reduced by other more important cues (e.g., animacy and the passive marker) or magnified when no other important information is available. Our finding suggests that cues do not necessarily fall into clusters of linguistic types within the hierarchy of cue importance to sentence meaning (e.g., all word order cues before all morphological cues, and vice versa). The results were compatible with interactive activation models of sentence comprehension, while posing problems for models that assume a modular architecture in which morphological, semantic, and syntactic sources of information are insulated from one another at various points in parsing and interpretation.

The reaction time data in Li et al (1993) attest to the utility of reaction time studies at the sentence level, since they can reveal aspects of processing that are often not available or obscure in results from choice response measures. However, the picture-choice method adopted in that study was not particularly suitable for tapping into the time course issues of sentence comprehension, e.g., at which point in time a correct interpretation is reached. While our study and several others (Miao, 1981; Miao et al, 1986; Li et al, 1992; Li et al, 1993) have provided converging evidence suggesting that animacy dominates word order cues in competition, and that grammatical devices such as *ba* and *bei* are important in sentence comprehension in Chinese, it is not yet clear how these cues interact across the course of the sentence. The current study uses a sentence gating technique to evaluate comprehension at separate points of the sentence by asking the listener to identify the agent of the sentence on the basis of sentence fragments as well as full sentences. Specifically, we are interested in the activation and interaction of different cues at each point in time during the course of sentence processing. For example, when does the listener identify a sentence role for a particular type of sentence? What determines whether the listener will identify a sentence meaning early or late? How do different cues interact to jointly contribute to the final interpretation of the sentence? In order to provide answers to these questions, we have concerned ourselves in particular with structural information of the sentence, i.e., 'global' syntactic information such as word order cues and 'local' grammatical information such as the semi-morphological markers *ba* and *bei*.

Method

In line with our earlier work (Li et al., 1993), the experiment presented below involved systematic variation of important cues to sentence meaning in Chinese, including word order, the object marker *ba*, and the passive marker *bei*. We did not vary the indefinite marker *yi*, which proved to be a relatively weak cue in our earlier study. And we did not vary animacy in this particular experiment, because the number of permutations required to test comprehension of successive sentence fragments is very large and because the relative importance for animacy vs word order is clear from our and others' studies. The task that has been selected for this experiment is a sentence gating task, a sentence-level adaptation of the gating method pioneered by Grosjean (1980) and others for the study of word recognition (Cotton & Grosjean, 1984; Tyler, 1984; Tyler & Wessels, 1983). Although Grosjean (1980) had pointed out the possibility of the method for sentence processing, there has not been much research which used this method for sentence-level study. Recently, we applied the gating method to a grammaticality judgment task with bilingual English-Chinese speakers (Liu, Bates, & Li, 1992), and the results there proved encouraging: the sentence gating method was used successfully to locate the temporal dimensions for which subjects make a grammaticality-judgment decision in bilingual processing.

In the word-level gating task, subjects are presented with fragments of a word, one at a time, in increasing length, until the whole word has been presented. For example, the word *trespass* may be presented as *t, tr, tre, tres, tresp, trespa*, and *trespass*; subjects are required to guess what the possible word will be at every fragment. In our experiment, sentences were similarly gated into smaller fragments (albeit on a word-by-word basis); subjects hear increasingly longer fragments each time until the whole sentence has been presented. For example, if the sentence was *Xiaogou chidiao xiaoya* (the dog eats the duckling), subjects would hear the following in sequence and need to respond each time: *Xiaogou* (dog), *Xiaogou chidiao* (dog eat), and finally *Xiaogou chidiao xiaoya* (dog eat duckling). Subjects were not asked to identify the sentence (a impossible task since the number of sentences are infinite in any given language), but instead to determine which constituent represents the agent of the sentence (the actor or doer of the action). The task of determining the agent of the sentence is a task that requires the listener to understand the relationships between sentence constituents and assign a semantic role, which is different from other tasks such as ' word-monitoring' where the focus of the task is on individual words (e.g., Marslen-Wilson & Tyler, 1980). Similarly to the word-monitoring task, however, this method allows us to probe into comprehension processes in finer detail at different temporal locations of the sentence, and it has advantage over the simple picture choice task in which only full sentences were presented, as in Li et al. (1993).

Subjects

Twenty native adult Mandarin Chinese speakers from mainland China participated in this experiment (11 females and 9 males ranging in age from 23 to 36 years, mean age=29). They were students or visiting scholars at the Chinese University of Hong Kong during the time of our experiment. Most of the subjects were from the Northern-dialect areas in China (i.e., the Mandarin-speaking areas). All of them use Mandarin Chinese intensively in daily life, and only one subject could speak Cantonese.

Materials

Three independent variables were manipulated systematically in our experiment: (a) Sentence type, divided into simple sentences with no markers, sentences with the object marker *ba*, and sentences with the passive marker *bei*; (b) Word order, arranged in NNV, NVN, and VNN sequences; (c) Fragment length, from the first fragment of the sentence (i.e., the first word) to the complete sentence with all fragments. Within the *ba* and *bei* sentences, the position of the marker was also varied (first word of the sentence, before the verb, and after the verb). The crossing of these variables yielded a total of 21 sentence types. To compensate for the imbalance in number between simple sentences vs *ba* and *bei* sentences, we used eight tokens for each type of simple sentences, and four tokens for each type of *ba* and *bei* sentences. Each sentence was then gated into different fragments on a word-by-word basis. The total number of the resulting fragments and sentences was 396. Our computer program generated random combinations of nouns and verbs for each test sentence. The nouns and verbs used for the test sentences are given in Appendix I. A complete list of the sentence types and their fragments is given in Appendix II.

Experimental Apparatus

All test sentences were recorded on a digital audio tape by a female native Chinese speaker who was unaware of the experimental purpose. The sentences were read at a normal rate and with a smooth and flat intonation. They were then digitized and segmented at a sampling rate of 22 kHZ on a Macintosh IIsi, using the SoundEdit program. During playback, subjects heard the sound materials through an amplified speaker.

The CMU button box is a device that registers responses and reaction times for pushing built-in buttons or for other external inputs. It utilizes a crystal oscillator producing time measurements accurate to one millisecond. It has three buttons mounted in a row on a sloped-front box and connects to the modem port of the Macintosh machine. The present experiment used only the left and the right buttons for registering responses and response times. The middle button was used by subjects to rest their index finger before and after each response.

The experiment was run on a Macintosh IIsi model. Pictures were digitized with an AST Turbo scanner and displayed on a high resolution RGB monitor. Each picture was displayed in a 7 x 11-cm frame. The presentation of pictures and sound stimuli was controlled by the PsyScope software, a general-purpose program for psychological experiments developed by Cohen, MacWhinney, Flatt, and Provost (1993).

Task

As subjects heard a sentence fragment played back on the speaker, they simultaneously saw on the computer screen a pair of pictures that corresponded to the two animals described in the sentence. Their task was to determine, as quickly and accurately as possible, which of the two pictures (appearing on the left or the right side of the screen, counterbalanced) represented the agent in the sentence. They indicated this choice by pressing the left or the right button on the button box. The onset of the sentence fragment started the button box timer for subjects response times to that fragment. Subjects were given a maximum of 2 seconds to respond after the sentence fragment had been played. This amount of time was sufficient to allow full response for most subjects under most of the conditions, while putting time pressure on response speed. There was a 2-second silence with a blank screen before the next sentence began to play.

Procedure

All subjects were tested individually in a dimly lit room. Before the experiment began, the experimenter explained the task to the subject in Chinese. Subjects were instructed to press the button as soon as they made a decision according to their understanding of the sentence fragments. They were told that in some cases the sentence might sound a little odd, but that they still needed to select one of the two animals as the actor. It was made clear that there was no right or wrong answer, and the subject only needed to pay attention to the sentence and make decisions.

The experiment consisted of four blocks of testing (99 sentence fragments in each) and was divided into two sessions with two blocks in each session. Each block contained randomized combination of different sentence types with different orders in successive fragmentation. Subjects were brought in for the first session and a few days later for the second session. In the first session, they were asked to practice with the button pressing, and then practice with a warm-up task in which 18 fragments from 5 sentences were given. In the second session, the same warm-up test was practiced before the experiment. In order to make sure that subjects were not satiated with the test sentences, they were given a five-minute break in the middle of each session . Each session took about 30 minutes.

Results and analyses

The dependent variable for our study is subjects' response times (RTs) to each of the sentence fragments. Our computer program calculated RTs from the beginning of the fragment when it was played on the speaker to the time when subjects pushed the button. The RTs were averaged across all responses regardless of which N was chosen as the agent. These raw RTs cannot be used directly for our analysis, however, since different sentence fragments have different lengths in time. A new set of RTs were derived from subtracting the raw RTs from the time lengths of each of the sentence fragments and were then used for subsequent statistical analysis. In the new RTs, a positive value indicates that a response was made with a given amount of time after the end of the sentence fragment has been heard, whereas a negative value indicates that the response was made sometime earlier relative to the end of the sentence fragment.

The Earliness of Sentence Comprehension

Consistent with our speculation at the beginning, a first important finding is that as in the recognition of words, listeners can and indeed do in many cases, comprehend a sentence before they have heard the complete acoustic-phonetic signal.[1] Figure 1 plots the mean response times for simple sentences, ba sentences, and *bei* sentences when the full sentence is presented. It shows that except for the simple sentences and *ba* sentences in the NNV order, all other sentences elicited responses before the end of the sentence. For a simple NVN sentence with three disyllabic words (the average length of these sentences is 1747 msec in our experiment), subjects responded some 288 msec before the end of the sentence. If we consider a correction factor of 100 msec (see footnote 1), then what the results mean is that subjects made a decision about the meaning of the sentence somewhere around the middle point of the third word (the third word lasts about 686 msec on the average). This decision point for other types of sentences were in general later, but mostly before the end of the sentence.

The results for the *ba* and *bei* sentences in Figure 1 were averaged across different positions of occurrence of the marker. When we break down further to examine the response times for various subcategories, we found that some sentences led to even faster responses. The *bei* NVN sentences, in which the marker occurs at the initial position, elicited a response some 493 msec before the end of the sentence, which is one-third into the sentence if we consider a correction factor of 100 msec. For these sentences, subjects would have made an internal choice just as they began to hear the last noun of NVN.

---

1. By 'comprehend' we mean that subjects have made an internal decision about the meaning of the sentence. It is important to note that the internal comprehension of sentences should be faster than the response times registered from a button pressing. Classical mental chronometry research tells us that it takes some time from the internal mental decision to the execution of responses for a given cognitive task. Marslen-Wilson (1987) hypothesizes that such a time lag for word recognition would be in the rang of 50-75 msec. The time lag for sentence comprehension should be even longer then.

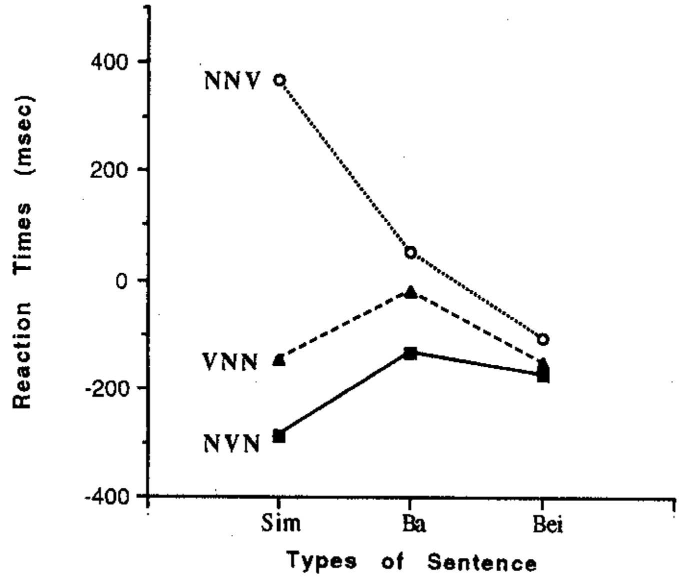

Figure 1. Mean reaction times for three types of sentences with different orders and markers

Obviously, not all sentences elicited fast responses to the same extent, and the results varied as a function of both word order and sentence types. First, we found a significant main effect of word order ($F(2,38)=30.80$, $p<.001$). Apparently, NNV sentences had much slower responses than the other two orders. Simple NNV sentences were responded to some 371 msec after the end of the sentence, which is 659 msec later (about the length of a word) than the counterpart NVN sentences (the lengths of these two types of sentences were about the same; in fact, NNV sentences were even shorter than the NVN sentences, i.e., 1710 msec vs. 1747 msec). An answer to this discrepancy may be readily found in our discussion of word order cues in Chinese (Li, Bates, and MacWhinney, 1993). As also mentioned above, NVN order is the canonical word order in Chinese that corresponds to a default SVO interpretation, whereas NNV order is highly marked in its semantics and pragmatics, and can be associated with either an OSV interpretation or an SOV interpretation. The indeterminacy between the two interpretations of NNV in the absence of a grammatical marker may well have led to the slowest response of simple NNV strings. Note that when there is a *ba* or *bei* marker, response speeds would be much facilitated; we will return to the effects of *ba* and *bei* in more detail later.

Second, we found a significant main effect of the type of sentences ($F(2,38)=7.31$, $p<.001$). Figure 1 shows that the *ba* sentences elicited in general slower RTs than the simple and the *bei* sentences (except the simple NNV sentences). This difference may be due to the fact that the function of *ba* as an grammatical marker is constrained by various syntactic and semantic factors. In contrast to *bei* whose uniform function is to passive the sentence, *ba* as an object marker is associated with definiteness and resultative verb compounds. The lower validity of *ba* as a pure object marker thus could not facilitate the comprehension process uniformly as the passive marker *bei* could. This result is consistent with our findings on *ba* and *bei* in Li et al. (1993).

Finally, there is also a significant interaction between word order and sentence types (F(4,76)=20.11, p<.001). The major source of the interaction comes from the slower responses for NNV orders in *ba* and especially simple sentences vs. the other relatively faster responses.

Speed of Response and the Length of Fragments

As in word recognition, listeners have to start with the acoustic-phonetic information in the sentence; they cannot identify the meaning of that sentence until some minimal amount of information is available. This bottom-up priority in lexical and sentential processing has been emphasized in Marslen-Wilson and Tyler (1980) and Marslen-Wilson (1987). But exactly how much information do they need? In our experiment, the subject receives increasingly more information in a successive order, and the decision as to which noun is the agent of the sentence would be made only when there is sufficient information available. In Figure 2, we can see that the response speed to each fragment point of the sentence decreases as a smooth function of how much information is present. In all cases, subjects responded faster than they did at an earlier point in time when they had only incomplete sentence fragments. There is a significant main effect of the length of fragments (F(2,38)=96.72, p<.001) as well as a main effect of word order (F(2,38)=26.40, p<.001). The more information there is, the faster the responses. Full sentences elicited faster response than when there were only two fragments, and two fragments faster than only one fragment.

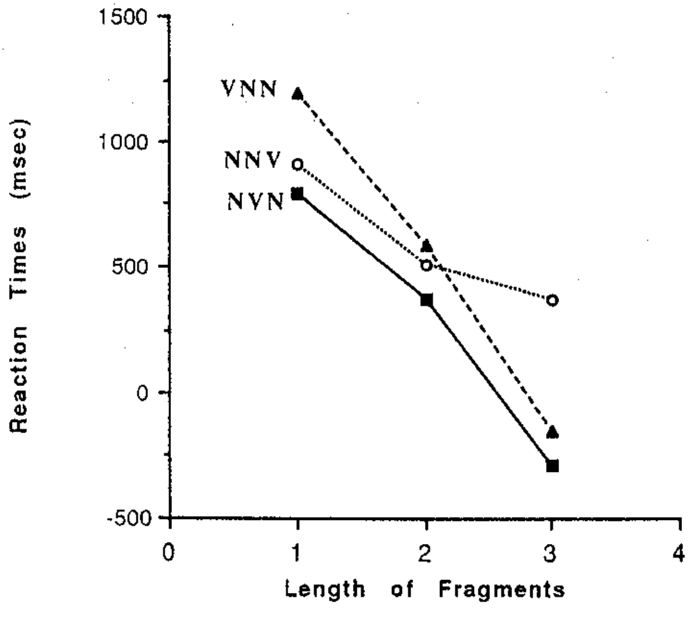

Figure 2. Mean reaction times for simple sentences with different lengths of fragments

The same results for simple sentences also hold for sentences with *ba* and *bei* markers. Figure 3 shows the mean response times for *ba* and *bei* sentences, averaged across word order conditions. There are significant main effects of the length of fragments in both cases ($F(3,57)=99.52$, $p<.001$ for *ba*; $F(3,57)=89.78$, $p<.001$ for *bei*). What these results suggest is that as the sentence unfolded in time, subjects became more and more confident in making a choice about the sentence, and thus responded more and more rapidly in the task. It also shows that for the average *ba* and *bei* sentences, subjects do not wait until they have the complete acoustic-phonetic information; instead, they respond shortly before or just around the end of the sentence. More acoustic-phonetic information leads to faster responses across the course of the sentence, but the final decision does not necessarily require that the information is complete. This is consistent with our discussion above about the earliness of sentence comprehension.

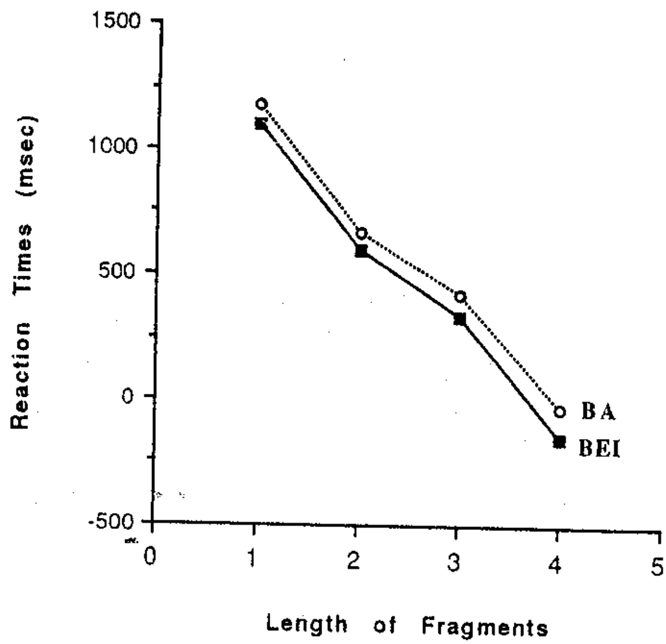

Figure 3. Mean reaction times for *ba* and *bei* sentences with different lengths of fragments

Although the general patterns of decrease in response times are similar across different boundaries of the sentence in Figure 3, detailed analysis reveals that the general patterns do not always hold for individual sentences. In Figure 4, the results of VNN sentences with the *ba* marker occurring before the last noun (i.e., VN*ba* N) are plotted against those with the bei marker occurring at the same position (i.e., VN*bei* N). It can be immediately noticed that the two markers had very different effects on the sentences. The results differed significantly as a function of the marker ($F(1,19)=51.12$, $p<.001$) as well as the length of the fragments ($F(3,57)=59.73$, $p<.001$).

The major source of the difference between VN*ba* N and VN*bei* N was in the last fragment. For the VN*bei* N sentences, there was a continuous and sharp decrease of response times as more fragment information became available, but for the VN*ba* N sentences, the decrease of response times took a different journey at the position where the *ba* marker occurred. Whereas the VN*bei* N sentences elicited a response on the average 453 msec before the end of the complete sentence, the VN*ba* N sentences not only elicited no responses before the end, but had slowed-down responses as compared with the incomplete fragments.

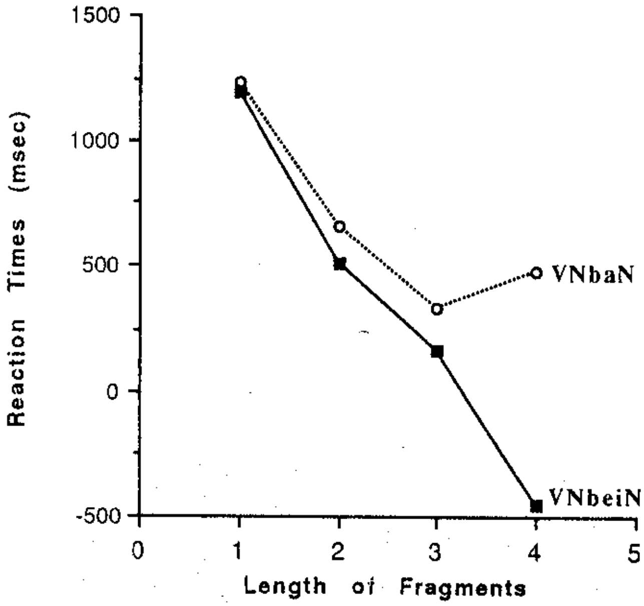

Figure 4. Mean reaction times for VNN sentences with *ba* and *bei* markers

The VN*ba* N and VN*bei* N sentences had about the same lengths (1778 vs. 1735 ms), but resulted in quite different response speeds (the final response times differed by 938 ms, about the time for the first two words in VNN). This discrepancy brings us to the question of what determines the speed of response at various time points of the sentence, and what determines whether listeners will or will not be able to comprehend the sentence before they have the complete acoustic-phonetic information.

Predicting What's Next

In natural language comprehension, listeners may be unconsciously engaged in the task of guessing what comes next (Grosjean, 1980; Elman, 1990; Kutas & Hillyard, 1988). In our experimental task, this ability of guessing may be more pronounced so that each time before the sentence fragment was heard, subjects may have made a prediction as to what is the most probable fragment next, based on his or her understanding of the previous fragments. As the sentence unfolded from one fragment to its completion acrossdifferent points, the fragments that are consistent with subjects' guessing would elicit faster responses, while the fragments that are inconsistent and incoherent would lead to slower responses.

On what grounds do listeners base their predictions? The Competition Model proposes that they use the individual cues and cue configurations provided by their language. Moreover, it is the interactions (competition and collaboration) among cues that determine the decisions and decision times. Earlier, we have discussed the relationship between cues and response speed, and we have also seen in the example of NNV sentences the role of specific cue configurations in sentence interpretation. The slower response times for NNV sentences, as compared with the faster responses to NVN sentences, were caused (or so we believe) by the conflict between an OSV interpretation and an SOV interpretation in the language. Now, how is the listener's prediction governed by cues? In the case of NVN, when subjects have heard the first two fragments of NV, they may well predict the next word to be N since they could use the canonical word order cue and interpret the whole string as SVO.[2] If the actual word turns out to be the predicted one, the response would speed up, as in the results of our experiment (see Figure 2). In contrast, in the case of NNV, when subjects encountered the first two fragments NN, they had two possible interpretations, OS or SO, and even if they could predict the next word to be V, among other choices (*ba* or *bei*), they could not readily interpret the status of the two nouns by means of canonicity of word order or other cues. That is why we did not see in Figure 2 a much facilitated response for NNV sentences at the end, compared with NVN and VNN.

A similar logic can be applied to the data in Figure 4. As demonstrated in Li et al (1993), VN is a very strong cue in Chinese for the identification of the object; in fact, the VO interpretation for VN strings is even stronger than SV for NV strings in Chinese (for example, in Figure 2, VN fragments elicited a sharper decrease of about 200 msec more than the NV fragments). When subjects encountered a single V, they could immediately predict that the coming word (if any) will be a noun that takes the object role (although the reader should keep in mind that VS fragments can occur in Chinese in marked cases, see Lu, 1980 -- hence this is a probabilistic rather than an absolute prediction). When subjects had actually heard the VN fragments, they would have already interpreted the noun after the verb as the patient, according to the strong VO cue. That is why we saw in Figure 4 a sharp decrease of response times for the VN phrases in both VN*ba*N and VN*bei*N sentences, because it is consistent with subjects' prediction. Next, when subjects heard the third word of the sentence, which was the *ba* or the *bei* marker, they would not know what to do with it because they could not predict its occurrence according to the cue configurations of the language (VN cannot be followed by a grammatical marker). However, because the *ba* or the *bei* marker is not a constituent to be assigned a semantic role, it did not interfere with the existing interpretations that subjects had reached, and thus there

2. Note that our experimental conditions permitted more options than just N after NV, because both *ba* and *bei* could occur after NV.

was still a facilitation in response times. At this point, subjects may have predicted the next word to be a noun, since the functions of these markers are to mark the immediately following noun. But the effect of the marker did not show up until subjects had actually heard the last noun after the marker, at which time the VO word order cue of the previous fragments would interact with the *ba* or the *bei* phrase, causing the discrepancy that we saw in Figure 4. In VN*bei*N, the VO interpretation for the first two words was consistent with the agenthood of the *bei* phrase, thus there was a great facilitation and subjects responded well before the end of the sentence. In contrast, in VN*ba* N, the VO interpretation was incompatible with the objecthood ofthe *ba* phrase, thus there was some reason to delay immediate understanding of the whole string. To put it in Competition Model terms, the first case represents a convergence between the word order cue and the grammatical marker (the two cues point to the same direction for agenthood), while the second case presents a competition between the two cues (they point to different directions for role assignment). As the competition principles would predict, converging cues facilitate sentence interpretation and lead to faster response times, while competing cues inhibit immediate comprehension and slow response times down. The results here are completely consistent with the findings of Li et al. (1993) regarding the relationship between cues and responses times.

Similar findings can be observed with various sentence types. Figure 5 presents another example in which the occurrence of the *bei* marker, unlike in Figure 4, hindered sentence comprehension.[3] In the simple NVN sentences, response times decreased as more fragments were heard across the course of the sentence, whereas in the NV*bei*N sentences, response times increased after the occurrence of the *bei* marker.[4] This increase is probably due to the conflict between the SVO interpretation for the NVN strings and the agenthood of the *bei* N phrase. Note that unlike in the VN*ba* N/VN*bei* N sentences where the effect of the grammatical marker was seen only after the noun appeared, the *bei* marker in this case had its effect immediately. This is perhaps because VN strings usually come at the end of a sentence, while NV strings usually have more to follow (especially in the case of a transitive verb, like in our experiment). If sentence comprehension does involve a series of predictions of this kind, then the presence of the *bei* marker after V should greatly disrupt the comprehension process, reflected in a marked slowing of reaction times.

Finally, let us compare simple sentences with *ba* and *bei* sentences to identify the role of these markers in sentence comprehension. In Li et al (1993), we found a robust effect of the bei marker for the identification of the sentence agent, and a significant but weaker effect of the *ba* marker for the identification of the sentence object. Earlier, we discussed the functions of NNV sentences in Chinese and their correlates in sentence comprehension, as seen in Figure 1. In Figure 6, we further plotted

3. In this graph, response times to the last noun have been stretched to the fourth position of the length of fragments in order to see the effects of *bei* clearly on the third position.

4. The faster responses to the first two words in NV*bei* N vs. those in NVN may be due to the difference in the lengths of the original sentence (the former is about 100 msec shorter than the latter).

simple NNV sentences against N*bei* NV and N*ba* NV sentences, the orders in which the *ba* and *bei* markers occur naturally in the language (Chao, 1968; Li & Thompson , 1981).

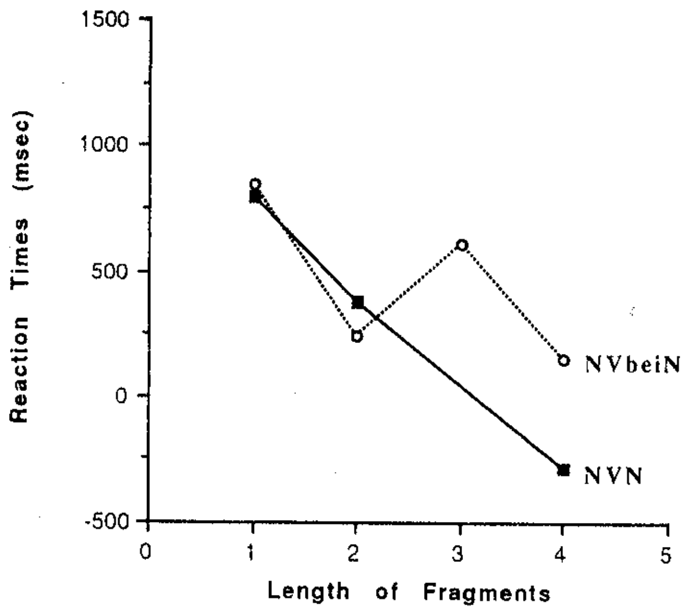

Figure 5. Mean reaction times for simple NVN sentences and NV*bei* N sentences

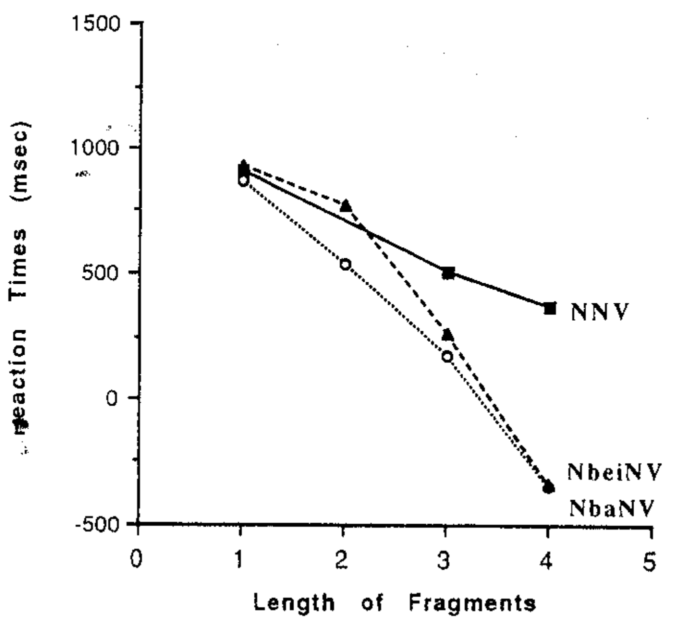

Figure 6. Mean reaction times for NNV sentences with and without markers

Although a direct statistical comparison is not possible for the data in Figure 6 because of the different number of fragments involved in the sentences, it is clear that the *ba* and *bei* markers played an important role in sentence comprehension in Chinese. In Figure 6, the indeterminacy between OSV and SOV interpretations for the simple NNV strings were greatly reduced when the *ba* or the *bei* marking occurred before the second noun. In other words, the marker served to disambiguate the two possible interpretations and indicate whether the second noun is the agent (by *bei*) or the patient (by ba). As seen in Figure 6, the response times for the second noun in N*bei* NV and N*ba* NV sentences were about 300 msec faster than that in the simple NNV sentences without the marker. This facilitation was certainly due to the occurrence of the marker. Moreover, the facilitation spread further to the full sentence, leading to an early response of some 340 msec before the end of the sentence (about 700 msec faster than the simple NNV!). After the subjects had heard N*ba* N or N-*bei* N strings, they might well predict the last fragment to be a verb, which was indeed the only possibility.

To summarize the results with the RT data, we have found significant main effects of word order and length of fragments, and interaction effects between the two. There is clearly evidence for the idea that listeners can comprehend a sentence before all the acoustic-phonetic signal is received. For a normal NVN sentence, subjects are able to make a choice on the sentence around the middle point of the last noun. For NNV and VNN sentences, comprehension is later. In general, the more acoustic-phonetic information is available, the faster subjects are able to make their choices. The early or late comprehension of sentences is constrained by specific cue configurations that are associated with specific interpretations in a given language (Chinese in our case). When the association is consistent with the listener's prediction, responses are facilitated; when the association is inconsistent with the listener's prediction or is hard to predict, responses are slowed down. We have also found that the *ba* and *bei* markers contribute significantly to sentence interpretation in Chinese, consistent with our findings elsewhere (Li, Bates, & MacWhinney, 1993). These markers interact with word order cues to determine the time course of sentence comprehension: when the cues converge, they lead to early comprehension; when the cues compete, comprehension is delayed.

General discussion

This study uses a sentence gating task to investigate the time course of sentence comprehension in Chinese. Experiments with several sentence types (simple, *ba*, and *bei* sentences) suggest that listeners can, like in the process of word recognition, identify the meaning of a sentence before the complete acoustic-phonetic information has become available. More importantly, the early or late comprehension of a sentence is constrained by the particular cue configurations across the course of the sentence. If the cues are consistent with the listener's prediction (which is based on patterns of specific linguistic knowledge) at different points in time, they lead to faster responses and early comprehension; if the cues are inconsistent with the listener's prediction, they lead to slower responses and late comprehension. Moreover, if several cues converge to point to the same direction for meaning, sentence comprehension is facilitated; if several cues compete with each other and point to different directions for meaning, sentence comprehension is slowed down. Our study further confirms the predictions of the Competition Model regarding the relationship between cues and response times in sentence comprehension. In general, the early comprehension of a sentence is either due to the salience of a particular cue (such as the canonicity of word order cue and the strong VO interpretation for VN strings), or due to the collaboration between cues (such as the converging cues of word order and grammatical marker in VN *bei* N sentences).

The test sentences and fragments in our experiment contain a mixture of grammatical and ungrammatical sets because of the crossing of different variables. While it is well accepted that researchers can use nonwords to tap into the process of word recognition, it is still highly disputable about the nature of ungrammatical sentences in studies of sentence comprehension. Some readers might argue against the use of ungrammatical and/or incomplete sentences on grounds that these sentences can invoke 'non-linguistic' or 'supra-linguistic' (all in all, 'abnormal') processing strategies. We believe that our subjects use the same processing strategies for ungrammatical and incomplete as well as grammatical and complete sentences. For the incomplete sentences, for example, subjects may be engaged in a fragment analysis that is similar for both grammatical and ungrammatical sentences. McClelland and Rumelhart (1981), using the example of Glushko (1979), shows that the presentation of a nonword such as MAVE can activate real words that share parts of the orthography/pronunciation such as MAKE, MALE, HAVE, GAVE, and SAVE. We can likewise argue that an ungrammatical sentence such as *ba* NVN can activate interpretations associated with various fragments such as *ba* NV, NV, NVN, VN, etc. The final interpretation of the string will depend on the activation level of each fragment, and on the competition and collaboration between various fragments. Previous work by Bates, MacWhinney and colleagues (see MacWhinney & Bates, 1989) shows that people have very systematic and language-specific ways of interpreting ungrammatical and incomplete sentences, due no doubt to their degree of overlap with available/similar grammatical sentences. Our work on sentence interpretation in bilingual English-Chinese speakers also indicates that subjects can clearly separate grammaticality from interpretation (Liu, Bates, & Li, 1992); in other words, subjects may judge a particular sentence to be ungrammatical, but they could still process the sentence in the normal way, using native comprehension strategies that are relevant to their particular languages.

It may be useful to point out here that Marslen-Wilson and Tyler (1980), even a decade ago, argued for the principle of "obligatory bottom-up processing of all knowledge sources" for word recognition and sentence comprehension. That is, during the processing of words and sentences, when one information source is not available, others are still operating to analyze the speech signal. For example, when a meaningful interpretative analysis is not possible, the structural and lexical knowledge sources will still function. This principle can explain why subjects could deal with sentence materials in random word order in Marslen-Wilson and Tyler's experiment. That implies that the processing system does not break down in case of abnormal sentences. *In other words, abnormal sentences do not necessarily induce abnormal analysis.* The processor has no way of treating ungrammatical sentences differently from grammatical ones *a priori*, unless the relevant components of the sentence have been processed. According to Marslen-Wilson and Tyler, even though during processing there might be times when an analysis of the input is delayed because of the inherent ambiguity or indeterminacy, "there is nothing intrinsic to the structure of the processing system itself that would prevent immediate on-line interpretation."

Although we have made comparisons between sentence comprehension and word recognition in various parts of the paper, it is important to note that there are fundamental differences between the two. When we say 'comprehension of a sentence', we mean very different things as we say 'recognition of a word'. To recognize a word is to access that word in the mental lexicon and select its unique phonological shape and semantic content, which is distinguished from any other word in the language. To comprehend a sentence, however, is to identify the relationships between different sentence constituents (e.g., the relationships between nouns and nouns, nouns and verbs); in other words, it is to identify the meaning which is conveyed by the structural arrangement of the sentence constituents (e.g., to understand who did what to whom). There is no such thing as a unique identity of a sentence, like the unique identity of a word, since the number of sentences in any given language is infinite, in contrast to the finite set of words in a language. This difference determines that our results argue for no mechanism such as a 'cohort activation' of sentences in sentence comprehension, since the cohort set will be too large for the processing system to manage. However, much like in word recognition processes as advocated by the Cohort Model (Marslen-Wilson, 1987), in sentence comprehension, there is also parallel activation of all relevant information sources, i.e., all relevant fragments that share the parts with the target sentence. In sum, during word recognition and sentence comprehension, the processing system continuously accesses the acoustic input and activates relevant information sources in parallel, likewise for both levels; however, it evaluates the input against a set of possible word candidates for word recognition (because there is a limited set of lexical items), but it predicts what comes next for sentence comprehension on the basis of cue configurations of the target language.

It should also be noted here that although we found early comprehension of sentences in many cases, this earliness is in no way directly comparable to the earliness of word recognition. According to Grosjean (1980) and Marslen-Wilson and Tyler (1980), listeners are able to identify a word in normal context when perhaps only half or less of the acoustic-phonetic information of the word has been heard. Whereas in sentence comprehension, as discussed here, subjects do not make their decisions before they have heard a sufficient amount of sentence fragments, i.e., two-thirds of the sentence. In the most rapid case, subjects responded on the average 493 msec before the end of the sentence, which means that they comprehended the sentence just as they started to hear the third word (and if we consider a correction factor). This difference between words and sentences in recognition probably indicates that in addition to putting the fragments together, the processor has more work to do for sentences than for words in accessing the meaning of the whole from the parts (e.g., lexical, syntactic, and semantic analyses).

Finally, we would like to point out that the current study includes only bottom-up information for the comprehension process (i.e., sentence fragments of increasing length), but not top-down information (i.e., context in which the sentence occurs). It would be interesting to see how sentence comprehension proceeds as a function of the interaction between bottom-up and top-down information, as what has been found in word recognition studies. Future research is needed to address this issue.

Acknowledgments

This research was supported by the Incentive Research Fund and the Croucher Foundation Fund from the Chinese University of Hong Kong. I am very grateful to Elizabeth Bates and Chen Hsuan-Chih for their constructive comments on an earlier version of the paper, to Tan Lee Yong, Wong Wing, Robert Buffington, and Larry Juarez for their assistance to the experiments. I would also like to thank MacWhinney and colleagues for their development of the PsyScope experimental software. Please send correspondence to: Ping Li, Department of Psychology, The Chinese University of Hong Kong, Shatin, NT, Hong Kong. E-mail: pingli@cuhk.hk.

References

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

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

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

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.

Chao, Y-R. (1968). *A grammar of spoken Chinese.* Berkeley: University of California Press.

Cole, P. (1987). Null objects in universal grammar. *Linguistic Inquiry,* 18, 597-612.

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

Elman, J. (1990) Finding structure in time. *Cognitive Science,* 14, 179-211.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance,* 5, 674-691.

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

Huang, J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry,* 15, 531-574.

Kutas, M., & Hillyard, S.A. (1988). Contextual effects in language comprehension: studies using event-related potentials. In F. Plum (Ed.) *Language, communication, and the brain.* New York: Raven Press.

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

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

Li, P., Bates, E., & 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.

Liu, H., Bates, E., & Li, P. (1992). Sentence interpretation in bilingual speakers of English and Chinese. *Applied Psycholinguistics,* 13, 451-484.

Liu, H., Bates, E., & Li, P. (1992). Grammaticality judgment in Chinese-English bilinguals: A gating experiment. In *Proceedings of the 14th Annual Conference of the Cognitive Science Society,* 95-99, Indiana University, Bloomington.

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.

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

Marslen-Wilson, W., & Tyler, L. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.

McClelland, J. & Rumelhart, D. (1981) An interactive activation model of context effects in letter perception. *Psychological Review*, 88, 375-407.

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

Miao, X-C., Chen, G., & Ying, H. (1986). Sentence comprehension in Chinese. In M. Zhu (ed.), *Studies in child language development*. Shanghai: East China Normal University Press.

Sun, C-F. (1991). Transitivity and the *ba* construction. Paper presented at the 3rd North American Conference on Chinese Linguistics, Cornell University, Ithaca, May.

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

Zhu, D-X. (1982). *Yufa jiangyi*. (Introduction to Grammar). Beijing: Commercial Press.

Appendix I: Nouns and Verbs Used for The Test Sentences

Nouns

banma 'zebra', *daxiang* 'elephant', *daishu* 'kangaroo', *gongji* 'cock', *gouxiong* 'bear', *haitun* 'dolphin', *houzi* 'monkey', *hudie* 'butterfly', *jinyu* 'goldfish', *laohu* 'tiger', *laoshu* 'mouse', *mianyang* 'sheep', *qier* 'penguin', *xiaogou* 'dog', *xiaoma* 'horse', *xiaomao* 'cat', *xiaoniao* 'bird', *xiaoniu* 'cow', *xiaotu* 'rabbit', *xiaoya* 'duckling', *xiaozhu* 'pig', *wugui* ' turtle';

Verbs

chi-diao 'eat-up', *da-bai* 'hit-defeat', *fang-zou* 'let-go', *gan-pao* 'drive-run', *ti-dao* 'kick-down', *tui-kai* 'push-away', *yao-zhu* 'bite-stop', *zhua-zhu* 'seize-stop', *zhuang-dao* 'bump-down'.

Appendix II: A Complete List of the Test Sentences and Fragments

Type	Fragments and Sentences
SIMPLE	N......NV.....NVN
	N......NN.....NNV
	V......VN.....VNN
BA	ba..... baN.....baNV......baNVN
	ba..... baN.....baNN......baNNV
	ba..... baV.....baVN......baVNN
	N......Nba.....NbaV......NbaVN
	N......Nba.....NbaN......NbaNV
	V......Vba.....VbaN......VbaNN
	N......NV.....NVba......NVbaN
	N......NN.....NNba......NNbaV
	V......VN.....VNba......VNbaN
BEI	bei.... beiN.... beiNV..... beiNVN
	bei.... beiN.... beiNN..... beiNNV
	bei.... beiV.... beiVN..... beiVNN
	N......Nbei.... NbeiV..... NbeiVN
	N......Nbei.... NbeiN..... NbeiNV
	V......Vbei.... VbeiN..... VbeiNN
	N......NV..... NVbei..... NVbeiN
	N......NN..... NNbei..... NNbeiV
	V......VN..... VNbei..... VNbeiN