Inhibitory Processes of Chinese Spoken Word Recognition

Michael C.W. Yip

Department of Educational Psychology, Counselling and Learning Needs
The Hong Kong Institute of Education, Hong Kong SAR

ABSTRACT

The present study was designed to examine the inhibitory processes of spoken word recognition of Chinese homophones during sentence comprehension, by using a cross-modal naming experiment. In this experiment, listeners were asked to name aloud a visual probe as fast as they could after hearing a sentence, which ended with a spoken Chinese homophone. Results confirm that preceding sentential context has exerted an early effect on disambiguating among different alternative meanings of the ambiguous word. Secondly, the contextually inappropriate meanings of the ambiguous word would be inhibited rapidly during sentence comprehension. Thirdly, the present results also demonstrated that the inhibitory mechanism could sustain to a longer duration following the occurrence of the ambiguous word.

1. INTRODUCTION

Investigation to the sentential context effects in spoken word recognition has always been the focal point for the research of lexical ambiguity resolution [1-9]. The fundamental question is the time-course of the context effects operated during different various stages of lexical access? So far, answers for this question lead to two major contrastive hypotheses. One is the exhaustive access hypothesis, which argues that all meanings of an ambiguous word will be accessed momentarily following the occurrence of the ambiguous word. The other one is the context-dependency hypothesis, which argues that only the contextually appropriate meaning of an ambiguous word will be accessed momentarily following the occurrence of the ambiguous word, and the sentential context can only help to select the appropriate meaning at a late post-access stage. This hypothesis assumes that language processing is a modular, bottom-up process in which non-lexical, contextual information does not penetrate lexical access [10].

Empirical examinations for the above hypotheses have been widely researched in different languages (Chinese, Dutch, English, and Italian) during the last three decades with different patterns of results. However, only a few studies were systematically examined the inhibitory processes of the different alternative meanings of the ambiguous word thereafter [12]. Previous studies concluded that the contextually inappropriate meanings of the ambiguous word would be suppressed actively and rapidly, within 200 milliseconds, following the occurrence of the ambiguous word with a view to maintain a normal sentence comprehension [13]. In addition, relevant experimental study also demonstrated that the inhibitory processes of those unselected or those contextually inappropriate meanings would still be sustained for a period of time after our human language processor selected the contextually appropriate meaning, might be longer than a second [14]. The present study attempts to investigate this rarely addressed question by using Chinese as a testing case. Chinese is a language that differs significantly from most other languages (e.g., in its use of lexical tones, its morphemic monosyllabicity) and lexical ambiguity is pervasive in Chinese (in terms of its extensive homophony). Therefore, it offers many unique and interesting psycholinguistic properties in its phonological, lexical, and syntactic structures [15] to crucially investigate the issue. From the Modern Chinese Dictionary [16], about 80% of the monosyllables (with tonal differentiation) in Chinese are ambiguous between two different meanings, and about 55% have five or more alternative meanings. For example, a single Cantonese syllable si1 has up to 14 meanings (e.g., teacher, lion, silk, corpse, poem, private, think, etc.) [17]. Then, upon hearing the syllable si1 in a highly semantically constraint sentence, do native Cantonese listeners activate only the most contextually appropriate meaning out of the total 14 meanings and inhibit the other remaining 13 meanings of the single syllable simultaneously [18]? What would be the inhibitory processes thereafter? Would there be any other lexical variables affected the inhibitory processes at the same time? How long the inhibition can be sustained?

In the present study, we attempt to answer these aforementioned questions, by using a cross-modal naming experiment, in order to seek new evidence for making an accurate and complete picture on the relationship between the sentential context effects and other lexical factors operated in processing the spoken language.

2. EXPERIMENT

2.1. Method

Participants. Fifty-six native Cantonese speakers (32 male and 24 female, mean age =23.1) who reported no speech or hearing deficits participated in this experiment. All participants were students at the Open University of Hong Kong. They took part in the experiment on a voluntary basis.

Materials and design. Thirty spoken Cantonese homophones were selected, each with at least two different meanings in the same tone (syllables with different tones are not considered homophones in the present study). Each homophone was embedded in four different types of sentences, in a two (Dominance: dominant vs. subordinate) by two (Reiteration: reiterated vs. not reiterated) mixed design (see the below...
examples). The four conditions are: (DR) The first part of the preceding sentential context will bias to the dominant meaning of the homophone, which occurred at the middle of the sentence; and the second part of the context will be changed to bias the subordinate meaning of the same homophone, which occurred once again at the end of the sentence. (SR) The first part of the preceding sentential context will bias to the subordinate meaning of the homophone, which occurred at the middle of the sentence; and the second part of the context will be changed to bias the dominant meaning of the same homophone, which occurred only at the end of the sentence. (SN) Preceding sentential context will bias to the subordinate meaning of the homophone, which occurred only at the end of the sentence.

Altogether, two main variables were manipulated in this experiment:

(1) Dominance: The visual probe was related to either the dominant meaning of the spoken homophone, or to the subordinate meaning of the spoken homophone. The dominance information was based on Ho and Jiang’s study [19].

(2) Reiteration: The spoken homophones were either reiterated in the middle position of the whole sentence or not.

An example si1 (silk/poem) in the four different types of sentential context is given below.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Contextual Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>(DR)</td>
<td>Sai3 lo5 lam5 ye5 lam5 do3 tau4 faat3 do1 hei2 saa3 si1 tai2 lai4 kui5 gam1 yan6 do1 get2 laan4 yok3 do2 yat1 sau2 si1. Young brother is thinking very hard that his hair seems like gradually becoming silk, I do not think that today he can finally compose a poem.</td>
</tr>
<tr>
<td>(SR)</td>
<td>Dong1 ga1 je1 hai5 su1 fong2 luo5 min6 yan4 gan2 sau2 si1 ma1 me4 jau6 tui3 joh2 siw6 fong2 d06 tong3 gan2 san1 maai5 di1 si1. When elder sister is reading aloud a poem inside her studying room, mother at the same time, in her bedroom, is ironing her newly-bought silk.</td>
</tr>
<tr>
<td>(DN)</td>
<td>Lei1 gaan1 uk1 yi5 ging1 fong1 fai3 joh2 ho2 doh1 lin4 cheung4 gok3 sueng6 min6 jung6 yau5 ho2 doh1 jil1 jue1 hai5 do6 to3 gan2 si1. The house has been neglected for so many years, and there are so many spiders at the corner that are spinning the silk.</td>
</tr>
<tr>
<td>(SN)</td>
<td>Wai6 joh2 ga1 sam1 dui3 jung1 gok3 man4 hok6 ge1 ying6 sil1 ngo5 lei1 pai4 b001 joh2 ho2 doh1 sau2 ho2 chu1 min5 ge3 si1. In order to deepen my understanding of traditional Chinese literature, I recently recite a lot of famous poem.</td>
</tr>
</tbody>
</table>

The four visual probes to si1 in these sentences are: saam1 “shirt” (SR and DN), and man4 “prose” (DR and SN).

From the construction of these materials, we set out the following predictions for the present study, they are: the response latencies of condition (DN) are faster than (SN) as well as (SR) are faster than (DR) due to a robust dominance effect. In addition, the response latencies of condition (DN) are faster than (SR) as well as the response latencies of condition (SN) are faster than (DR) due to the inhibitory effects of the inappropriate meanings of the ambiguous words at the first occurrence. The rationale is simply because of that by comparing the response latencies to conditions (SR) and (DN), we can see that the activation level of the inappropriate meaning of the ambiguous word (only the contextually appropriate meaning is accessed when hearing the ambiguous word at the first instance in the middle of the sentence) will gradually be decreased. The inhibition of the inappropriate meaning will be reflected by slower response latencies than the time observed in DN condition (as no inhibition to the inappropriate meaning of the ambiguous word has been occurred beforehand). If this is really the case, it will certainly confirm the inhibitory effects actually take place during the sentence processing and also these effects can sustain to a longer duration.

Following the procedure of relevant studies [20], a separate group of 20 native Cantonese speakers was asked to judge the degree of constraint of the preceding sentential context on the homophone. They were given all the testing sentences with the preceding biasing context but without the homophone, and were asked to fill in the word. They were told to think of a Chinese word that would naturally complete the sentence. Their responses were scored on a 1–4 scale, based on the scale proposed by Marslen-Wilson and Welsh [21]: 1 was given for a word identical to the test word, 2 for a synonym, 3 for a related word, and 4 for an unrelated word. Responses were pooled across the 20 judges, and the mean rating was 1.8. This score was above the high constraint condition in Marslen-Wilson and Welsh’s study [21]. An effort was also made to have the sentence context of equal length, and the average length of the test sentences, counting the homophones, was 26 words (ranging from 24 to 27 words). In addition, we have made an effort to eliminate any kind of intra-sentential priming from other individual words within the whole sentence frame when constructing the sentence context as much as possible.

Other than the sentential context, we carefully selected the appropriate visual probes as follows. First, all the visual probes were based on a semantic relatedness norm experiment from another separate group of one hundred native Cantonese speakers [20]. In this simple experiment, the participants were asked to immediately think of three Chinese words that have the same or closely related meaning to each homophone, and the mostly frequent words they listed were used to be the related visual probe for each spoken homophone. All the visual probes in each experimental condition were carefully matched with the same category of initial phonemes and individual frequency information.

**Experimental Apparatus.** All the test sentences were read by a female native Cantonese speaker at a normal conversation speed and tape-recorded digitally in a SONY MD. All the spoken sentences were then transformed and digitized into a Macintosh PowerBook. A computer program called PsyScope [22] controlled the presentation of the materials. A microphone, which was used to register listeners' vocal responses and hence calculated the naming latencies, was connected to the notebook computer. A remote controlled SONY IC-recorder was also used and controlled by the experimenter in another partition of the experimental room to check for accuracy.

**Procedure.** Each participant randomly received an equal number of sentences (fifteen) for each condition in the 2
(Dominance) x 2 (Reiterated) mixed factorial design. The order of presentation for the sentences was pseudo-randomly arranged such that the visual probes did not consecutively bias spoken homophones. The order of presentation was counterbalanced across all participants.

All participants did the experiment individually in a quiet experimental room. Before the experiment, the experimenter explained the task in Cantonese to the listener. First, they were told that they would be hearing a sentence through a pair of headphones, and then at the end of the sentence, they would see a Chinese character (visual probe) appeared on the computer screen. Their task was simply to name aloud the Chinese character into the microphone as quickly and as accurately as possible. Before the actual experiment began, they were given a practice session in which they heard a set of separate but similar sentences. The whole experiment took about twenty minutes.

### 2.2 Results

Mean response latencies, counting from the onset of the visual probe to the vocal response, as a function of Dominance and Reiteration are presented in the following figure. Errors (i.e. listeners named the visual probes with a word that is totally different from the target word) were very rare (approximately 0.05 across all conditions), and therefore error data were not analyzed further in the present study.

![Figure 1. Mean response latencies as a function of Dominance and Reiteration.](image-url)

A 2 (Dominance) x 2 (Reiteration) repeated measure ANOVA was conducted on the response latencies to the visual probes. Results revealed two clear main effects on both variables: Dominance and Reiteration. However, there was no interaction effect observed.

First, results suggested that frequency of the individual meanings of the ambiguous word marginally influenced the response time, $F(1.55) = 4.13, p < .05$. Collapsed over levels of the other variable, the mean response time to access the dominant homophone meaning was 814.5 milliseconds and the time to the subordinate homophone meaning was 825 milliseconds. Results indicated that the dominant meaning of a given homophone would be activated first in a highly constrained sentential context. Second, there was a main effect on the variable of Reiteration, $F(1.55) = 32.83, p < .01$. Collapsed across the other variable, the mean response time to reiterated condition was 881 milliseconds and the time to the not reiterated condition was 758.5 milliseconds with a difference of 122.5 milliseconds. The reiteration effect obviously pointed to a strong inhibitory effect occurred to the inappropriate meanings following the first occurrence of the ambiguous word, in which an appropriate meaning was firmly selected due to the preceding constraint context. These inhibitory effects could even sustain to a long duration, more than 2 seconds, a bit longer than Simpson and Kang’s finding [14].

In addition, the null interaction effect suggested that there might be other lexical variables operated during the continuous sentence processing that the present study did not controlled. One of the possible confounds is the variable of homophone density [6,7,19]. This variable might be more or less affected the continuous competition among different alternative meanings of the ambiguous word. The higher the homophone density, the higher the competition among them and thus reducing the inhibitory effects [23,24]. Other potential confound is likely to be the relatively frequencies of the dominant and subordinate meanings of the ambiguous word [25].

### 3. General Discussion

The present study further examines the lexical ambiguity in sentence processing from a different perspective, to examine the inhibitory processes of the ambiguous words in sentence comprehension. We used spoken Chinese homophone in the present study as a rigorous test case because of the pervasive homophony phenomenon in this language. Clearly, the present results show that preceding sentential context aids the processing of Chinese homophones at an early stage. These findings were actually consistent with other relevant studies in Chinese sentence processing [5-9].

In addition, the present results again indicated a strong dominance effect during the processing of the ambiguous word. The dominance effect was strong enough to aid the early selection among different alternative meanings of the ambiguous word, as revealed in the faster response time in DN and SR conditions than in SN and DR conditions respectively. This result is obviously against the assumption of modular approach of language processing: frequency effect could only occur at a late selection stage [1,2]. In addition, the reiteration effects clearly confirmed the inhibitory processes occurred rapidly following the occurrence of the ambiguous word. More importantly, the inhibitory processes could sustain to a longer duration, over 2 seconds (around twelve syllables’ time) after the first appearance of the ambiguous word in a sentence, in line with other relevant studies [13,14].

Ongoing experiments are being designed in our laboratory to further examine how the other variables, such as homophone density effect, relative frequencies information, and Inter-stimuli-interval (ISI), influence the inhibitory processes of the ambiguous word during sentence processing.
ACKNOWLEDGEMENTS

This research was supported by a Research and Development Fund from the Open University of Hong Kong. The author would also like to thank Katherine Leung for her assistance in materials preparation to the present study. The author is now at the Department of Educational Psychology, Counselling and Learning Needs, The Hong Kong Institute of Education. Email: myip@ied.edu.hk or myip@ouhk.edu.hk.

REFERENCES


2876