Regional variation in syntax modulates EEG responses in sentence processing

Zhong Chen[1], Chung-Lin Yang[2], Yuduo Cai[2], Ruitong Jiang[3]
[1] Rochester Institute of Technology, [2] University of Rochester, [3] Carnegie Mellon University
z.chen@rit.edu

Speakers across different communities express their identities through variations in language. Understanding those variations, therefore, plays a critical role in developing robust language processing theories. Much of previous neurocognitive research on variations has focused on phonological and lexical domains [1-2], while only a few have investigated the impact of morphosyntactic and syntactic variations on auditory sentence processing [3-4]. In this work, we examine the placement of the Chinese co-verb *gěi* and show that the regional syntactic variation modulates EEG responses during incremental reading comprehension. We interpret our results in the context of usage-based theories [5-7] which emphasize the connection between familiarity and linguistic behavior.

In addition to being a main verb in Chinese, *gěi* could also mark a dative structure and take on either a recipient or beneficiary [8-9]. Although Chinese syntax allows a *gěi*-phrase to be positioned before or after the main verb, a regional preference at the production level was reported based on speakers' home vernacular [10]. Chinese informants from the northern provinces produced a much higher percentage of pre-verbal *gěi*. The post-verbal variant started to emerge south towards the non-Mandarin dialectal regions and was the dominant structure favored by speakers from the southernmost areas.

Using a 2X2 design, we measured readers' EEG responses when they came across a *gěi*-phrase in either a preferred or a less preferred position. Our results include the EEG data of twenty-six Chinese-speaking college students divided into two equal groups, following Ramsey's basic North-South Division of Chinese dialects [11]. As shown in Table 1, a one-sentence context introduces an event to allow a dative structure in the target sentence. Each subject then read 84 randomized target sentences balanced between the two conditions, along with 108 filler sentences of similar lengths and complexity. Each segment in the sentence was presented on the monitor for 400ms, followed by a 400ms interval, making a total stimulus-onset asynchrony of 800ms. A true-or-false comprehension question appeared after 20% of the trials.

Figure 1 illustrates the ERP results from the left hemisphere at the pre-verbal and post-verbal *gěi*, focusing on two regions of interest (ROIs), left frontal and left posterior, and spanning 200ms pre-stimulus to 900ms post-stimulus. Similar to [4], we analyzed the ERP amplitudes in two time windows – 200–400ms and 500–900ms – to examine the (early) left anterior negativity (LAN/ELAN) and the late positivity, respectively.

In the pre-verbal condition, the southern group showed significantly stronger negativity in the 200–400ms window than the northern group in the left-frontal region ($p < .05$), suggesting that *gěi* is less expected early in the sentence and more surprising to those from the south. In contrast, speakers from the north are more familiar with *gěi* here, either as a co-verb or a main verb following the preceding adverbial phrase. In the 500–900ms window, the northern group yielded a significantly stronger late positivity than the southern group in both ROIs ($p < .05$), reflecting additional parsing effort [12] among northern speakers as they decide whether to analyze *gěi* as a dative co-verb or a main verb, since both options are equally viable given the preceding context. On the other hand, southern speakers are likely to see *gěi* as a main verb, as they are less used to a dative structure early in the sentence. In the post-verbal condition, there were no significant differences in the mean amplitude between the two groups in either time window.

Our observations align with usage-based theories such that experience with a specific construction, defined by its frequency, leads to corresponding alterations in language comprehension. In this case, the choice of where to place the Chinese co-verb *gěi* results in distinct neural activities among speakers from varied regional backgrounds.

Context: 母亲节到了，女儿和妈妈一起逛花店。 Mother's Day is coming. The daughter and her mom visit a flower shop together.
Condition a, pre-verbal *gei*, preferred by northern speakers: 女儿 | 细心地 | **给** | 妈妈 | 选了 | 一束 | 康乃馨 | 让 | 妈妈 | 笑得 | 很开心。 daughter | carefully | **gěi** | mom | selected | a bouquet of | carnations | make | mom | smile | happily
Condition b, post-verbal *gei*, preferred by southern speakers: 女儿 | 细心地 | 选了 | 一束 | 康乃馨 | **给** | 妈妈 | 让 | 妈妈 | 笑得 | 很开心。 daughter | carefully | selected | a bouquet of | carnations | **gěi** | mom | make | mom | smile | happily
The daughter carefully selected a bouquet of carnations for her mom, which made her mom smile happily.

Table 1. An example item of the experiment stimuli

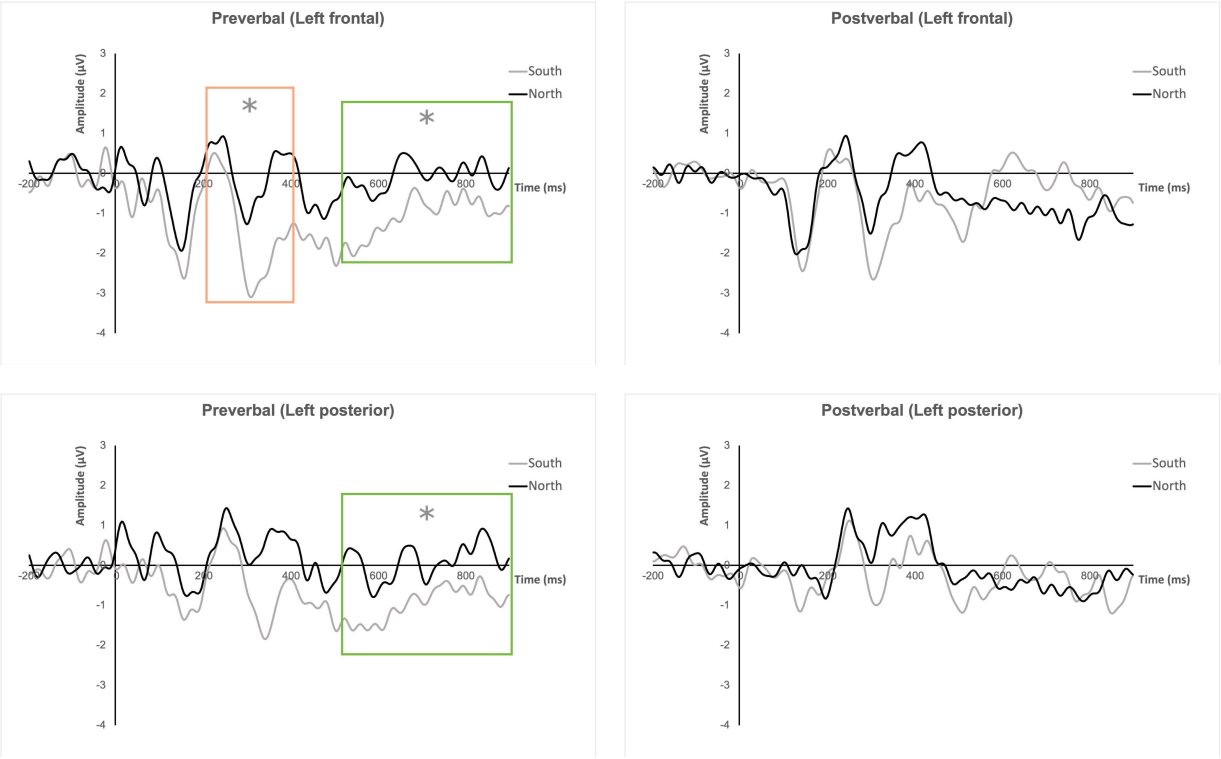

Figure 1. Grand mean waveforms for southern speakers (**gray**) compared to northern speakers (**black**) at the critical word *gěi* in either the preverbal or the postverbal condition.

References:

[1] Goslin et al., 2012, *Brain & Language*; [2] Voeten & Levelt, 2019, *Frontiers in Neuroscience*; [3] Garcia, 2017, *PhD Thesis Columbia University*.; [4] Zaharchuk et al., 2021, *Brain & Language*; [5] Bybee, 2006, *Language*; [6] Goldberg, 2006, *Oxford University Press*; [7] MacDonald, 2013, *Frontiers in Psychology*; [8] Li & Thompson, 1981, *University of California Press*; [9] Her, 2006, *Lingua*; [10] Peng, 2020, *Chinese Language and Discourse*; [11] Ramsey, 1989, *Princeton University Press*; [12] Leckey & Federmeier, 2020, *Psychophysiology*.