In G. Cottrell (ed.), *Proceedings of the 18th Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates.

Lexical Ambiguity and Context Effects in Spoken Word Recognition: Evidence from Chinese

Ping Li and **Michael C. Yip**

Department of Psychology, University of Richmond, Richmond, VA 23173, ping@cogsci.richmond.edu

Department of Psychology, Chinese University of Hong Kong, Shatin, NT, Hong Kong, michaelyip@cuhk.hk

Abstract

Chinese is a language that is extensively ambiguous on a lexical-morphemic level. In this study, we examined the effects of prior context, frequency, and density of a homophone on spoken word recognition of Chinese homophones in a cross-modal experiment. Results indicate that prior context affects the access of the appropriate meaning from early on, and that context interacts with frequency of the individual meanings of a homophone. These results are consistent with the context-dependency hypothesis which argues that ambiguous meanings of a word may be selectively accessed at an early stage of recognition according to sentential context. However, the results do not suggest a pre-selection process in which the contextually appropriate meaning can be activated prior to the perception of the relevant acoustic signal.

Introduction

Lexical ambiguity has been the focus for the study of context effects in word recognition in the past 20 years (Onifer & Swinney, 1981; Small, Cottrell, & Tanenhaus, 1988; Simpson & Krueger, 1991; Swinney, 1979; Tabossi, 1988). Results from these studies point to two major hypotheses of lexical access. The exhaustive access hypothesis argues that all meanings of a homophone will be accessed momentarily following the occurrence of the word; semantic context can only help to select the appropriate meaning at a post-access decision/selection stage. This hypothesis assumes a modular account of lexical processing in which context does not penetrate lexical access (Fodor, 1983). In contrast, the context-dependency hypothesis argues that the contextually appropriate meaning of a homophone can be selectively accessed early on, if prior context provides strong bias to one of the meanings. It assumes an interactive process in which lexical and contextual information can mutually influence each other at a very early stage (McClelland, 1987).

Few studies have examined these hypotheses in Chinese, a language that is extensively ambiguous on a lexical-morphemic level. According to the Modern Chinese Dictionary (Institute of Linguistics, 1985), 80% of the monosyllables are ambiguous, and half of them have five or more homophones. For example, the 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. Should we assume that upon hearing “yi”, Chinese listeners activate all 90 or more meanings of the syllable? We should, if we accept the exhaustive access hypothesis in its strict sense: according to this hypothesis lexical access is an autonomous and capacity-free process. However, if we follow the context-dependency hypothesis, only the contextually appropriate meaning will be activated when listeners hear the syllable.

The present study is designed to investigate how context, frequency, and homophone density information constrain the access of lexical meanings, using Chinese homophones as a crucial test case.

Method

Subjects

Sixty native Cantonese speakers (38 females and 22 males, mean age = 19.6) who reported no speech or hearing deficits participated in the experiment. All subjects were students at the Chinese University of Hong Kong. They took part in the experiment as a laboratory requirement for credit in an introductory psychology course.

Materials

Thirty spoken homophones (all nouns, see Appendix) were selected, each with at least two different meanings in the same tone (syllables in different tones are not considered homophones in this study). Each homophone was embedded in two sentences with prior context 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 60 test sentences with the prior 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 (1978): 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 1.6. This score was above the high constraint condition in Marslen-Wilson and Welsh (1978). An effort was also made to have prior context of equal length, and the average length of the test sentences (counting the target homophone) was 14 words (ranging from 12 to 17 words).

Four 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 a second meaning not biased by the context, or (c) unrelated control.

(2) Dominance. The prior context biased either the dominant meaning (more frequent) or the subordinate meaning (less frequent) of a homophone. The frequency counts were based on Ho & Jiang (1994).

(3) Homophone Density. A given homophone had either many potential competitors (four or more) or few (two to three). No previous studies have examined this variable to our knowledge.

(4) SOA (stimulus-onset-asynchrony). The visual probe occurred at a given SOA relative to the spoken homophone, either at the onset or at the offset of the homophone.

A homophone example (*zeong*) and the two corresponding test sentences are given below.

(1) Sentence:
Ngo nam hai dungmatjyun leoi-min zeoi daai ge zau hai ***zeong***.
I think that the biggest in a zoo is the elephant.
(literally: I think in zoo inside most big 's then is elephant.)
Probes:
(a) *syu* "mouse" (biased)
(b) *gwan* "rod" (unbiased)
(c) *zat* "quality" (unrelated control)

(2) Sentence:
Popo waa keoi hanglou m fongbin soeng jiu jat-zi ***zeong***.
Grandma says that she has trouble walking and wants to have a stick.
(literally: grandma says she walk not convenient think want one stick.)
Probes:
(a) *gwan* "rod" (biased)
(b) *syu* "mouse" (unbiased)
(c) *zat* "quality" (unrelated control)

All the visual probes were based on a semantic relatedness judgment task with a separate group of 20 native Cantonese speakers. They were asked to think of three nouns that have the same or closely related meaning to each homophone, and their most frequent response was selected as the visual probe for the homophone (see Appendix for a complete list of the probes used in the experiment).

Design

Subjects were divided into two groups of 30 according to two different SOA conditions. Within each SOA condition, the 30 subjects were again randomly assigned to six groups of five. Each group randomly received an equal number of sentences for each SOA condition in the 2 (Dominance) x 2 (Homophone Density) x 3 (Probe Type) design. This yielded a total of 12 different experimental conditions. The order of presentation for the sentences was pseudorandomly arranged such that the visual probes did not consecutively bias spoken homophones. The order of presentation was counterbalanced across subjects. No subject heard the same target homophone twice.

Experimental Apparatus

The test sentences were read by a native Cantonese speaker at a normal conversation rate, and were tape-recorded and digitized into a PowerMac computer. A sampling rate of 22kHZ with a 16-bit sound format was used for digitizing. The presentation of auditory and visual stimuli was controlled by the PsyScope program (Cohen, MacWhinney, Flatt, & Provost, 1993). Subjects' naming latencies were recorded by the CMU button-box (Cohen et al., 1993). A unidirectional microphone to register listeners' vocal response was connected to the button-box through the box's voice-activated relay.

Procedure

A cross-modal naming technique (e.g., Li, in press; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982) was used. Subjects saw a fixation point, and immediately heard on a pair of headphone the sentence in which the homophone was embedded. A visual probe then occurred on the computer screen according to the SOA condition. Their task was to, as accurately and quickly as possible, name the visual probe aloud into a microphone. Subjects were given a maximum of two seconds to respond, counting from the onset of visual probe. This length of time was sufficient for most subjects to give their responses while at the same time putting them under time pressure.

All subjects did the experiment individually. Before the test began, they were given a practice session in which they heard a set of separate but similar sentences. The experiment took about twenty minutes.

Data Analysis

The dependent variable was subjects' response latencies to each visual probe. The latency was measured from the onset of the visual probe to the subject's vocal response.

Results

A 2 x 3 x 2 x 2 (SOA x Probe Type x Dominance x Homophone Density) ANOVA revealed two main effects: SOA ($F_{1,59} = 94.41$, $p < .001$), and Probe Type ($F_{2,58} = 5.13$, $p < .01$). There was also an interaction between SOA and Dominance ($F_{1,119} = 4.78$, $p < .05$). The main effect of Probe Type indicates that the contextually biased meaning affected subjects' naming of the visual probe. It shows that context effects could take place immediately following the occurrence of the homophone, or even within the acoustic boundary of the spoken word. These effects occurred much earlier in our experiment than what has been previously argued for (e.g., about 1.5 seconds following the occurrence of the ambiguous word, see Onifer & Swinney, 1981; Swinney, 1979). The main effect of SOA shows that although context did not pre-select the meaning before any acoustic information of the homophone (in the Onset SOA condition), listeners did show sensitivity to the contextually biased meaning with minimal acoustic information (in the Offset SOA condition). The interaction effect of SOA by Dominance was due to the fact that the frequency of the meanings of a homophone was reflected only in the Onset SOA condition. In order to see more clearly the various main effects and interactions in each SOA condition, we conducted two separate 3 x 2 x 2 (Probe Type x Dominance x Homophone Density) ANOVAs.

Figure 1 presents the results for response latencies as a function of Probe Type, Dominance, and Homophone Density in the Onset SOA condition. ANOVA shows that only the main effect of Dominance ($F_{1,29} = 4.91$, $p < .05$) was significant. Individual comparisons revealed that this effect was due to differences in the unbiased, low density items ($F_{1,29} = 4.89$, $p < .05$). This effect indicates that when there were few competitors within a homophone, frequency was the main factor for determining subjects' response speed in the unbiased condition. However, when context provides a bias, then frequency effect became weaker.

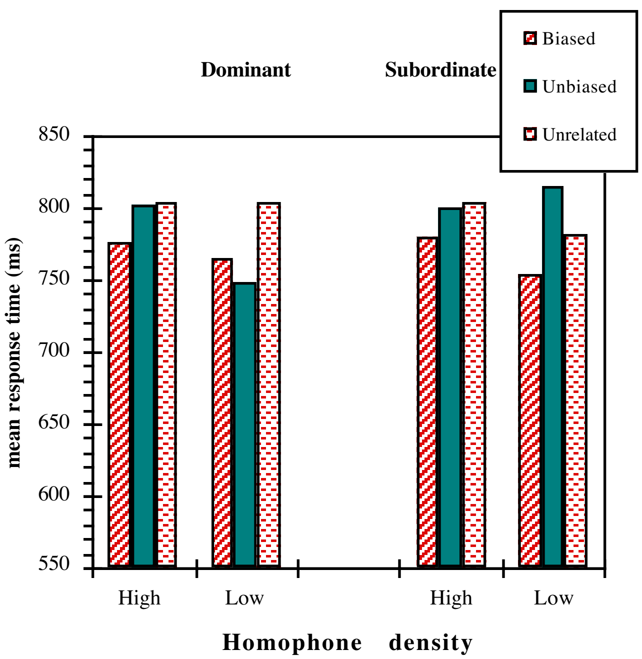

Figure 1: Mean Response Latencies (ms) as a function of Probe Type, Dominance, and Homophone Density in the Onset SOA condition

Figure 2 presents the results for response latencies as a function of Probe Type, Dominance, and Homophone Density in the Offset SOA condition. ANOVA shows that there was a main effect of Probe Type ($F_{2,28} = 8.84$, $p < .001$), indicating that context had an effect in listeners' identification of the contextually biased meaning of a homophone, which in turn facilitates the naming of the corresponding visual probe. There was also a significant interaction between Probe Type and Dominance ($F_{2,58} = 4.62$, $p < .05$), showing that during recognition context interacted with the frequency of individual meanings of a homophone. In general, dominant meanings were accessed faster than subordinate meanings ($F_{1,59} = 4.83$, $p < .05$). However, when context provided a strong bias, then they did not differ ($F_{1,59} = 2.65$, n.s.).

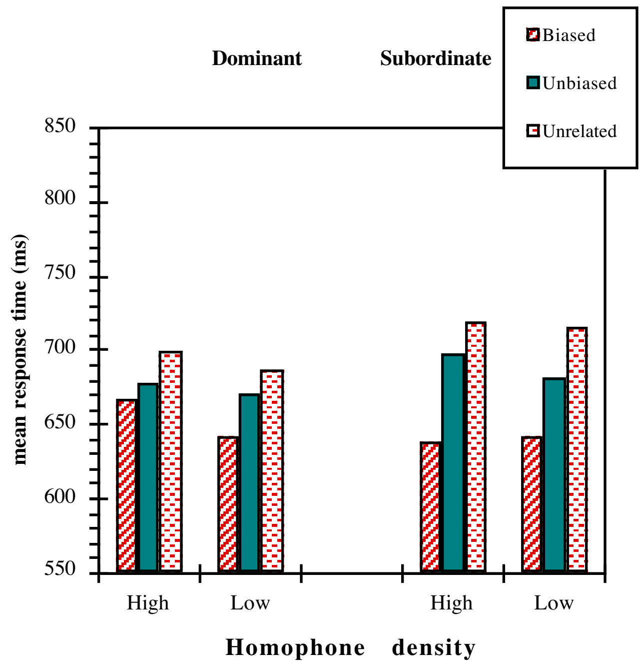

Figure 2: Mean Response Latencies (ms) as a function of Probe Type, Dominance, and Homophone Density in the Offset SOA condition

Note that in both SOA conditions, there was no clear main effect of Homophone Density, nor clear interactions between Dominance and Density. The absence of these effects in our experiment could be due to two confounding factors. First, in this study we used written frequency to approximate the frequency differences for the spoken homophones, due to the unavailability of spoken frequency information for Cantonese. Second, we suspect that the frequency of the visual probes could confound the results. We did not control for the frequency of the visual probes because the visual probes were derived from independent evidence in a semantic relatedness judgment task (see Method), as there have been no semantic associate norms for Cantonese. Our suspicion was confirmed in a separate analysis treating the frequency of the visual probes as an independent variable ($F_{1,59} = 5.97$, $p < .05$).

To summarize, the above results provide a time-course picture of the context effect in spoken word recognition of homophones. In the Onset SOA condition, there is no effect of context as a function of Probe Type, whereas in the Offset SOA condition, contextually appropriate probes are named significantly faster than contextually inappropriate probes. Dominance of the homophone meanings also plays an important role in the recognition process, and it interacts with context in affecting listeners' response speed. In other words, when context provides no clear bias to the individual meanings of a homophone, dominant meanings in general are activated faster; when context provides a clear bias, dominant and subordinate meanings elicit similar naming latencies.

General Discussion

This study attempts to provide new evidence on an old problem. Chinese represents a significantly different language from Indo-European languages, and its lexical properties make the language ideal for testing the lexical ambiguity issue. Our results indicate that given a visual probe at the offset of a spoken homophone, listeners show sensitivity to the contextually biased meaning. These results are consistent with the context-dependency hypothesis which argues that ambiguous meanings of a word may be selectively accessed at an early stage of recognition according to prior sentential context (Simpson, 1981; Simpson & Krueger, 1991; Tabossi, 1988). They point to a much earlier context effect than what has been previously assumed in modular accounts (e.g., Onifer & Swinney, 1981). Recently, Moss and Marslen-Wilson (1993) argue 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 word offset. They suggest that the initial access for semantic information about a word could occur much earlier, and that the selection of the contextually appropriate meaning need not occur after the word. Our results are consistent with these accounts, in that context effects can be clearly observed at the offset of a spoken homophone, within the temporal acoustic boundary of the word.

Although our results support the context-dependency hypothesis in general, they do not provide support for a top-down process of pre-selection of the contextually appropriate meaning without any acoustic information. In our experiment when the listener sees a visual probe at the onset of the spoken homophone, the context shows no clear effect. This result is consistent with the "bottom-up priority" principle of the cohort model (Marslen-Wilson, 1987) which argues that in order for lexical access to take place, there has to be limited bottom-up acoustic information. Our results from the Onset SOA condition seem to suggest that there is a very brief moment of lexical access of multiple meanings following the onset of the acoustic signal. However, this moment is short-lived, and other information such as frequency could start to play a role rapidly thereafter.

The current study also indicates the importance of frequency or dominance information in spoken word recognition of homophones. Although the frequency information used in this study is only an approximation because they are based on written norms, the results do suggest an interactive picture in general, in which frequency and context can mutually interact for the identification of the correct meaning. For example, frequency of the individual meanings of a homophone shows an effect in the low density items in the Onset SOA condition, counter to hypotheses which argue that frequency effects can only occur later on in the selection stage (e.g., Onifer & Swinney, 1981). Ongoing experiments with several paradigms are being designed in our laboratories to examine context, frequency, and homophone density in greater detail so that we may provide a more comprehensive account of context effects in spoken word recognition.

Acknowledgements

This research was supported by a Direct Research Grant from the Chinese University of Hong Kong. We are very grateful to Elizabeth Bates, François Grosjean, and Greg Simpson for their very helpful discussions of the materials presented here.

References

Cohen, J.D., MacWhinney, B., Flatt, M., and Provost, J. (1983). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, and Computers*, 25, 257-271.

Fodor, J.A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Ho, H.-H. & 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, P. (in press). Spoken word recognition of code-switched words by Chinese-English bilinguals. *Journal of Memory and Language*.

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

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

McClelland, J.L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Moss, H., & 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., & 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., & 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., & Krueger, M.A.(1991). Selective Access of Homograph Meanings in Sentence Context. *Journal of Memory and Language*, 30, 627-643.

Small, S., Cottrell, G., & Tanenhaus, M. (1988). *Lexical ambiguity resolution: perspectives from psycho-linguistics, 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.

Appendix: Homophones and Probes Used in the Experiment

Homophone	DOM	Probe	SUB	Probe
baan1	class	zu3	dot	zi4
baan2	board	yi3	edition	yin4
bo1	ball	qiu2	slope	cao3
bou3	cloth	mian2	newspaper	shi4
ci4	pond	he2	prose	zi4
coeng1	gun	dan4	window	men2
coeng4	wall	zhuan1	playground	cheng2
dou2	island	hai3	gamble	shu1
fo3	goods	cang1	lesson	shu1
gau2	dog	mao1	nine	shu4
jyun4	circle	fang1	ape	hou2
kei4	flag	jun1	date	ri4
kwan4	dress	yi1	group	ji2
maa5	horse	niu2	yard	chi3
min6	face	yan3	noodle	fen3
mou6	fog	yu3	tomb	si3
ping4	vase	zun1	plateau	di4
saam1	clothes	wa4	three	si4
sau2	hand	tui3	tanker	chuan2
seoi3	tax	qian2	year	nian2
si1	silk	tiao2	poem	wen2
sin3	thread	you4	fan	liang2
soeng1	container	he2a	frost	xue3
tong4	sugar	tian2	lesson	xue2
waa6	language	shu1	painting	se4
wan4	cloud	tian1	soul	gui3
wo1	nest	bei4	wok	huo4
wu4	lake	shui3	arc	wan1
zoeng3	debit	zhai4	curtain	ying2
zoeng6	elephant	shu3	stick	gun4

Note: DOM -- Dominant meaning; SUB -- Subordinate meaning. The probes are semantically related to the corresponding meanings.