

Phonetic Tone Divergence and Convergence in a Cue-Distractor Paradigm

Zifeng Liu^{1,2} Giuseppina Turco^{1,2} Ioana Chitoran^{1,2}

¹Université Paris Cité, France ²Laboratoire de Linguistique Formelle, CNRS, France

Introduction: Recent studies in speech production show that concurrent activation of multiple candidates in the planning stage can yield gradient deviations in production. Two patterns have been identified: 1) Convergence, in which production shifts toward the competitor [1]; 2) Divergence, in which production is driven away from the competitor [2]. Meanwhile, it remains unclear how tones are engaged in concurrent activation and how such activation affects production. Existing studies on tone planning have focused mainly on the planning [3], with few studies investigating the interaction between planning and production [4]. The present study addresses this gap by examining whether phonetic divergence or convergence emerges when target tone planning is disrupted by tone distractors in a cue-distractor paradigm [5].

Methodology: We analyzed the production data collected during a Cue-Distractor task on Mandarin tones. In this task, participants were repeatedly prompted to produce two target tones, Tone 2 (rising) and Tone 4 (falling), carried by the syllable [p^hu]. At a given time after the prompt (SOA 150/200/300 ms), a distractor was played. Each target was paired with four types of distractors: 1) Congruent distractor; 2) Two incongruent tone distractors; 3) Two incongruent vowel distractors; 4) Non-speech noise distractor (Table.1). Each target-distractor pair was repeated 15 times with 3 SOAs, yielding 540 trials (2 targets * 6 distractors * 3 SOAs * 15 repetitions). In addition, 90 trials without any distractor were included as a baseline. In total, each participant produced 630 trials. 44 native speakers of Mandarin were instructed to read out the word written in *pinyin* on the screen as quickly and accurately as possible, ignoring any sound from headphones (Fig. 1). F0 trajectories produced in tone distractor conditions were extracted using STRAIGHT in VoiceSauce [6]. The overall F0 range, F0 time-course, and scalar projection scores across conditions were measured.

Predictions: 1) No systematic F0 deviation across distractor conditions, as participants barely produced target tones inaccurately (0.61% errors) despite the distraction; 2) Convergence: F0 trajectories will move toward the distractor, consistent with [1]; 3) Divergence: F0 trajectories will exhibit increased F0 range, resisting distraction, as in [2].

Results: 1) F0 range of target tones was evaluated across distractor conditions by linear mixed-effect models. For target Tone 2, F0 range was significantly larger in the distant-tone condition than in the baseline. For target Tone 4, F0 range was significantly larger in the close-tone condition than in the baseline (Fig.2, top left). 2) Generalized additive models (GAMM) were fitted for each target tone to localize the temporal F0 deviations. For target Tone 2, GAMM revealed a significant deviation at the end of the F0 in the distant-tone condition (Fig.2, bottom left). For target Tone 4, GAMM predicted a significant F0 downward shift in the congruent condition, and a significant deviation in the second half of the F0 in the close-tone condition (Fig.2, bottom right). 3) Scalar projection scores (Fig.2, top right) further confirmed the direction of deviations. For target Tone 2, F0 remained intact under congruent conditions (near 0 values), but diverged from the distractor under incongruent conditions (negative values). For target Tone 4, F0 converged toward the distractor in the congruent condition (positive values) and diverged in incongruent conditions (negative values).

Discussion: Contrary to the first prediction, concurrent activation of tones in planning led to phonetic deviations in production, conditioned by the types of distractors. When distractors were congruent, F0 trajectories either remained stable (for target Tone 2) or converged toward distractors (for target Tone 4). When distractors were incongruent, both tones showed a divergent pattern. The gradient F0 deviations suggest that tone planning involves concurrent activation, with movement plans being continuously adjusted under distraction. Our results support a dynamic planning-production interaction in speech production.

Table 1. Target-Distractor pairs.

Conditions	Auditory Distractors	Target [p ^h u2]	Target [p ^h u4]
Congruent	Same syllable	p ^h u2	p ^h u4
Incongruent Tone	Close Tone	p ^h u1	p ^h u1
	Distant Tone	p ^h u4	p ^h u2
Incongruent Vowel	Close Vowel	p ^h o2	p ^h o4
	Distant Vowel	p ^h a2	p ^h a4
Non-speech	A beep noise	Beep	Beep
Baseline	No Distractor	NA	NA

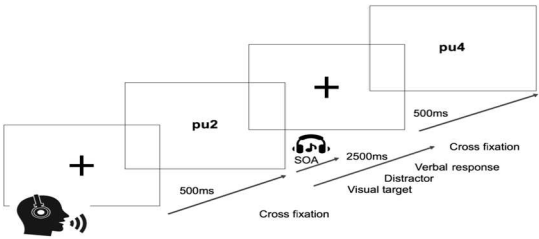


Fig. 1. Schematic timeline of a single trial in the Cue-Distractor task. The visual target displayed after a 500-ms cross. An auditory distractor was presented at SOA 150, 200, or 300 ms relative to target onset. Participants then produced the target aloud.

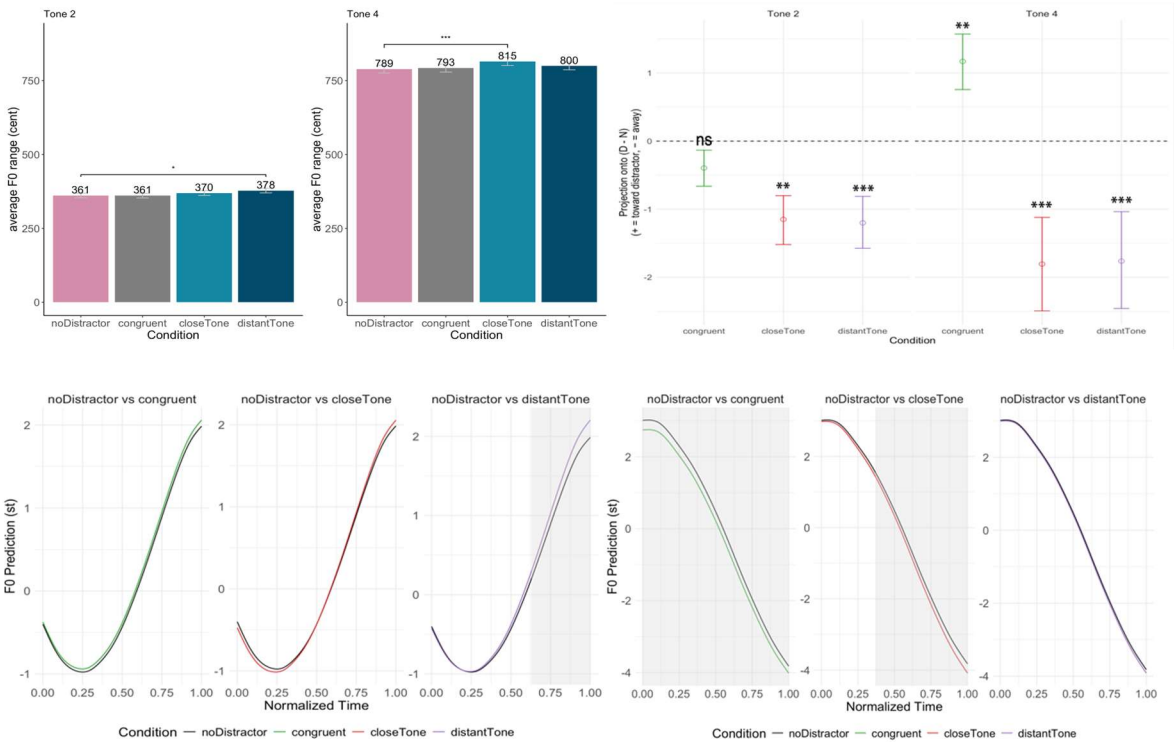


Fig. 2. Top left: mean F0 range by conditions for two target tones. Top right: scalar projection scores by conditions for target tones. Positive values indicate convergence toward the distractor, negative values indicate divergence. Bottom left: GAMM-predicted time-normalized Tone 2 trajectories by condition. Grey shading marks significant deviations. Bottom right: GAMM-predicted time-normalized Tone 4 trajectories by condition. Grey shading marks significant deviations.

References

- [1] Tobin, Stephen, Marc Hullebus, and Adamantios Gafos. "Immediate phonetic convergence in a cue-distractor paradigm." *The Journal of the Acoustical Society of America* 144.6 (2018): EL528-EL534.
- [2] Stern, Michael C., and Jason A. Shaw. "Neural inhibition during speech planning contributes to contrastive hyperarticulation." *Journal of Memory and Language* 132 (2023): 104443.
- [3] Chen, Xiacong, and Caicai Zhang. "Setting the "tone" first and then integrating it into the syllable: An EEG investigation of the time course of lexical tone and syllable encoding in Mandarin word production." *Journal of Memory and Language* 140 (2025): 104575.
- [4] Zeng, Yuyu, Chang Wang, and Jie Zhang. "Cascading activation in spoken word production drives incomplete neutralization: An internet-based study of Mandarin 3rd tone sandhi." *Journal of Phonetics* 112 (2025): 101428.
- [5] Roon, Kevin D., and Adamantios I. Gafos. "Perceptuo-motor effects of response-distractor compatibility in speech: Beyond phonemic identity." *Psychonomic bulletin & review* 22.1 (2015): 242-250.
- [6] Kawahara, Hideki, et al. "Nearly defect-free F0 trajectory extraction for expressive speech modifications based on STRAIGHT." *Interspeech*. 2005.