Perceptuomotor effect of lexical tones in Mandarin

Zifeng Liu[1,2], Ioana Chitoran[1,2], Giuseppina Turco[1,3]

[1]Université Paris Cité (France), [2]Clillac-ARP (France), [3]CNRS Laboratoire de Linguistique Formelle (France)

Introduction: Using the Cue-Distractor paradigm, a speech production-perception (PP) link has been explored at the phonemic as well as sub-phonemic level (e.g., VOT or vowel features) [1,2,3,4]. In this paradigm, participants' reaction times (RTs) to utter a visually prompted syllable can be modulated by the distance between the intended target and the auditory distractor, dubbed as the perceptuomotor (PM) effect. The more similar the target and the distractor are, the faster they can utter the target and vice versa. Findings from this line of research reveal that speech planning occurs dynamically at a fine-grain level beyond phonemes. However, this link has only been studied on consonants and vowels, few studies address the PP link on lexical tones [5,6,7]. This preliminary study aims to investigate if a PM effect can be provoked in tones independent of segments. The answer to this question will provide more insights into the phonological role of lexical tones in speech planning.

Methodology: To address this question, Tone 2 (rising tone) was used as the target. Two other tones were used as tone distractors, labeled as *distant tone distractor* (Tone 4, falling tone) and *close tone distractor* (Tone 1, high level tone) according to their relative acoustic distance to the target (Table.1). The target tone was carried by the syllable [p^{h}u] for its high lexical frequency of every syllable-tone combination. To make sure that the result would not be an artifact of the selected vowel, two additional vowels were introduced as *distant vowel distractor* and *close vowel distractor* based on their feature distance to the target vowel (Table 2). All the distractors were recorded by a 24-year-old native female speaker of Mandarin. Their duration and intensity were normalized. A synthesized nonspeech beep noise was included to verify the general effect of the distractors in the task compared to trials without distractors (Table 2). Each target-distractor pair was repeated 15 times with 3 SOAs (150, 200, 300 ms), yielding 315 trials (7 conditions * 3 SOAs * 15 repetitions). Additionally, 315 fillers with Tone 4 as a visual cue were included to prevent participants from predicting the target. The target was presented in *pinyin* with the corresponding tone number (i.e., pu2). 6 participants were instructed to read out the word written in *pinyin* on the screen as quickly and accurately as possible, but not too fast to make too many mistakes, ignoring any sound they heard from headphones (Fig.1). Responses that occurred before the display of distractors or 750 ms after the display of distractors were rejected. RTs were measured from the display of the visual cue to the onset of the response.

Predictions: *1)* If tones are planned in a similar way to segments and independently of segments, we expect to find a PM effect. Namely, RTs to utter the target tone will lengthen as the target-distractor distance increases in both tone distractor and vowel distractor conditions. *2)*If tones are not planned at all, or not planned separately from segments, RTs to utter the target should be comparable among different tone distractor conditions but different from the vowel distractor conditions due to the tone-vowel interaction (Table 2).

Results: Compared with no distractor condition, participants responded slower across all distractor conditions. Among tone distractor conditions, RTs to the target Tone 2 lengthened progressively as the tone distance varied from congruent to distant. RTs were the longest among all conditions when participants heard a distant tone distractor. As for the vowel distractor conditions, RTs were longer when a close vowel distractor was introduced than when a distant vowel distractor was introduced. We did not report further results from confirmatory statistical analysis due to the limited number of participants.

Conclusions: These preliminary results reveal a potential PP link for lexical tones. The RTs to produce a tone might be modulated by the distance between the target tone and the distractor tone. The closer the target and the distractor are, the faster the target is produced. The modulation of RTs also implies that tones are segment-like units and they are actively engaged in the process of speech planning. Meanwhile, we cannot confirm any interaction between tone and vowel. We are recruiting more participants to provide more solid evidence for the current findings and expect to elucidate a more distinct pattern regarding the PP link on tones.

Table 1. Relative tone distance between the target and distractors. F0 values were extracted in a log2 scale. Tone distance is defined as the averaged Euclidean distance of the f0 difference and velocity difference between the target f0 trajectory and the distractor f0 trajectory at every 5ms timestep.

	Tone 1(High Tone, H)	Tone 3 (Low tone, L)	Tone 4 (Falling tone, F)
Tone 2 (Rising tone, R)	0.513	0.784	1.122

Table 2. Predictions of RTs based on different distractor conditions. Congruent distractor is identical to the target. Vowel distractors differ only from vowels of target syllables, keeping the tone consistent between vowel distractors and target syllables. The digit in parentheses indicates the feature distance. A non-linguistic distractor, beep noise, is also included and the baseline is that without distractor.

Target [p^{h}u R]

Conditions	No distractor	Beep	Congruent	Tone distractors		Vowel distractors	
				Close	Distant	Close (1)	Distant (4)
	Ø	Beep	[p^{h}u R]	[p^{h}u H]	[p^{h}u F]	[p^{h}o R]	[p^{h}a R]
RTs (Pred.1)	Fastest	2nd	3rd	4th	5th	3rd	3rd
RTs (Pred.2)	Fastest	2nd	3rd	3rd	3rd	4th	5th

Fig. 1. An example of the RT measurement when a participant is asked to say the syllable "p^{h}u" with Tone2 (Rising tone). The distractor is introduced **AFTER** (SOA:150-300ms) the display of the target but **BEFORE** the response of participants.

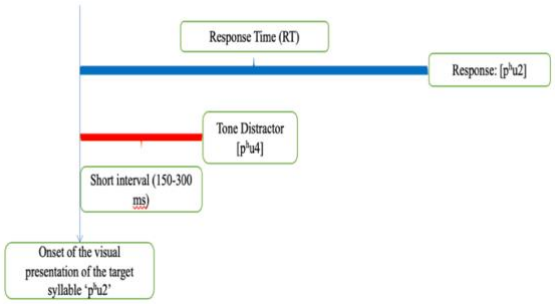
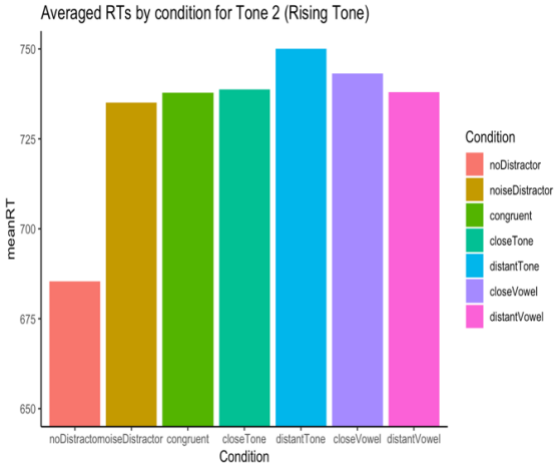


Fig. 2. Averaged RTs among 6 participants by condition for the target Tone 2.



References

[1] B. Galantucci, C. A. Fowler, and L. Goldstein, "Perceptuomotor compatibility effects in speech," *Atten. Percept. Psychophys.*, vol. 71, no. 5, pp. 1138–1149, 2009.

[2] P. Ghaffarvand Mokari, A. Gafos, and D. Williams, "Perceptuomotor compatibility effects in vowels: Effects of consonantal context and acoustic proximity of response and distractor," *JASA Express Lett.*, vol. 1, no. 1, p. 015204, 2021.

[3] K. D. Roon and A. I. Gafos, "Perceptuo-motor effects of response-distractor compatibility in speech: Beyond phonemic identity," *Psychon. Bull. Rev.*, vol. 22, pp. 242–250, 2015.

[4] K. D. Roon and A. I. Gafos, "Perceiving while producing: Modeling the dynamics of phonological planning," *J. Mem. Lang.*, vol. 89, pp. 222–243, 2016.

[5] J. Chen, C. T. Best, and M. Antoniou, "Native phonological and phonetic influences in perceptual assimilation of monosyllabic Thai lexical tones by Mandarin and Vietnamese listeners," *J. Phon.*, vol. 83, p. 101013, Nov. 2020, doi: 10.1016/j.wocn.2020.101013.

[6] A. Reid *et al.*, "Perceptual assimilation of lexical tone: The roles of language experience and visual information," *Atten. Percept. Psychophys.*, vol. 77, no. 2, pp. 571–591, Feb. 2015, doi: 10.3758/s13414-014-0791-3.

[7] Y. Zeng, R. Fiorentino, and J. Zhang, "Experimental investigation of Mandarin lexical tone production," in *ExLing 2022*, 2022, p. 201. Accessed: Mar. 13, 2024. [Online]. Available: https://exlingsociety.com/wp-content/uploads/2022/09/ExLing-2022-Paris-proceedings-1.pdf#page=185