

A Preliminary Study on Focus and Prosodic Phrasing in Kunshan Wu

Wenwei Xu Yiya Chen

Leiden University Centre for Linguistics, Leiden Institute for Brain and Cognition

Although focus is widely understood to influence prosodic prominence, its effects on prosodic phrasing are less clearly established. Despite existing theoretical proposals [1–5], recent experimental research has yielded mixed results [6–8], and cross-linguistic evidence for a systematic focus–phrasing mapping in naturalistic language remains limited.

In Wu Chinese, tone sandhi applies obligatorily to phonological words, and phonological voicing surfaces only within sandhi domains, while domain-initial obstruents are devoiced [9]. Together, tone sandhi and voicing alternation provide reliable diagnostics for prosodic phrasing. Leveraging these cues, this preliminary study examines the extent to which contrastive focus modulates prosodic phrasing in Kunshan Wu, and whether such effects are constrained by morphosyntactic structure. To this end, we investigated locative (LC) and verb–object (VO) phrases in Kunshan Wu, both of which permit either holistic sandhi (Type I phrasing) or division into two sandhi domains (Type II phrasing) that host the first element (W1: noun/verb) and the second element (W2: locative/object) separately. Based on prior work [3, 10], we hypothesized that contrastive focus on W2 inserts a prosodic boundary before W2, leading to Type II phrasing, whereas focus on W1 promotes dephrasing, giving rise to Type I phrasing.

To encourage naturalistic production, we generated 40 pictures (24: locative; 16: activity) using ChatGPT and recorded the first author describing each picture with mismatches at W1, W2, or the entire phrase (henceforth Baseline). All stimuli were produced with Type I phrasing to ensure corrective focus is elicited by the discourse context rather than prosodic priming. Thirty Kunshan Wu speakers then completed an elicitation task, in which they verbally corrected the video–picture mismatch across 120 randomized trials (target sentences presented after the video; see **Table 1** for examples). Responses were first auditorily judged for phrasing by the first author, blind to the trial conditions. The recordings were then annotated for voicing in Praat [11]. To date, data from four speakers have been processed. **Table 2(a–c)** shows the frequency distribution of *mismatch condition* $\times$ *phrasing* for LC *on*-phrases, LC *in*-phrases, and VO phrases, respectively. **Table 2(d)** shows the frequency distribution of *mismatch condition* $\times$ *voicing* for object-initial voiced onsets in VO phrases. The agreement rate between phrasing judgment and voicing annotation is 87.5%. Statistical inferences were based on the posterior probability of direction (*pd*) estimated from Bayesian multilevel logistic models (see **Table 3**).

The results reveal distinct distributional patterns across phrase types and mismatch conditions. First, LC *on*-phrases uniformly received Type I phrasing regardless of mismatch condition, in line with the fact that the locative 浪向 ‘on’ in Kunshan is a bound morpheme without a standalone pronunciation. Among phrases that allow variable phrasing, W2-mismatch evoked more Type II phrasing for both LC *in*-phrases (W2 vs. W1/Baseline: *pd* = 1) and VO phrases (W2 vs. W1: *pd* = 1; W2 vs. Baseline: *pd* = .976), suggesting that W2-focus tends to induce boundary insertion. Under W2-mismatch, LC *in*-phrases almost always patterned as Type II phrasing (*pd* = .998), whereas VO phrases showed no clear preference for phrasing (*pd* = .790). In contrast, W1-mismatch increased Type I phrasing only for VO phrases (W1 vs. Baseline: *pd* = .983), with this preference also evident under W1-mismatch (*pd* = .994), pointing to a dephrasing effect of W1-focus. LC *in*-phrases were consistently realized with Type I phrasing under both W1-mismatch and Baseline. Finally, results from onset voicing largely corroborated those from phrasing judgment.

In sum, our findings indicate that contrastive focus only probabilistically affects prosodic phrasing (i.e., boundary insertion and dephrasing) in Kunshan Wu, with these effects conditioned by morphosyntactic structure.

Table 1. Example trials.

Phrase type	Target	Mismatch condition (stimulus)		
		Baseline	W1	W2
LC <i>in</i>	书店里向有只狗 'There is a dog in the bookstore'	大街浪向~ 'on the street'	菜场里向~ 'in the market'	书店外头~ 'outside the bookstore'
LC <i>on</i>	坐垫浪向有只狗 'There is a dog on the cushion'	盒子里向~ 'in the box'	台子浪向~ 'on the table'	坐垫下头~ 'under the cushion'
VO	小明浪听电话 'Ming is answering the phone'	~打游戏 'playing games'	~等电话 'waiting for a call'	~听音乐 'listening to the music'

Table 2. Frequency distribution of phrasing and voicing.

Phrase Condition	(a) LC <i>on</i>		(b) LC <i>in</i>		(c) VO		(d) voicing	
	I	II	I	II	I	II	+voice	–voice
Baseline	47	0	46	2 (4.2%)	40	24 (37.5%)	17	15 (46.9%)
W1	48	0	47	1 (2.1%)	55	9 (14.1%)	24	8 (25%)
W2	46	2 (4.2%) ¹	4	44 (91.7%)	26	38 (59.4%)	7	25 (78.1%)

1. Two Type II tokens from the same speaker, which were judged to be unnatural.

Table 3. Predictions of Bayesian multilevel logistic models.

Phrase Condition	LC <i>in</i> ¹		VO ¹		voicing ²	
	95% CI ³	<i>pd</i>	95% CI ³	<i>pd</i>	95% CI ³	<i>pd</i>
Baseline	[0.06%, 9.54%]	1	[6.7%, 75.3%]	.816	[13.9%, 82.2%]	.565
W1	[0.05%, 9.67%]	1	[1.0%, 33.8%]	.994	[4.97%, 58.6%]	.940
W2	[72.4%, 99.4%]	.998	[25.0%, 93.0%]	.790	[45.3%, 95.7%]	.964
W2 vs. W1	[169, 47746]	1	[6.78, 107]	1	[2.9, 82.0]	.999
W2 vs. Base.	[154, 49834]	1	[1.02, 18.0]	.976	[0.8, 29.3]	.963
W1 vs. Base.	[0.05, 15.0]	.533	[0.03, 0.72]	.983	[0.05, 1.97]	.901

1. Response: 0 = Type I, 1 = Type II;

Response ~ Condition*Phrase_type + (1|Word) + (1+Condition+Phrase_type|Participant)

2. Response: 0 = voiced, 1 = devoiced;

Response ~ Condition + (1|Word) + (1+Condition|Participant)

3. Predicted probabilities are inverse-logit transformed; between-condition contrasts are odds ratios.

References

- [1] Beckman, M. and Pierrehumbert, J. 1986. Intonational structure in English and Japanese. *Phonology* 3, 255–309.
- [2] Féry, C. 2013. Focus as prosodic alignment. *Nat Lang Linguist Theory* 31, 683–734.
- [3] Jun, S.-A. 1996. *The phonetics and phonology of Korean prosody: Intonational phonology and prosodic structure*. Garland Publishing.
- [4] Jun, S.-A. and Fougeron, C. 2000. A phonological model of French intonation. In: A. Botinis (ed.) *Intonation: Analysis, Modeling and Technology* (pp. 209–242). Kluwer Academic Publishers.
- [5] Selkirk, E. 2000. The Interaction of Constraints on Prosodic Phrasing. In: Horne, M. (eds) *Prosody: Theory and Experiment. Text, Speech and Language Technology*, vol 14. Springer, Dordrecht.
- [6] Jeon, H.-S. and Nolan, F. 2017. Prosodic Marking of Narrow Focus in Seoul Korean. *Lab. Phonol.* 8(1): 2, 1–30.
- [7] Hatcher, R., Joo, H., Kim, S., and Cho, T. 2024. Focus-induced tonal distribution in Seoul Korean as an edge-prominence language. *J. Phon.* 107, 101353.
- [8] Wagner, M. and McAuliffe, M. 2024. The effect of focus prominence on phrasing. *J. Phon.* 77, 100930.
- [9] Chen, Y. 2011. How does phonology guide phonetics in segment– f0 interaction? *J. Phon.* 39(4), 612–625.
- [10] Selkirk, E. and Sheng, T. 1990. Prosodic domains in Shanghai Chinese. In S. Inkelas & D. Zec (eds.) *The phonology-syntax*. University of Chicago Press.
- [11] Boersma, P. and Weenink, D. 2025. Praat: doing phonetics by computer. <https://www.fon.hum.uva.nl/praat/>