Two Steps to Parasitic Agreement in Hindi-Urdu

Kenneth Hanson

Stony Brook University

1 Overview

Hindi-Urdu (HU) is famous for its verbal agreement system, which involves case-sensitivity, long-distance/cross-clausal agreement (LDA), and **parasitic agreement**, in which non-finite verbs (embedded or otherwise) agree if and only if the finite verb does. This last item is **difficult to motivate** under common Minimalist assumptions. If non-finite verbs bear unvalued ϕ-features which probe, then it should not matter whether or not some higher item also agrees. If they do not, then it is not clear why they should agree at all.

Here, I build on Bhatt's (2005) intuition that all verbs are valued together when the probe on finite T finds a goal by providing an analysis in the formal system of Hanson (2024a,b). Two logical steps are needed: 1) finite T agrees with the closest visible DP, and 2) all verbs along this path agree. Each of these is a **tier-based strictly 2-local** (TSL-2) pattern, like many others in phonology and syntax (Graf 2022b), and each on its own is unexceptional. This provides a **potential explanation** for this otherwise puzzling phenomenon: if language is capable of producing each pattern independently, we expect them to show up together at least occasionally.

2 Basic Data

This data comes from Bhatt (2005). The finite verb/auxiliary agrees with the structurally highest nominative/unmarked DP in its domain, which may be the matrix subject (1), matrix object (2), or embedded object (3). Any participles/infinitives agree with the same DP. If there is no potential goal, default (MSG) agreement occurs (4).

(1) Subject agreement (unmarked subject/object)
Rahul *kitaab* *paṛh-taa* *thaa*
Rahul.M book.F read-HAB.**MSG** be.PST.**MSG**
'Rahul used to read (a/the) book.'

(2) Object agreement (ERG sbj. + unmarked obj.)
Rahul-ne ***kitaab*** *paṛh-ii* *thii*
Rahul-ERG **book.F** read-PFV.**F** be.PST.**FSG**
'Rahul had read the book.'

(3) LDA (ERG sbj. + unmarked embedded object)
Vivek-ne *[**kitaab** paṛh-**nii**]* *chaah-**ii***
Vivek-ERG **book.F** read-INF.**F** want-PFV.**FSG**
'Vivek wanted to read the book.'

(4) Default agreement (ERG subject + ACC object)
Rahul-ne *kitaab-ko* *paṛh-**aa*** ***thaa***
R-ERG book-ACC read-PFV.**MSG** be.PST.**MSG**
'Rahul had read the book.'

LDA appears to be optional in some sentences, but this is only apparent. Bhatt (2005) shows that restructured clauses require LDA, while other infinitives block it. Following Keine (2019), I assume the former to be *v*Ps, and the latter TPs, as in (5–6).

(5) *v*P allows LDA to cross
Ram-ne *[$_{vP}$* ***roṭii*** *khaa-**nii**]* *chaah-**ii***
R-ERG **bread.F** eat-INF.**F** want-PFV.**FSG**
'Ram wanted to eat bread.'

(6) No LDA across TP
Ram-ne *[$_{TP}$* *roṭii* *khaa-**naa**]* *chaah-**aa***
R-ERG bread.F eat-INF.**M** want-PFV.**MSG**
'Ram wanted to eat bread.'

(5–6) also demonstrate parasitic agreement; finite agreement is impossible without infinitival agreement and vice versa (illicit examples omitted for brevity). Finally, if the matrix subject is nominative, LDA is blocked, but the matrix participle does agree. The basic generalization is therefore: every verb (finite or not) agrees iff it occurs along the **path** from finite T to its goal.

3 TSL-2 patterns

A pattern is TSL if it can be described with **strictly local constraints** over a **tier of salient elements**, treating the rest as invisible. For TSL-2, each constraint may reference only two adjacent elements on the tier.

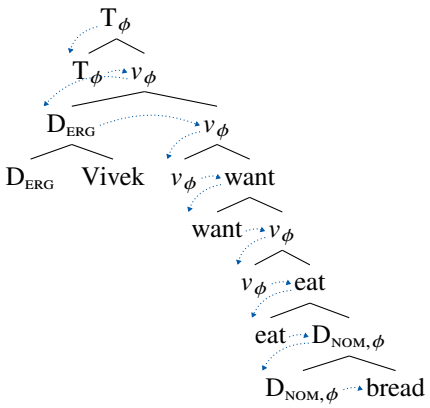
	Word	Tier
✓	sakasa	ss
✓	ʃakaʃa	ʃʃ
✗	sakaʃa	sʃ
✓	sataʃa	stʃ

Example. Suppose we have a language with sibilant harmony which is blocked by [t], similar to Slovenian. The tier consists of all segments in the set {s, ʃ, t}, which are the sibilants and the blocker. All other segments are invisible. The constraints on the tier are {*sʃ, *ʃs}. As a result, harmony is enforced except when [t] intervenes.

Notice: TSL-2 enforces a model of **relativized locality** in which a single intervener of the right type is enough to break any long-distance dependency. Although originally proposed for phonology, this is exactly what we want for syntax as well (e.g. islands, horizons, minimality).

4 Constructing the tier

Agreement proceeds in two steps, each with its own tier. Here, we focus on Step 2 (concord). Following Hanson (2024b) we analyze the string representing the **complement spine** of the tree, which corresponds to the *search path of the probe* on finite T (Hanson 2024a). On the right is a bare phrase structure diagram for (3) at the relevant point in the derivation. A ϕ diacritic is added to items which agree successfully *in the present derivation*. Omitting the technical details, we follow the complement spine, adding each head *in the order first encountered* (the position of the maximal projection) producing:



$\mathbf{T_{FIN,\phi}} \cdot \boldsymbol{v_\phi} \cdot \mathrm{D_{ERG}} \cdot \mathrm{want} \cdot \boldsymbol{v_\phi} \cdot \mathrm{eat} \cdot \mathbf{D_{NOM,\phi}} \cdot \mathrm{bread}$

We construct the tier containing all items *which can bear the diacritic*, which are {$\mathrm{T_{FIN}}$, Aux, *v*, $\mathrm{D_{NOM}}$}, plus any blockers, including $\mathrm{T_{INF}}$. (PRO is omitted, but would be invisible if present.) When agreement is successful, we have a chain of agreeing elements on the tier, starting with T and ending with D. When agreement fails, whether because there is no visible goal or because infinitive T intervenes, no elements in this chain bear the ϕ diacritic. This is summarized below.

Ex.	Configuration	Path (Tier elements highlighted)
1	Sbj. Agr.	$\mathbf{T_{FIN,\phi}} \cdot \mathbf{Aux_\phi} \cdot \boldsymbol{v_\phi} \cdot \mathbf{D_{NOM,\phi}} \cdot \mathrm{read} \cdot \mathbf{D_{NOM}} \cdot \mathrm{book}$
2	Obj. Agr.	$\mathbf{T_{FIN,\phi}} \cdot \mathbf{Aux_\phi} \cdot \boldsymbol{v_\phi} \cdot \mathrm{D_{ERG}} \cdot \mathrm{read} \cdot \mathbf{D_{NOM,\phi}} \cdot \mathrm{book}$
3	LDA	$\mathbf{T_{FIN,\phi}} \cdot \boldsymbol{v_\phi} \cdot \mathrm{D_{ERG}} \cdot \mathrm{want} \cdot \boldsymbol{v_\phi} \cdot \mathrm{read} \cdot \mathbf{D_{NOM,\phi}} \cdot \mathrm{book}$
4	Dflt. (No goal)	$\mathbf{T_{FIN}} \cdot \mathbf{Aux} \cdot \boldsymbol{v} \cdot \mathrm{D_{ERG}} \cdot \mathrm{read} \cdot \mathrm{D_{ACC}} \cdot \mathrm{book}$
6	Dflt. (Blocked)	$\mathbf{T_{FIN}} \cdot \boldsymbol{v} \cdot \mathrm{D_{ERG}} \cdot \mathrm{want} \cdot \mathbf{T_{INF}} \cdot \boldsymbol{v} \cdot \mathrm{eat} \cdot \mathbf{D_{NOM}} \cdot \mathrm{bread}$

Now, our task is to show that we can correctly distinguish well-formed and ill-formed tiers by means of strictly local constraints.

5 The constraints

As discussed above, agreement is all or nothing. We enforce this by banning mismatched pairs of (non-)agreeing elements within the chain from T to D:

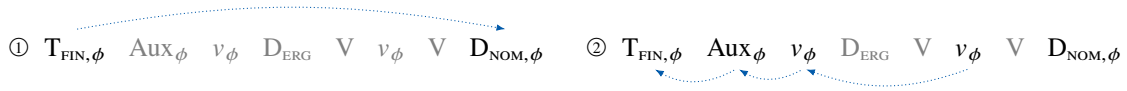
$$\{ *\mathrm{T_{FIN,\phi}} \cdot v, *\mathrm{T_{FIN}} \cdot v_\phi, *v_\phi \cdot v, *v \cdot v_\phi, *v_\phi \cdot \mathrm{D_{NOM}}, *v \cdot \mathrm{D_{NOM,\phi}}, *v_\phi \cdot \mathrm{T_{INF}}, *\mathrm{T_{INF}} \cdot v_\phi, \ldots \}.$$

This can be reduced to two constraint templates which encode the relevant generalization:

$$\{ *X_{\phi} \cdot Y, *X \cdot Y_{\phi} \}$$

As a result, agreement in sentences like (6) is impossible, since T_{INF} projects yet never agrees. Next, to rule out agreement without a goal (a nonsensical version of (4)), we add $* v_{\phi} \cdot \ltimes$, where $\ltimes$ indicates end-of-string. Finally, to ensure that agreement takes place when possible in (1/2/3), we need a tier which contains just $T_{\pm FIN}$ and D_{NOM}. On this tier we ban $* T_{FIN} \cdot D_{NOM}$, where T fails to agree with a visible goal. This cannot be done on the existing tier, since all bigrams in, e.g., $T_{FIN} \cdot Aux \cdot v \cdot D_{NOM}$ are licit with respect to our concord rule.

To summarize: the tier containing just $\{T_{\pm FIN}, D_{NOM}\}$ sets up the relation between T and its goal, and the tier which also includes the verbs handles feature spreading.



Note that there is no temporal ordering relation between the tiers themselves. The formal analysis is therefore compatible with a checking implementation which operates in either direction. If a valuation analysis is desired, then LDA involves the combination of a top-down and bottom-up process, analogous to certain phonological processes (see below).

6 Typology, cross-domain parallels

As discussed by Bhatt (2005), HU has relatives in which infinitive agreement is not parasitic. In such languages, $T_{INF} \cdot v_{\phi}$ is not banned; also, the two tiers can be collapsed. Similar variation can be observed in extraction morphology along a movement path (Graf 2022a). It is unclear to me to what degree phonological processes mirror parasitic agreement, but certain unbounded circumambient processes (Jardine 2016; McCollum et al. 2020), which can be seen as the combination of a left-to-right and a right-to-left process, might be comparable.

7 Conclusion

Parasitic agreement is a natural outcome of the space of possibilities afforded by TSL-2 computations. Simple agreement involves AB pairs on a tier; bounded concord allows ACB, ACCB, ACCCB, etc. Putting the two together creates parasitic agreement. By analyzing linguistic patterns in this way, we gain insight into otherwise mysterious phenomena.

References

Bhatt, Rajesh (2005). Long Distance Agreement in Hindi-Urdu. *Natural Language & Linguistic Theory* 23, pp. 757–807.

Graf, Thomas (2022a). Diving deeper into subregular syntax. *Theoretical Linguistics* 48.3–4, pp. 245–278. DOI: 10.1515/tl-2022-2043.

Graf, Thomas (2022b). Subregular linguistics: bridging theoretical linguistics and formal grammar. *Theoretical Linguistics* 48.3–4, pp. 145–184. DOI: 10.1515/tl-2022-2037.

Hanson, Kenneth (2024a). A Tier-Based Model of Syntactic Agreement. *Proceedings of CLS 60*. Chicago, IL: The Chicago Linguistic Society.

Hanson, Kenneth (2024b). Tier-Based Strict Locality and the Typology of Agreement. *Journal of Language Modeling*. To appear.

Jardine, Adam (2016). Computationally, tone is different. *Phonology* 33.2, pp. 247–283. DOI: 10.1017/S0952675716000129.

Keine, Stefan (2019). Selective Opacity. *Linguistic Inquiry* 50.1, pp. 13–62. DOI: 10.1162/ling_a_00299.

McCollum, Adam G. et al. (2020). Unbounded circumambient patterns in segmental phonology. *Phonology* 37.2, pp. 215–255. DOI: 10.1017/S095267572000010X.