INTERSPEECH 2009 BRIGHTON

Context effects and the processing of ambiguous words: Further evidence from semantic incongruence

Michael C.W. Yip

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

10.21437/Interspeech.2009-543

Abstract

A cross-modal naming experiment was conducted to further verify the effects of context and other lexical information in the processing of Chinese homophones during spoken language comprehension. In this experiment, listeners named aloud a visual probe as fast as they could, at a pre-designated point upon hearing the sentence, which ended with a spoken Chinese homophone. Results further support that context has exerted an effect on the disambiguation of various homophonic meanings at an early stage, within the acoustic boundary of the word. This contextual effect was even stronger than the tonal information. Finally, the present results are in line with the context-dependency hypothesis that selection of the appropriate meaning of an ambiguous word depends on the simultaneous interaction among sentential, tonal and other lexical information during lexical access.

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 [1-10]. 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 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, and the sentential context can only help to select the appropriate meaning at a 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 [11]. The other one is the context-dependency hypothesis, which argues that only the contextually appropriate meaning of the ambiguous word will be accessed early on if the preceding sentential context provides a strong bias to the appropriate meaning due to its increased activation. This hypothesis undoubtedly assumes that language processing is operated by an interactive approach in which information can flow both bottom-up and top-down simultaneously and that lexical access and sentential context can mutually influence each other at a very early stage [12].

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, the overall results seemed to be incongruent in patterns. The present study attempts to further address this question again in Chinese language. Chinese is a language that differs significantly from most other languages (e.g., in its use of lexical tones, its morphemic mono-syllabicity) and lexical ambiguity is pervasive in Chinese (because of its extensive homophony). Therefore, it offers many unique and interesting psycholinguistic properties in its phonological, lexical, and syntactic structures [13] to crucially investigate the issue. From the Modern Chinese Dictionary [14], 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.) [15]. 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 of the single syllable simultaneously? They should if they follow the context-dependency hypothesis, only the contextually appropriate meaning will be activated when listeners hear the syllable due to the robust context effects. Apart from the contextual effect, the use of tonal information in Chinese can also help to differentiate the alternative meanings associated with the homophone and thus reducing the potential candidates due to the lexical nature of Chinese tone. For example, the identical Chinese syllable /ma/ with different tones: /ma1/; /ma4/; /ma5/; /ma6/ literally mean “mother”; “hemp”; “horse”; “to scold” respectively. Therefore, how does tonal information interact with sentence context to disambiguate different homophone meanings? What is the time-course of the tonal information starts to play a role in differentiating alternative meanings for the homophone? The present study was designed to answer all these questions, by using a cross-modal naming experiment, so as to seek additional evidence for making a full picture on the interplay between the sentential context effects and lexical tone during spoken sentence processing.

2. Experiment

2.1. Method

Participants. One hundred and forty-four native Cantonese speakers (50 male and 94 female, mean age =22.3) 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. Sixty 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 preceded by a sentence context that was biased toward one of the meanings of the homophone. The homophone either matched the sentence context or did not match due to an incorrect tone, creating a total of 120 sentence materials. Following the procedure of relevant studies [10], a separate group of 20 native Cantonese speakers was asked to judge the degree of constraint of the preceding sentential context on the matched 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 [16]: 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 [16]. An effort was also made to have the sentence context of equal length, and the average length of the test sentences, including the homophones, was 14 words (ranging from 12 to 17 words). In addition, we have made an effort to eliminate any kind of intra-sentential priming from other individual words within the entire 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 [10]. 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 matched with the same category of initial phonemes and individual frequency information.

Altogether, there were four variables manipulated in this experiment:

(1) Context: The preceding context was (a) matched to the context (homophone carried a correct tone that fit the semantics of the sentence context) or (b) mismatched to the context (homophone carried an incorrect tone that did not fit the semantics of the sentence context).

(2) Relatedness: The visual probes were either semantically related to the spoken homophone or unrelated.

(3) Homophone Density: A given homophone had either many potential semantic competitors (four or more alternative meanings) or few semantic competitors (below three alternative meanings). The density information was based on study of Zhang and Zhang [15].

(4) SOA: The visual probe occurred at three different SOAs (stimulus-onset-asynchrony) relative to the spoken homophone: (a) the vowel onset point (VOP), or (b) the acoustic offset of the homophone (OS), or (c) 300ms after the acoustic offset of the homophone (300ms). The vowel onset point (VOP) for each homophone was derived from other previous gating results [8]. The VOP for each spoken homophone was different (average VOP was 58% of the entire duration of the word, around 226msec.).

An example of the high density homophone *ji6* (word) and the two types of test sentences are given below.

(a) Match to context:
Nei gau ging hai zoeng zi soeng min se zo di me je ji6.
Actually, which word are you writing on the paper?

(b) Mismatch to the context:
Gam do zung ngaan sik lei min ngo zeoi sang ke zau hai ji6.
(*The correct homophone should be ji2 "purple".)
Among all the different colors, I hate word the most.

The four visual probes in these sentences are: *bat1* "pen" (related-match), *sue1* "book" (related-mismatch), *booi1* "glass" (unrelated-match), and *sin1* "fresh" (unrelated-mismatch).

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 by the SoundEdit software. A computer program called PsyScope [17] controlled the presentation of the materials. A microphone, which was used to register listeners' vocal responses and counted 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 participants' accuracy.

Procedure. Participants were randomly assigned into three groups of 48. Each group randomly received an equal number of sentences in the 2 (Context) x 2 (Relatedness) x 2 (Homophone Density) x 3 (SOA) factorial design. Each participant received 60 sentences in total. 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 also counterbalanced across all participants. No participant heard the same homophone twice.

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 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 context, relatedness and SOA are presented in the following table*. Errors (i.e. listeners named the visual probes with a word that is totally different from the target word) were very rare (approximately 0.03 across all conditions), and therefore no analysis was conducted on the error data in the present study.

Context	**Match**		**Mismatch**	
	Related	Unrelated	Related	Unrelated
VOP	795.8	826.1	789.7	833.6
OS	754.8	807.2	788.6	813.3
300 msec	738.1	813.1	736.7	820.9

*Note: Because the effect of homophone density was at all absent in the experiment, data for this variable were not computed in the above table.

A 3 (SOA) x 2 (Context) x 2 (Relatedness) repeated measure ANOVA was conducted on the response latencies to the visual probes. There were main effects on all of the three variables and there were also significant interaction effects of Context and SOA as well as among the three variables. In the followings, we will divide into three parts to analyze the pattern of data, based on the levels of SOA, for easier interpretation.

First, in the condition of VOP, results found that context did not greatly influence the response time, $F(1,143) = 1.8$, $p > .05$. It was because the different meanings of the homophone seemed to be accessed before the vowel onset point. This result might propose that there was a very short period of time (probably before VOP) that more than one meaning of the homophone would be accessed after the occurrence of the ambiguous word. This speculation was confirmed by the interaction of SOA by Context, $F(2,286) = 15.22$, $p < .05$. However, there is a significant effect of relatedness, $F(1,143) = 16.53$, $p < .05$. This result was a kind of semantic priming effect that listeners usually responded to related words faster than the unrelated control words [18].

Second, in the condition of OS, results found that context started to play a role in selecting the correct meaning of the homophone due to significant different responses times between match and mismatch context conditions, $F(1,143) = 13.46$, $p < .05$. Together with the significant relatedness effect $F(1,143) = 21.87$, $p < .05$, as well as the interaction between Context and relatedness, $F(2,286) = 38.13$, $p < .05$ all these results indicated that the processing time to access the correct homophone meaning would be shortly after the VOP of the homophone and possibly within the acoustic boundary of the word [5-10] when the context matches the meaning of the homophone.

Thirdly, in the condition of 300msec, results indicated clear relatedness $F(1,143) = 31.56$, $p < .05$ but a nil context effect. The patterns of results were reasonable and compatible with previous studies. In this stage, listeners would base more on the acoustic information of the homophone to make up a correct response.

In addition, there was the overall main effect on the variable of SOA, $F(2,286) = 52.7$, $p < .001$. Collapsed over other variables, the mean response time to each SOA condition was 811msec (VOP), 791msec (OS) and 777msec (300msec); and the fastest response latencies occurred at the 300msec SOA. The SOA effect implied that the access to a correct homophone meaning appeared at a relatively late period of time (after the VOP to 300msec following the occurrence of the ambiguity) even the preceding sentence offered contextual influence. This result is consistent with Simpson and Krueger's finding [19].

One more thing is that the present patterns of results pointed out that the context effects were even stronger than the tonal influence of the homophone itself, a kind of semantic incongruence. That is, even the listeners hear a pseudo-homophone (correct segmental information with incorrect supra-segmental information of the same syllable) that do not match the current context, they still access the contextually appropriate homophone meaning first and then the acoustically appropriate pseudo-homophone meaning emerge thereafter when the tonal information of the homophone interact with the current semantics of the sentence context. This process is most likely taking place between the time of VOP and OS, which is consistent with other previous gating results [6].

3. General Discussion

The present study further examines the lexical ambiguity in sentence processing from a cross-language perspective, we used spoken Chinese homophone as a rigorous test case because of the pervasive homophony phenomenon in Chinese language. Clearly, the semantic incongruence results confirm that sentence context really aids the processing of Chinese homophones at an early stage, shortly after the acoustic onset of the word (or vowel onset of the word) and most likely within the acoustic boundary of the spoken word. These findings further confirm that Chinese speakers could identify the appropriate meaning at an early point of time that was consistent with other relevant studies in Chinese sentence processing [5-10,13].

In addition, the present results indicated an interaction effect between context and tonal information. The interaction firstly demonstrated that contextual information mutually interacted with the tonal (or other lexical information) of the Chinese homophones together during sentence processing. The role of tonal information clearly shows up late, compared with the strong context effects, during the temporal processing of the homophone because native Cantonese listeners usually realized the tone of Chinese words at the onset of their vowel sound. These results are compatible with the gating results reported in our previous studies [6].

One point should be noted that the effect of homophone density was absent in the present study once again. Possible explanations for this result may be: (1) due to the fact that the semantic judgment norms gathered in this study are ineffective in reflecting the nature of this variable; or (2) the variable of homophone density may not be simply measuring the actual number of semantic competitors based on the same phonological representation since listeners may actually not have knowledge to some potential words in their mental lexicon [20]. Ongoing experiments with different methodologies (such as eye-tracking measure) or different experimental designs [21] are being designed in our laboratory to further examine the nature of homophone density effect as well as the contextual influence in Chinese spoken word recognition and sentence processing.

4. Acknowledgements

This research was supported by the Research and Development Fund from the Open University of Hong Kong and the Research Support Fund from the Hong Kong Institute of Education. I would like to thank Ping Li and three anonymous reviewers for their constructive comments to the study and to Katherine Leung for her assistance in materials preparation to the experiment. The author is now at the Department of Educational Psychology, Counselling and Learning Needs, The Hong Kong Institute of Education. Email: mcwyip@ied.edu.hk

5. References

1. 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.

2. Simpson, G. B. (1981) Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20,* 120-136.
3. Swinney, D. A. (1979) Lexical Access during Sentence Comprehension: (Re)Consideration of Context Effects. *Journal of Verbal Learning and Verbal Behavior, 18,* 645-659.
4. Tabossi, P. (1988) Accessing Lexical Ambiguity in Different Types of Sentential Contexts. *Journal of Memory and Language, 27,* 324-340.
5. Li, P., & Yip, M. C. W. (1996) Lexical ambiguity and context effects in spoken word recognition: Evidence from Chinese. In G. Cottrell (ed.), *Proceedings of the 18th Annual Meeting of the Cognitive Science Society,* 228-232.
6. Li, P., & Yip, M. C. W. (1998) Context Effects and the Processing of Spoken Homophones. *Reading and Writing, 10,* 223-243.
7. Li, P., Shu, H., Yip, M. C. W., Zhang, Y., & Tang, Y. (2002). Lexical ambiguity in sentence processing: Evidence from Chinese. In M. Nakayama (ed.), *Sentence Processing in East Asian Languages*. Stanford, CA: CSLI Publications. pp. 111-129.
8. Yip, M. C. W. (2000). Spoken word recognition of Chinese homophones: The role of context and tone neighbors. *Psychologia, 43*, 135-143.
9. Yip, M. C. W. (2002). Access to homophonic meanings during spoken language comprehension: Effects of context and neighborhood density. *Proceedings of ICSLP'2002, 1665-1668*.
10. Yip, M. C. W. (2008). Lexical and sentential priming of ambiguous words. *Psychologia, 51*,196-205.
11. Fodor, J. A. (1983) *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
12. 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: Erlbaum.
13. Zhang, Y., Wu, N., & Yip, M. C. W. (2006). Lexical ambiguity resolution in Chinese sentence processing. In P. Li, L. H. Tan, E. Bates, & O. J. L. Tzeng. (eds.), *Handbook of East Asian Psycholinguistics (Vol. 1: Chinese)*, pp. 268-278. Cambridge, UK: Cambridge University Press.
14. Institute of Linguistics, The Academy of Social Sciences. (1985) *Xianda Hanyu Cidian* (Modern Chinese Dictionary). Beijing: Commercial Press.
15. Zhang, L.-Y. & Zhang, S.-Y. (1987). Guoyin Yueyin Suoyin Zihui (A Pronunciation Dictionary of Cantonese and Putonghua). Hong Kong: Chung Hwa Bookstore.
16. Marslen-Wilson, W. D., & Welsh, A. (1978) Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29-63.
17. Cohen, J. D., 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.
18. Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds). *Basic processes in reading: Visual word recognition.* Hillsdale, NJ. Lawrence Erlbaum Associates. pp. 264-336.
19. Simpson, G. B., & Krueger, M. A. (1991) Selective Access of Homograph Meanings in Sentence Context. *Journal of Memory and Language, 30,* 627-643.
20. Ahrens, K. (2002). Timing issues in lexical ambiguity resolution. In M. Nakayama (ed.) *Sentence Processing in East Asian Languages*. Stanford, CA: CSLI Publications. pp. 1-29.
21. Yip, M. C. W. (2008). Inhibitory Processes of Chinese Spoken word recognition. *Proceedings of Interspeech'2008,* 2873-2876.