THE ROLE OF PROSODY IN PRIMING ALTERNATIVES IN MANDARIN CHINESE

Mengzhu Yan, Sasha Calhoun, Paul Warren

Victoria University of Wellington, New Zealand
mengzhu.yan@vuw.ac.nz, sasha.calhoun@vuw.ac.nz, paul.warren@vuw.ac.nz

ABSTRACT

A growing body of psycholinguistic research carried out in Germanic languages shows that listeners infer contextual alternatives to an element in a sentence when it has a contrastive accent, even when the alternatives are not explicitly mentioned in the discourse. This paper reports a cross-modal priming experiment conducted to test the role of contrastive prominence in priming alternatives of subject nouns in Mandarin Chinese. The results showed that alternatives were recognized faster when the subject carried contrastive prominence than when it did not, suggesting a common mechanism of focus generating alternatives across languages when prosodic prominence is used to signal focus in the language.

Keywords: alternatives, contrast, focus, Mandarin Chinese, prosody

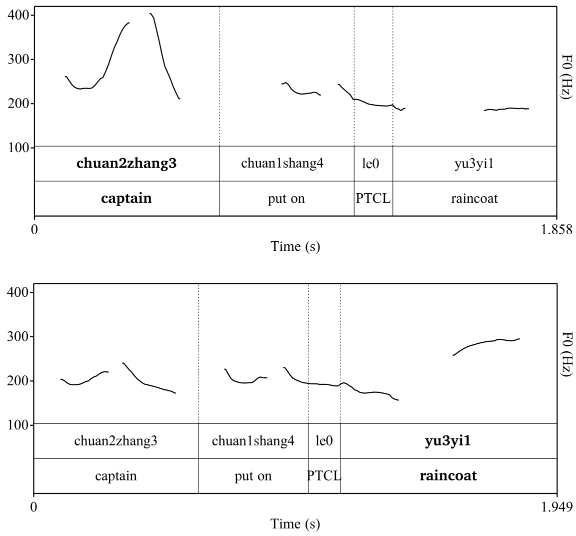
1. INTRODUCTION

It is usual in discourse that much of the information conveyed is implied, not explicitly stated. For example, in *The CAPTAIN put on the raincoat* the contrastive accent on *captain* not only tells us that the captain put on a raincoat, but also implies someone else (e.g., the *sailor*), could have put on a raincoat. Therefore, understanding implicature is crucial to successful communication [11, 15]. Further, studies show that listener do infer alternatives on hearing sentences like this [4, 13]. It is largely assumed that the mechanism behind this is (contrastive) focus marking: following the alternative semantics theory [20], contrastive pitch accents mark focus, which indicates contextually-relevant alternatives. Contrastive accenting is a common way of realising prominence in languages like English, i.e. a large movement in fundamental frequency (F0) associated with the stressed syllable of the prominent word, as well as longer duration and higher intensity [5, 16]. However, to our knowledge, almost all of the psycholinguistic studies in this area are on a handful of closely related languages, i.e. English, Dutch and German. It is not known if other types of prosodic focus marking have similar processing effects.

Much psycholinguistic literature has shown that focus facilitates language processing, such that focused words are attended to more and remembered better (e.g., [2, 8, 9]). More recently, we have seen psycholinguistic evidence of focus as indicating alternatives. One line of study used eye-tracking to investigate the role of contrastive accents in finding intended referents [10, 14, 23, 24]. For example, after first hearing *Click on the blue ball*, listeners had more fixations on a green ball when they heard *Now click on the GREEN ball* than when they heard *Now click on the green BALL*. These results show that listeners use contrastive accenting to rapidly identify referents that are available in their visual display. Later research has shown contrastive accenting also activates unmentioned and/or visually unavailable alternatives. To investigate this, cross-modal lexical decision priming experiments were used. Lexical decision tasks have been used since the 1970s [21] and have shown that single words prime their semantic associates, but priming is inconsistent in sentence contexts. [18] followed by [4] and others then looked closely at what sentence contexts prime semantically related words and which do not. The recent studies have consistently shown that unmentioned contextual alternatives (e.g., *pelican*) are primed when the prime word is contrastively accented (e.g., *In Florida he photographed a FLAMINGO*) [4, 12, 13]. [4] found that noncontrastive associates (e.g., *pink*) were weakly primed regardless of intonation. The explanation in [13] is somewhat different, i.e. all semantically related words are primed, but focus marking rapidly deactivates noncontrastive associates in later processing. Therefore, it remains unclear whether and how focus marking affects contrastive and noncontrastive associates differently.

However, to our knowledge, the priming effects of focus have only been looked at in Germanic languages. It is therefore interesting to see whether the effects can also be found in other languages that are typologically distant from Germanic languages, in this case Mandarin Chinese. Focus in Chinese is also marked through prominence, i.e., expanded F0 [7, 22, 25], longer duration [7, 25], and higher mean intensity [6]. However, prominence is not expressed as pitch accenting, as the lexical tone of a syllable determines its local F0 curve. Rather, prominence modulates the global F0 contour, which in turn influences the local F0 contour, but does not neutralize the tonal identities [25], as shown in Fig. 1. Therefore, we expected comparable priming effects of prosodically-marked focus in Chinese; though this had only previously been shown for contrastive accenting.

Figure 1: The effects of phonological prominence on F0 (bold shows the prominent word)



In this paper, we report on a cross-modal lexical priming study, carried out in Chinese, which investigates prosodic focus-marking as a means of indicating alternatives. In order to test this, our study involves two prominence conditions in canonical order, i.e. prosodic prominence on the subject (e.g., *The CAPTAIN put on a raincoat*) and prosodic prominence on the object (e.g., *The captain put on the RAINCOAT*) in Chinese, see Table 1). Identical visual targets (e.g., *captain*) were included to examine the priming of focused words. Contrastive and noncontrastive associates (both are related to the prime word but only the former can replace it in the context, e.g. *sailor* vs. *deck*) were also included to investigate the relative priming of the two related types under the expectation that contrastive associates would be more primed due to the focus mechanism. Finally, unrelated items (e.g., *pumpkin*) served as baseline controls.

2. THE EXPERIMENT

2.1. Participants

A total of 80 (50 females and 30 males) near-monolingual native Chinese speakers (mean age = 21.91, *SD* = 2.11, age range = 18-27) took part in the experiment in China. The participants received supermarket vouchers in recognition of their participation. None of them reported any hearing or reading difficulties.

2.2. Stimuli

80 experimental sentences (40 sentences * 2 prominence conditions) were constructed. For each sentence, a quadruplet of words was constructed consisting of the subject noun, contrastive associate, noncontrastive associate and control, as shown in Table 1. The same four targets appeared with both prominence conditions. Each sentence was paired with each word in that quadruplet, resulting in a total of 320 experimental stimuli. Eight lists of 40 experimental stimuli were distributed in a Latin square design, making five items per participant per condition. The lists were rotated across participants so that each participant saw only one list. An additional 120 filler sentences were included with word and nonword visual targets. Six practice sentences which had three word and three nonword visual targets were also prepared. Further, eight comprehension questions asking about the content of a previous filler were included in random positions to encourage participants to attend to the sentences.

Table 1: Prominence conditions and target types used in the experiment (underline shows contrastive prominence).

Prominence condition	PromS prosodic cue on subjects	PromO prosodic cue on objects
English	The captain put on the raincoat.	The captain put on the raincoat
Chinese	船长穿上了雨衣	船长穿上了雨衣
Target types	Identical: 船长captain Contrastive associate: 水手sailor Noncontrastive associate: 甲板deck Unrelated control: 南瓜pumpkin	

The semantic association strength of 75 word quadruplets were rated via an online questionnaire by 67 native Chinese participants. Participants were asked to rate the relationship from 1 '*not related at all*' to 7 '*highly related*' in the presence of the context sentence (e.g., *how related are 'captain' and 'sailor' in the sentence 'the captain put on the raincoat'*). Following the survey, 40 quadruplets were chosen in order to have similar relatedness scores between the identity prime and both types of associates, and also so that the identity prime and the unrelated control were as unrelated as possible. The mean relatedness score of chosen items was 4.97 for identity prime-contrastive pairs (e.g., *captain-sailor*), 5.06 for identity prime-noncontrastive pairs (e.g., *captain-deck*) and 1.81 for identity prime-unrelated pairs (e.g., *captain-pumpkin*). The last pairing was significantly different from the other two, which did not differ from one another.

The sentences were recorded in Praat [3] by a trained female native speaker (first author). All sentences were checked by two other native speakers for the location of prominence. Acoustic measures (mean F0, duration and mean intensity) for the experimental stimuli were fitted into a linear mixed effects model using lme4 [1] in R [19], showing that focused words were more prominent than unfocused words for the same words in different prominence conditions (e.g., *captain* in *The CAPTAIN put on the raincoat* vs. *The captain put on the RAINCOAT*) and also for different words in the same prominence condition (e.g., *captain* vs. *raincoat* in *The CAPTAIN put on the raincoat*) (all p values $<$ 0.001). The fitted values are given in Table 2.

Table 2: Fitted values of duration (ms), F0 (Hz) and intensity (dB) in critical prime stimuli

Word position	Prominence condition	Duration	F0	Intensity
Subject	PromS	697	333	80
Subject	PromO	542	213	71
Object	PromS	565	187	66
Object	PromO	729	285	76

The mean log frequency of the chosen items was 3.07 for identity primes, 2.88 for contrastive associates, 2.78 for noncontrastive associates and 2.88 for controls. Log frequency of words between items were not significantly different from each other (all p values $>$ 0.1), except for the noncontrastive - prime pair (t = -3.00, p = 0.016).

Variable duration of silence was added to the end of each sound file, so that the duration between the offset of the subject noun and the end of the sentence was 1500 ms, creating a stimulus onset asynchrony (SOA) of 1500 ms. 1500 ms allows all sentences to finish and also leaves a reasonable time between the end of the sentence and the onset of visual target.

The experiment was administered using Opensesame v.3.1 [17], and was run in a quiet computer room. The entire session was conducted in Chinese. Participants received written instructions on the computer screen, and the instructions were repeated orally by the experimenter. Six practice trials in a fixed order were played before the main experiment. Participants first heard a sentence, and while the sentence was being played, they focused on a fixation dot in the middle of the screen. Then they saw two characters, and had to decide whether these two characters make up a real word or not by pressing ‘m’ key (labelled as ‘是’ (yes)) for ‘yes’ response and ‘z’ key (labelled as ‘否’ (no)) for ‘no’ response as fast as they could (the ’yes’ response key was always pressed by the dominant hand). In the practice phase, participants received feedback on their responses and response times (RTs). No feedback was provided in the main experiment.

The experimental items were in four blocks with a break of at least 10 s in between. Stimuli within each block, and block order, were randomized. Eight comprehension questions were inserted after fillers randomly and evenly across blocks. The experiment lasted around 12 minutes. Demographic information was collected through a written questionnaire at the end of the experiment.

2.3. Results

Data from three participants were discarded due to low response accuracy ($<$ 80%). A further 57 incorrect critical trials (1.85%) from the remaining 3080 tokens were excluded for the RT analysis but not for the accuracy analysis. RTs of the remaining data were reverse transformed.

The overall accuracy on experimental trials is 98.15%. Accuracy scores of the experimental trials were fitted into a generalized linear mixed effects model (family: binomial) in lme4 [1]. The model included simple effects of prominence condition (PromS, PromO) and target type (identical, contrastive, noncontrastive, control) as well as their interaction as fixed effects, and by participant intercepts in the random structure. None of the fixed effects were significant, likely due to ceiling effects.

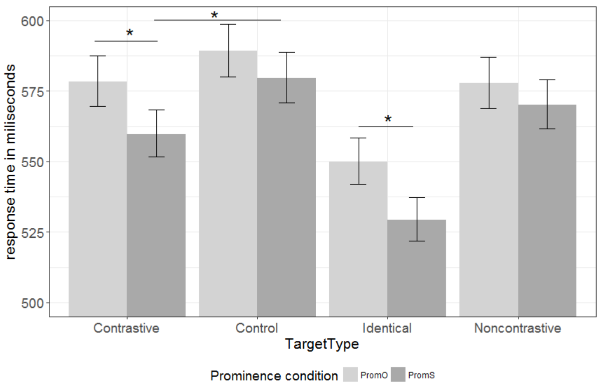
For the RT analysis, transformed RT was the dependent variable in a linear mixed effects regression. The fixed effects included the two key experimental factors (prominence condition and target type) and their interaction, the log frequency of target words, the position of a trial in the sequence of trials across the experiment, and the RT of the previous trial. The random structure included intercepts for participants and target words.

Trial position was significant, with participants getting faster over the experiment (F = 98.8, $p <$ 0.001). As expected, higher frequency words were recognized faster than lower frequency (F = 27.6, p

< 0.001). People responded more slowly when the previous trial RT was slow ($F = 114.9$, $p < 0.001$).

The back-transformed fitted RTs per condition are shown in Fig. 2. The model showed significant effects of target types and prominence conditions (target types: $F = 17.9$, $p < 0.001$; prominence conditions: $F = 20.3$, $p < 0.001$), but not a significant interaction ($F = 1.89$, $p = 0.129$). To examine which groups differed, planned comparisons involving target types and prominence conditions were conducted. The results showed that identical items and contrastive items were recognized faster when the prime word carried contrastive prominence (identical: $z = 3.83$, $p = 0.001$; contrastive: $z = 3.00$, $p = 0.025$). However, noncontrastive and control items were responded to equally fast in both prominence conditions (noncontrastive: $z = 0.91$, $p = 0.961$; control: $z = 1.30$, $p = 0.800$). Within prominence conditions, contrastive items were facilitated over control items in the PromS condition ($z = 2.71$, $p = 0.059$), but not in the PromO condition ($z = 1.25$, $p = 0.828$).

Figure 2: Back-transformed fitted RTs in ms to four target types in PromS and PromO conditions. Error bars show standard error of the means. Stars indicate significant differences.



3. DISCUSSION AND CONCLUSION

We conducted a cross-modal priming lexical decision experiment with two prominence conditions (PromS, PromO) and four target types (identical, contrastive, noncontrastive, control) related to subject nouns with an SOA of 1500 ms. As predicted, focused words (identical targets) are recognized faster than other types of words in both prosodically prominent and not prominent conditions, but priming was greater in the prominent condition. Alternatives are only primed when the prime word carries contrastive prominence, while noncontrastive associates and unrelated words are not primed in either of the conditions.

The results provide the first evidence of the effect of prosodic focus marking on the recognition of focused words and unmentioned alternatives in Mandarin Chinese. This is in line with the previous research findings for Germanic languages, and suggests that prosodic prominence as manifested in global F0 range is also effective, along with pitch accenting, in priming implicit contextual alternatives in languages that use prosodic prominence as one of the main markers of focus. Noncontrastive associates are not plausible replacements for the focused words in sentences, so listeners do not consider them as part of the alternative set, re [20]. Therefore, this study provides cross-linguistic psychological evidence for [20]'s theory, showing that contrastive associates differ from noncontrastive associates in the interpretation of focus.

[13] showed that semantic priming is different when the target is presented immediately after the prime word (SOA 0 ms) and in later processing (SOA 750 ms). Our results, with an SOA of 1500 ms are consistent with [13]'s account of later processing in the prosodically prominent condition: contrastive associates continued to be primed but noncontrastive associates were not due to the selection mechanisms that distinguish the two types of associates. Therefore, our study further supports these selection mechanisms with evidence after an even longer processing time.

As shown in Fig. 2, there was a general trend for PromS to be faster than PromO for all target types, though the difference (8.6 ms) was not significant for the control. The general difference may be because when the final word is prominent, people might be inferring alternatives to objects when seeing the visual target, so they are slower in responding to the visual words. This inhibitory effect was also found by [13]. However, overall, contrastive associates were more activated (18.6 ms faster) in the PromS condition due to the extra priming effects for alternatives.

Successful communication requires listeners to draw inferences about a speaker's implicit meaning. This research contributes to better understanding of an important part of these inferencing processes across diverse languages, looking at the linguistic mechanisms listeners are using to generate alternatives.

4. ACKNOWLEDGMENTS

This work was supported by the Marsden Fund Council (grant number 15-VUW-015) from New Zealand Government funding, awarded to the second author, and a Faculty of Humanities and Social Sciences Research Grant (grant number 218221), Victoria University of Wellington, awarded to the first author. Thank you to Minxia Wang at Henan Polytechnic University for hosting the first author, and help in recruiting participants. Finally thank you to all participants for taking part.

5. REFERENCES

[1] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* (67(1)), 1–48.

[2] Birch, S. L., Garnsey, S. M. 1995. The effect of focus on memory for words in sentences. *Journal of Memory and Language* 34(2), 232–267.

[3] Boersma, P., Weenink, D. 2018. Praat: Doing phonetics by computer [computer program]. version 6.0.37. http://www.praat.org.

[4] Braun, B., Tagliapietra, L. 2010. The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes* 25(7-9), 1024–1043.

[5] Breen, M., Fedorenko, E., Wagner, M., Gibson, E. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes* 25(7-9), 1044–1098.

[6] Chen, S.-w., Wang, B., Xu, Y. 2009. Closely related languages, different ways of realizing focus. In: *Proceedings of Interspeech 2009*. Brighton, UK: 1007–1010.

[7] Chen, Y., Gussenhoven, C. 2008. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36(4), 724–746.

[8] Cutler, A. 1976. Phoneme-monitoring reaction time as a function of preceding intonation contour. *Attention, Perception, & Psychophysics* 20(1), 55–60.

[9] Cutler, A., Fodor, J. A. 1979. Semantic focus and sentence comprehension. *Cognition* 7(1), 49–59.

[10] Dahan, D., Tanenhaus, M. K., Chambers, C. G. 2002. Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language* 47(2), 292–314.

[11] Féry, C., Krifka, M. 2008. Information structure: Notional distinctions, ways of expression. In: Sterkenburg, P. v., (ed), *Unity and diversity of languages*. Amsterdam: John Benjamins 123–136.

[12] Gotzner, N., Wartenburger, I., Spalek, K. 2016. The impact of focus particles on the recognition and rejection of contrastive alternatives. *Language and Cognition* 8(1), 59–95.

[13] Husband, E. M., Ferreira, F. 2016. The role of selection in the comprehension of focus alternatives. *Language, Cognition and Neuroscience* 31(2), 217–235.

[14] Ito, K., Speer, S. R. 2008. Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language* 58(2), 541–573.

[15] Krifka, M. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55, 243–276.

[16] Kügler, F., Calhoun, S. to appear. Typology of focus expression: prosodic encoding of information structure. In: Gussenhoven, C., Chen, A., (eds), *The Oxford Handbook of Language Prosody*. Oxford, UK: Oxford University Press.

[17] Mathôt, S., Schreij, D., Theeuwes, J. 2012. Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* 44(2), 314–324.

[18] Norris, D., Cutler, A., McQueen, J. M., Butterfield, S. 2006. Phonological and conceptual activation in speech comprehension. *Cognitive Psychology* 53(2), 146–193.

[19] R Core Team, 2018. R: A language and environment for statistical computing.

[20] Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1), 75–116.

[21] Swinney, D. A., Onifer, W., Prather, P., Hirshkowitz, M. 1979. Semantic facilitation across sensory modalities in the processing of individual words and sentences. *Memory & Cognition* 7(3), 159–165.

[22] Wang, B., Xu, Y. 2006. Prosodic encoding of topic and focus in Mandarin. In: Hoffmann, R., Mixdorff, H., (eds), *Proceedings of Speech Prosody 2006*. Dresden: TUD press 313–316.

[23] Watson, D. G., Tanenhaus, M. K., Gunlogson, C. A. 2008. Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science* 32(7), 1232–1244.

[24] Weber, A., Braun, B., Crocker, M. W. 2006. Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech* 49(3), 367–392.

[25] Xu, Y. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27(1), 55–105.