Structure-sensitive pronoun processing even in the absence of Principle B effects

Thuy Bui & Brian Dillon (University of Massachusetts Amherst)
tbui@umass.edu

Principle B of Binding Theory asserts that a pronoun cannot corefer with an antecedent in its local clause [4]. Previous studies have reported conflicting results regarding whether English speakers use Principle B to filter the interpretation of pronouns in real-time comprehension. While [5, 6, 11] have argued that only non-local antecedents are retrieved initially, [2] found that local antecedents can also be considered in early processing. In this study, we investigate this issue in Vietnamese, where coreference between a pronoun and a local referential antecedent is permitted (1). We ask whether structural constraints guide early pronoun resolution processes in this language, despite the lack of clear Principle B effects. We considered two possibilities:

(i) Free access to all antecedents: Vietnamese speakers could consider all grammatically licit antecedents immediately in real-time processing. They may retrieve any feature-matching antecedent, regardless of its syntactic position, as predicted by a cue-based retrieval mechanism in a content-addressable memory system [10].

(ii) Restricted antecedent search: Alternatively, Vietnamese speakers might still use structural cues to guide retrieval, and be biased away from local antecedents even without categorical Principle B effects. For instance, pronouns typically refer to the most prominent discourse referent, which is commonly the highest subject of a sentence [1, 8]. Since non-local antecedents in [2, 6] take this position, they may be more easily retrieved [see 8].

Crucially, (i) predicts that it would take longer to process a pronoun only when no antecedent matches the pronoun's features compared to any sentence in which at least one antecedent does. That is, there would be no difference between the non-local and the local antecedents, as long as they match the pronoun's features. Meanwhile, (ii) predicts that comprehenders would encounter processing difficulty only when the non-local antecedent mismatches the pronoun's features. Feature-matching local antecedents would not reduce processing difficulty, since they would not be considered initially.

We tested these predictions in an online self-paced reading experiment ($n_{participant}$ = 80, n_{item} = 36) on Ibex Farm [6]. We manipulated two factors: (i) the syntactic position of a referent relative to the pronoun and (ii) the feature match between a referent and the pronoun. We used honorificity to investigate the latter [9], since all Vietnamese pronouns are marked with honorific status (either HONorific or SUBhonorific), but not with gender. All referents had unambiguous classifiers specifying their honorificity features. We crossed [**NON-LOCAL** / **LOCAL**] with [**MATCH** / **MISMATCH**] in a 2 × 2 design, as in [2, 5; Table 1]. Mean RTs are in Figure 1, with linear mixed effects results in Table 2 and pairwise contrasts in Table 3. The data suggest three key results:

(i) A preference for non-local antecedents: Participants were immediately sensitive to an honorificity mismatch with non-local antecedents. RTs were shorter when the non-local antecedent matched the pronoun than when it did not, replicating previous work [2, 5, 6, 11].

(ii) A bias against local antecedents: RTs were slowed for a feature-matched local antecedent even when the non-local antecedent was unavailable. **NON-LOCAL MISMATCH, LOCAL MATCH** was significantly slower than **NON-LOCAL MATCH, LOCAL MISMATCH**, but was not reliably faster than **NON-LOCAL MISMATCH, LOCAL MISMATCH** in immediate spillover regions [5].

(iii) Competition for multiple matching antecedents: When both antecedents matched, processing times were slowed. This multiple match effect may reflect competition to determine a single antecedent for the pronoun. This led to additional processing time observed in the **NON-LOCAL MATCH, LOCAL MATCH** condition, contrasting with the significantly shorter RTs in the **NON-LOCAL MATCH, LOCAL MISMATCH** condition, replicating the findings in [2].

In sum, even though Vietnamese does not display robust Principle B effects, the pattern of RTs observed in this study replicates [2] closely. Our data suggest that despite the lack of Principle B effects, Vietnamese speakers still prefer non-local antecedents, showing a similar structural bias in processing to that observed in English.

(1) ***Thằng*** ***nhân*** ***viên*** ***đó*** *bầu* *cho* ***nó.***
SUB person worker that vote for SUB
'That employee-SUB voted for him-SUB.'

nó	=	***thằng nhân viên đó***
him-SUB	=	that employee-SUB

Table 1: Experimental conditions and sample materials.

NON-LOCAL	LOCAL	SENTENCES										
Match	Match	***Thằng*** SUB	***Tâm*** Tam	*nói* say	*là* that	***thằng*** SUB	***nhân*** person	***viên*** worker	***đó*** that	*bầu* vote	*cho* for	***nó.*** SUB
Match	Mismatch	***Ông*** HON	***Tâm*** Tam	*nói* say	*là* that	***thằng*** SUB	***nhân*** person	***viên*** worker	***đó*** that	*bầu* vote	*cho* for	***ổng.*** HON
Mismatch	Match	***Ông*** HON	***Tâm*** Tam	*nói* say	*là* that	***thằng*** SUB	***nhân*** person	***viên*** worker	***đó*** that	*bầu* vote	*cho* for	***nó.*** SUB
Mismatch	Mismatch	***Thằng*** SUB	***Tâm*** Tam	*nói* say	*là* that	***thằng*** SUB	***nhân*** person	***viên*** worker	***đó*** that	*bầu* vote	*cho* for	***ổng.*** HON
							trong in	*cuộc* session	*họp* meeting	*sáng* morning	*nay.* this	

'Tam-{SUB / HON} said that that employee-SUB voted for him-{SUB / HON} in the meeting this morning.'

Figure 1: Mean RTs (in ms) by condition and region. Error bars represent standard error by participants, corrected for between-participant variance [3].

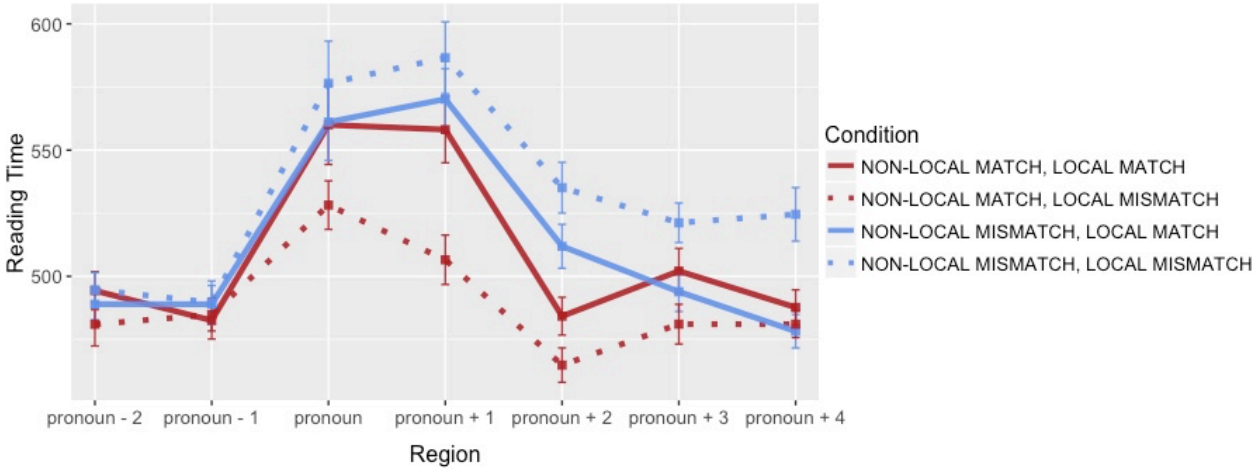

Table 2: Maximal linear mixed effects model fit to log-transformed RTs. Significant effects are boldfaced.

Fixed Effects	$t_{pronoun}$	$t_{pronoun+1}$	$t_{pronoun+2}$	$t_{pronoun+3}$
NON-LOCAL	−1.23	**−4.02**	**−5.15**	**−2.96**
LOCAL	0.89	0.96	0.55	−0.74
NON-LOCAL × LOCAL	1.40	**2.58**	**2.01**	**2.29**

Table 3: Relevant pairwise comparisons. Significant contrasts are boldfaced.

Contrasts			t_{pro+1}	t_{pro+2}	t_{pro+3}
NON-LOCAL MISMATCH, LOCAL MISMATCH	−	NON-LOCAL MISMATCH, LOCAL MATCH	0.95	1.06	**2.22**
NON-LOCAL MISMATCH, LOCAL MISMATCH	−	NON-LOCAL MATCH, LOCAL MATCH	1.89	**3.29**	**2.62**
NON-LOCAL MISMATCH, LOCAL MISMATCH	−	NON-LOCAL MATCH, LOCAL MISMATCH	**4.66**	**5.05**	**3.67**
NON-LOCAL MISMATCH, LOCAL MATCH	−	NON-LOCAL MATCH, LOCAL MISMATCH	**3.36**	**4.05**	1.56
NON-LOCAL MATCH, LOCAL MATCH	−	NON-LOCAL MATCH, LOCAL MISMATCH	**2.41**	1.83	1.25

References: **[1]** Ariel, M. (1990). *Accessing NP Antecedents*. **[2]** Badecker, W. & Straub, K. (2002). *JEP: LMC 28*. **[3]** Bakeman, R. & McArthur, D. (1996). *BRM 28*. **[4]** Chomsky, N. (1981). *Lect on Govt & Binding*. **[5]** Chow, W.-Y., Lewis, S., & Phillips, C. (2014). *Front Psychol 5*. **[6]** Cunnings, I & Sturt, P. (2018). *Lang Cogn Neurosci 33*. **[7]** Drummond, A. (2013). *Ibex Farm*. **[8]** Kush, D., Johns, C. L., & Van Dyke, J. A. (2018). *JEP: LMC*. **[9]** Kwon, N. & Sturt, P. (2016). Front Psychol 7. **[10]** Lewis, R. L. & Vasishth, R. (2005). *Cogn Sci 29*. **[11]** Nicol, J. & Swinney, D. (1989). *J Psycholinguist Res 18*.