

INVESTIGATING SCRAMBLING IN SYNTACTIC STRUCTURE OF URDU

Muhammad Farman

Lecturer SLA, UMT, LHR. University of Management and Technology.

muhammad.farman@umt.edu.pk

Mehwish Ahsan

MPhil Linguistics Scholar, Lahore Leads University. mehwishahsan422@gmail.com

Alisha Mahnoor

MPhil Linguistics Scholar, Lahore Leads University. alishamahnoor84@gmail.com

Abstract

Scrambling refers to the displacement of constituents from their canonical positions without altering core grammatical relations. This study investigates the scrambling within the syntactic structure of Urdu language within the theoretical framework of Minimalist Program, focusing on argument dislocation and the locality of movement. As a native speaker of Urdu, the researcher compiled a corpus of approximately 1500 naturally occurring sentences through purposive sampling of informal spoken discourse. Researcher selected 30 sentences from this dataset purposively for detailed syntactic analysis based on the presence of no canonical word order and constituent displacement. The findings reveal that scrambling in Urdu is a phase-bound, interface-convergent operation, and successive cyclic regulated by the phase heads C^0 and v^0 , requiring no additional projections such as $AgroP$, or other external factors. Movement is driven solely with core syntactic operations — Merge and Move — within the CP and VP domains. This argues, scrambling in Urdu is not an optional surface phenomenon but reflects universal computational principles of grammar. These results challenge antisymmetric and derived-SOV analysis, offering empirical support for a fully phase-driven model of displacement and underscoring the autonomy of syntactic computation in determining surface word order.

Keywords: Minimalist, Scrambling, Wh-movement, Remnant-VP, Weak crossover

Introduction

Scrambling describes the syntactic phenomenon where constituents, especially nominal like subjects, objects, or adjuncts, deviate from their canonical clause positions. This phenomenon can be found in many typologically diverse languages, but it is particularly common in Indo-Aryan languages like Urdu. Traditionally, scrambling has been considered “free” or optional as marking stylistic or discourse-driven variation that lacks a substantial syntactic impact (Mahajan, 1990). Still, these descriptions simplify the behavior of the structure and hide the underlying governing principles that bound it. In this paper, we address this problem by studying post-verbal scrambling in Urdu and claim that it is neither devoid of syntax nor uninterpretable. Scrambling, rather, is an operation induced by a derivation and is restricted by the internal syntactic rules as well as the interface requirements (Chomsky, 1995).

In generative grammar, the idea of syntactic structure refers to the hierarchical order of constituents formed by operations like Merge and Move (Chomsky et al., 2002) introduced the Program as Minimalist emphasizing that all operations within syntax must be motivated by formal features, constrained by principles of economy, and executed at the relevant interfaces, which in this case is Logical Form (LF) and Phonological Form (PF) – interfaces. Scholars like Dayal have advanced the debate on the intricacies of scrambling in South Asian languages arguing that movement operations within Indo-Aryan languages are not random but rather part of a system which is hierarchically organized. Their work captures the insight that scrambling, even in so-called “free word order” languages, is formally controlled and restricted by structural boundaries and interpretive requirements, thereby refuting the long-standing claim that such reordering is extra-syntactic (Lechner, 1998).

The syntactic structure of Urdu is better known as configurational than as flat. Urdu does permit considerable variation in the surface word order, but it keeps the core syntax with a fixed SOV (Subject–Object–Verb) hierarchy. This is shown by binding relations, scope interpretation, and

clause-internal dependencies. For example, many arguments may co-occur in the post-verbal field, and arguments can come in between the verb and the auxiliary, which is indicative of strict underlying syntactic rules. Such behavior cannot be accounted to surface freedom only; rather, it suggests a concrete layering with defined linear precedence intertwining with phases of derivation. Thus, the lack of variation in Urdu poses no threat to sign a structural slack, rather it depicts advanced principles that govern the movement, locality, and legibility of the interface (Saito, 1989).

This study is situated within the Minimalist framework, paying special attention to Derivation by Phase (Kidwai, 2000; Simpson & Bhattacharya, 2003). Under this approach, phonetic and semantic processes take place in discrete phases like vP or CP, with each phase being successively integrated into the system in a cyclic fashion. For us, scrambling in Urdu is recognized as an intricately bounded remnant movement process that takes place among the phases. Using modern approaches from minimalist theories, such as Labeling theory and Transfer, we justify why some positions post-verbal are valid syntactically but lacking in meaning. When scrambling is characterized as an operation constrained by linear precedence rather than movement, we show that the escape is neither optional nor semantically irrelevant at face value. It is instead a tightly controlled internal phenomenon shaped by the system's grammatical structure (Mahajan, 2011).

This study focuses on how prior models miss the fact that post-verbal scrambling in Urdu is actually controlled and not freely allowed. These accounts miss out on three important items: (1) scrambling of elements after the verb affects how scope and binding work; (2) numerous arguments can occur in post-verbal positions without breaking grammar rules; and (3) a few arguments may end up between the main verbs and the auxiliaries. They force people to rethink previous analyses and design a system able to explain hierarchy as well as shifts in word order. Through improving and anchoring remnant movement theory to minimalist ideas, the study offers a both unified and explanatory view of post-verbal displacement in Urdu (Chomsky, 2004).

The previous paragraph states that Urdu scrambling is influenced by surface linear order, visibility at the interface, and is not purely hierarchical. It is often the case that rightward displaced constituents do not reconstruct for scope or binding, which means that they are likely bound by a clause's surface structure. For example, post verbal Wh-elements are syntactically interfaced, but are interrogatively inert when post-auxiliary and thus stratum inert. No interpretation is possible post-augmentation in structure. Such elements must also however not be framed transclusively as "licensable." It underscores the importance of the interface alongside syntactic interpretation. It also indicates the frameworks arguably do not pay much attention towards piece of word structured surface, which is fundamental in construing semantics.

The argument is supported by Wh-scope ambiguity, binding contrasts, and argument order variation in matrix and embedded clauses which are all empirically corroborated by Urdu. These simplified observations are conceptually placed within bounds of minimalist syntactic theory with particular focus directed to phase theory's locality restrictions (Dayal, 1994).

Literature Review

In 2011, (Raza et al., 2011) examined clause-level NP discontinuity in Urdu, which has been cataloged in the literature with such labels as extraposition, extraction, and quantifier float. While many have longitudinally acknowledged these gaps within clauses, Raza is cited as one of the initial scholars to document constituent-level NP discontinuity where the stems of noun phrases do not join together in a contiguous manner. In Urdu, this type of discontinuity occurs at one joint level in a clause, where an NP with an internal head cannot come before its

arguments. Such facts were difficult to account for within classical syntactic frameworks, which is why a c-structure was proposed to be flat within Lexical Functional Grammar (LFG) grammar, which departed from ParGram's assumptions of hierarchically structured branching. In 2023, (Ali & Malik, 2023) investigated Tense Projection (TP) in the context of Feature Sharing by Pesetsky and Torrego (2007), applying the Head Movement Constraints (HMC) by Travis (1984). Their analysis showed that unlike English, T-to-V or V-to-T movement is interfered with by other constituents in Urdu. Most importantly, their study showed that the verb in Urdu does not lexically inflect for tense features, but for aspectual distinctions: habitual (ta), imperfective (raha), perfective (chuka), and progressive (raha), while tense is realized by independent constituents (present: hai, past: tha). These results strongly supported the feature sharing model with regard to the behavior of inflection in Urdu.

(Bhatt & Dayal, 2007) also studied word order and scrambling in Urdu which argued against the SVO analysis put forth by Mahajan (1997) and Simpson and Bhattacharya (2003). Through data on rightward scrambling, they proposed an account based on remnant-VP movement capturing two pivotal aspects: the relationship between linear order and scope, and the severely limited scope of rightward scrambled wh-expressions. That model differed from prior ones which assumed rightward movement of individual arguments, or treated scrambling as argument stranding within antisymmetric frameworks. They accounted for constraints on interpretation that previous models did not address.

In a broader theoretical context, (Haider, 2021) emphasized that grammar has a distinctively modular character, contending that phenomena such as scrambling are more appropriately understood as consequences of relational activity between different subsystems, particularly between IS and syntactic structure. They claimed scrambling is 'pragmatically utilized' rather than syntactically 'triggered,' drawing upon cross-linguistic evidence from Germanic and Slavic languages. Along these lines, Urdu scrambling is viewed as lacking a syntactic basis; rather, it stems from the language's structural indeterminacy wherein pragmatic subsystems take advantage of such indeterminacy. Their account supported the claim that scrambling does not give rise to some element of a syntactic dependency but rather emerges from interaction across modules.

(Kareem & Yaseen, 2023) has analyzed the descriptive and theoretical aspects of scrambling in Central Kurdish while characterizing scrambling as the movement of constituents from their base, neutral positions within a sentence into marked positions in the sentence. "Doing the Minimalist Program", they looked into which constituents underwent scrambling, what kind of movements and syntactic positions, what level of discourse semantics, and what type of constituents were scrambled. They showed that all arguments and adjuncts such as DP, PP, VP, and adjunct CP were all subject to scrambling. Central Kurdish exhibited short and long distance as well as mid distance scrambling, and diagnostic tests such as binding and crossover effects as well as parasitic gaps were used to try and determine the nature of each movement type. Discourse-wise, they observed that scrambling did not have any effect on the compositional meaning of the sentence, but rather the new meaning derived from the sentence exposed the interaction between syntax and discourse.

Methodology

This study employs a qualitative approach grounded in the Minimalist Program to investigate scrambling in Urdu syntax. As a native speaker, the researcher compiled a corpus of approximately 1500 naturally occurring sentences from informal spoken interactions. From this corpus, researcher selected 30 sentences from this dataset purposively for detailed syntactic analysis based on the presence of no canonical word order and constituent displacement. The analysis has applied the core Minimalist Operations; merge, move and phase theory emphasizing the role of phase heads (V^0 and C^0) in bounding displacement. This

methodological approach syndicates introspective linguistic competence with empirical data to provide a grounded and theoretically informed analysis of scrambling in Urdu.

Results and Discussion

The investigation seeks to examine the rightward movement scrambling configuration in Urdu, paying special attention to how it occurs in naturalistic speech. The results show that rightward scrambling is not left to chance; there are clear syntactic orders that it follows. First, all arguments as a subject-value, object, or indirect object can follow the main verb. Then, it is also possible for several arguments to occur at the same time in post-verbal positions within a single clause, which suggests that the rightward movement scrambling is multi-constituent. Third, arguments are often found in the gap between the verb and the auxiliary, which implies that the post-verbal domain is structured in Urdu. These accounts illustrate that rightward movement or scrambling is executed within certain parameters and cannot be reasoned solely with discourse or phonology. Evidence collected using informal dialogues indicates that rightward movement or scrambling in the language is syntactically motivated. The results are as follows;

(1) a. O V Aux S

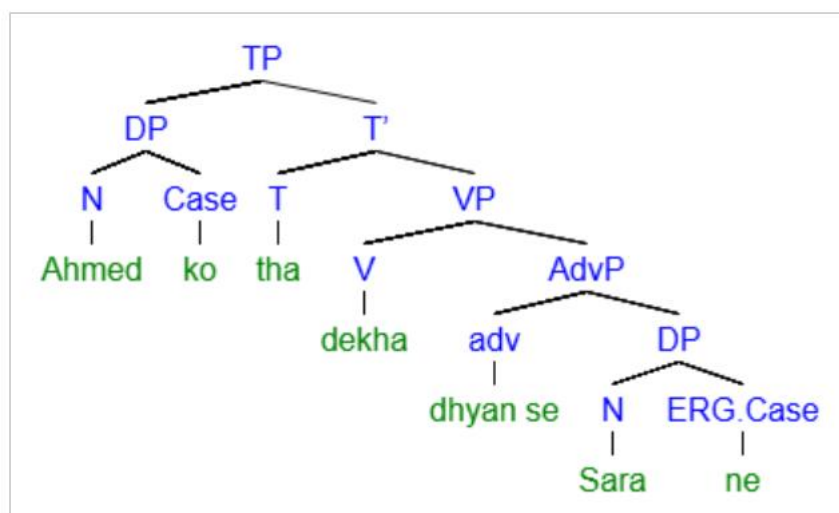
Ahmad-ko dhyān-se dekhā thā Sara-ne.

Ahmad-ACC care-with see.PFV be.PST Sara-ER

[TP[DP[N Ahmed]][ERG.Case ko]][T'[T tha]][VP[V dekha]][AdvP[adv dhyān se]][DP[N Sara]][ERG.Case ne]]

'Sara had looked at Ahmad carefully.'

Figure 4.1. a. O V Aux S



b. S V Aux O

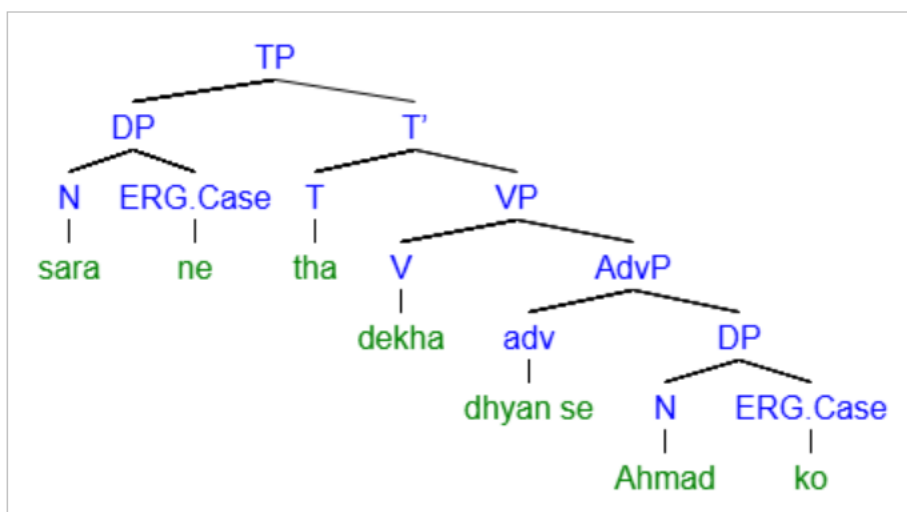
Sara-ne dhyān-se dekhā thā Ahmad-ko.

Sara-ERG care-with see. PFV be.PST Ahmad-ACC

'Sara had looked at Ahmad carefully.'

[TP[DP[N sara]][ERG.Case ne]][T'[T tha]][VP[V dekha]][AdvP[adv dhyān se]][DP[N Ahmad]][ERG.Case ko]]

Figure 4.2. S V Aux O



c. S DO V Aux O

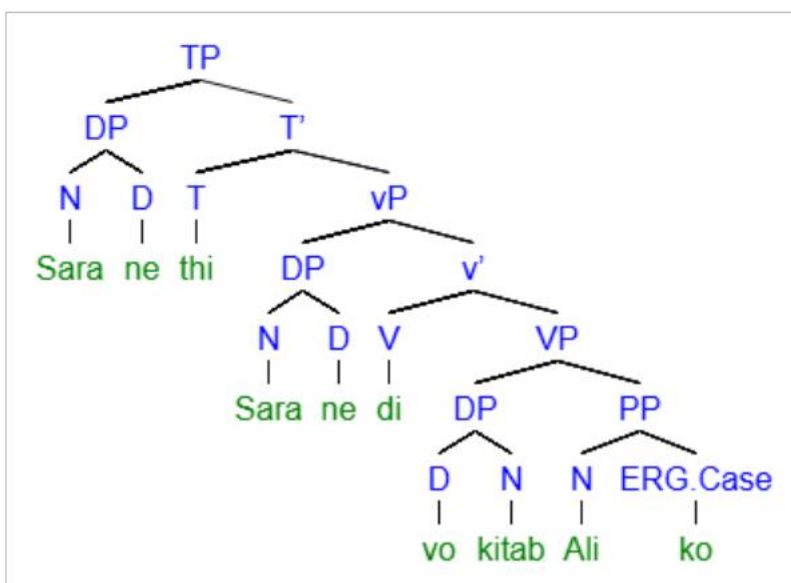
Sara-ne vo kitaab dī thī Ali-ko.

Sara-ERG that book.F give.PFV.F be.PST.F Ali-DAT

‘Sara had given that book to Ali.’

[TP [DP[N Sara][D ne]][T'[T thī][vP [DP[N Sara][D ne]][v'[V dī][VP[DP[D vo][N kitaab]][PP[N Ali][ERG.Case ko]]]]]

Figure 4.3. S DO V Aux O



(2) Multiple constituents post-verbally

DO V Aux S IO

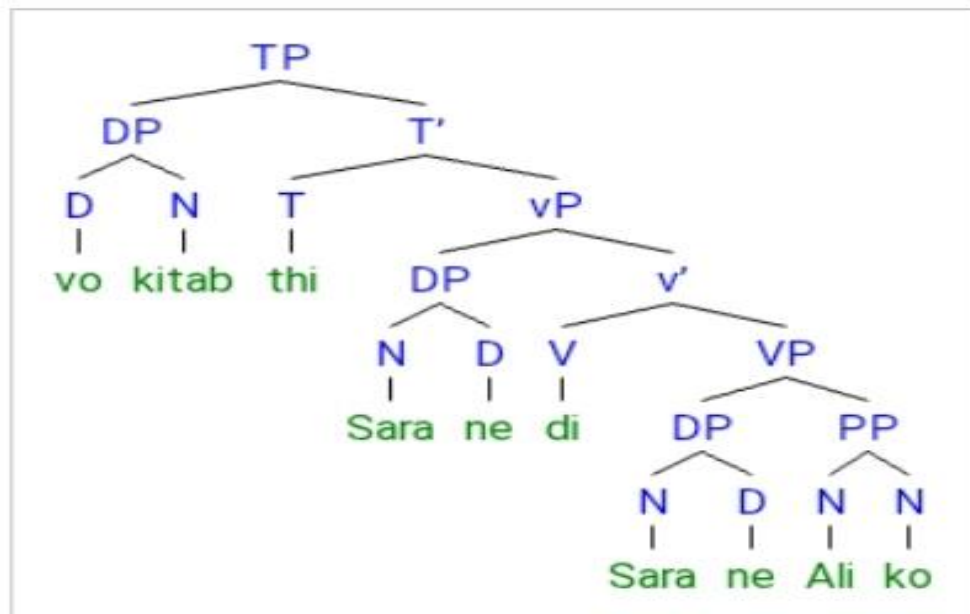
Vo kitaab dī thī Sara-ne Ali-ko.

that book.F give.PFV.F be.PST.F Sara-ERG Ali-DAT

‘Sara had given that book to Ali.’

[TP [DP vo kitaab] [T' [T thī] [vP [DP Sara-ne] [v' [V ∅] [VP [DP t Sara NE] [V' [V dī] [PP Ali ko]]]]]

Figure 4.4. DO V Aux S IO



(3) Argument between main verb and auxiliary

S DO V IO Aux

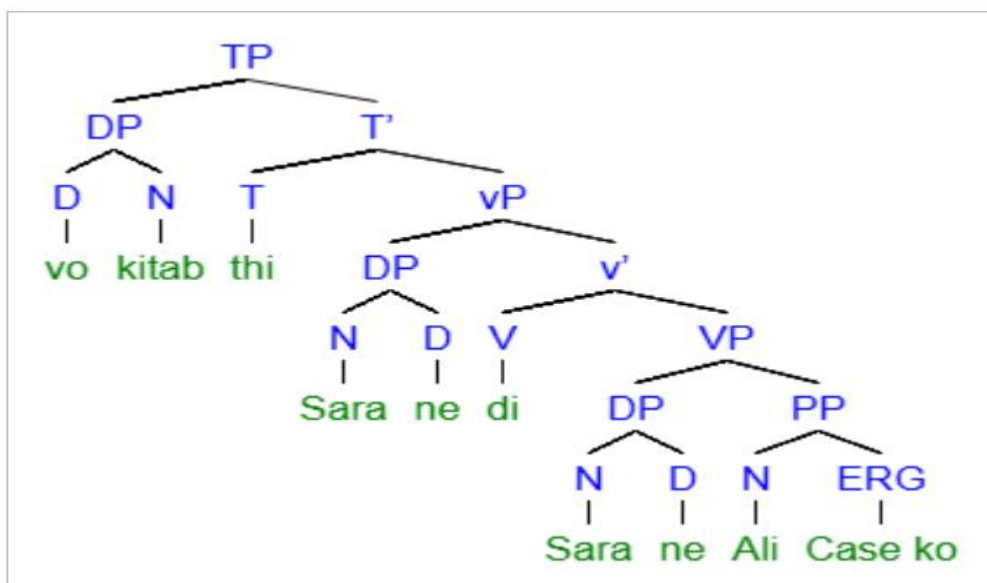
Sara-ne kitaab bhej-ī Ali-ko thī.

Sara-ERG book.F send.PFV.F Ali-DAT be.PST.F

‘Sara had sent the book to Ali.’

[TP[DP[D vo][N kitab]][T'[T thi][vP [DP[N Sara][D ne]][v'[V di][VP[DP[N Sara][D ne]][PP[N Ali][ERG Case ko]]]]]

Figure 4.5. S DO V IO Aux



4.1 Linear Order and Hierarchical Relations in Urdu

This section looks into the impact of linear order in scope sensitive phenomena in Urdu, particularly regarding scrambling operations and their consequences on interpretation. Rightward scrambling in Urdu, contrary to leftward scrambling, does not permit weak

crossover (WCO) configurations or new binding opportunities. As in the case of Hindi, there are no interpretations available for rightward scrambles. These results provide evidence that even with the relatively freer word order in Urdu, surface linear order is still important for the syntactic and interpretive relations for co arguments.

Consider the following contrast involving WCO effects:

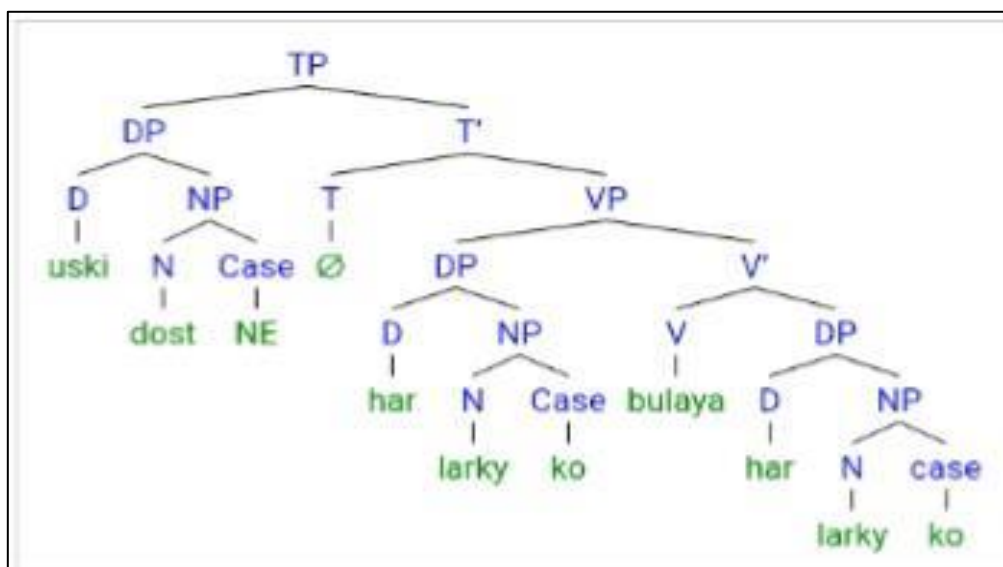
(1) a. [Us-kei dost-ne] [har larkay-ko] bulaya.

he-GEN friend-ERG every boy-ACC called

‘*His friend called every boy.

[TP[DP[D uski][NP[N dost][Case NE]]][T'[T Ø][VP[DP[D har][NP[N larkay][Case ko]]][V'[V bulaya][DP[D har][NP[N larkay][case ko]]]

Figure 4.6. S DO V



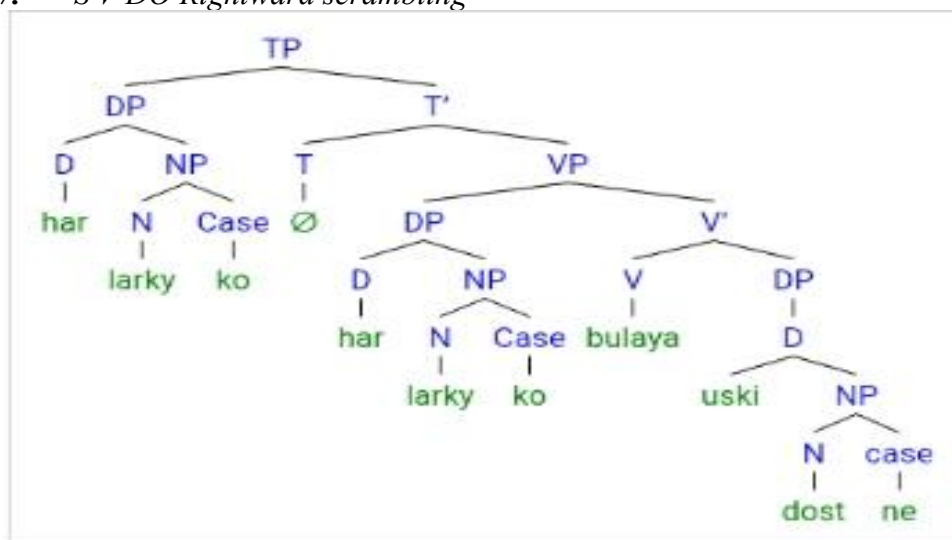
b. [Us-kei dost-ne] bulaya [har larkay-ko].

he-GEN friend-ERG called every boy-ACC

‘*His friend called every boy.’

[TP[DP[D har][NP[N larkay][Case ko]]][T'[T Ø][VP[DP[D har][NP[N larkay][Case ko]]][V'[V bulaya][DP[D uski][NP[N dost][case ne]]]

Figure4.7. S V DO Rightward scrambling



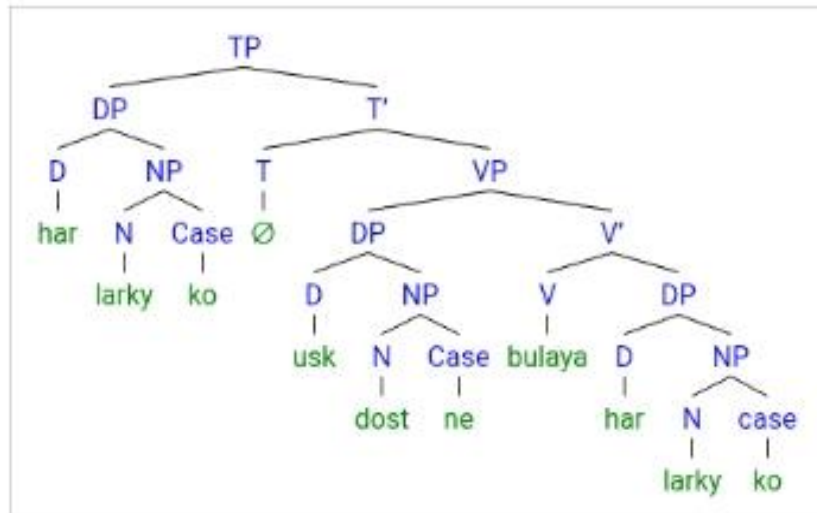
c. [Har larkay-ko][us-kei dost-ne] bulaya.

every boy-ACC he-GEN friend-ERG called

‘His friend called every boy.’

[TP[DP[D har][NP[N larky][Case ko]]][T'[T Ø][VP[DP[D usk][NP[N dost][Case ne]]][V'[V bulaya][DP[D har][NP[N larky][case ko]]]

Figure 4.8. *DO S V leftward scrambling*



In (1a) and (1b), the pronominal part of the subject phrase cannot be bound by the quantificational object, whether the object is following the verb canonically or scrambled to the right. However, (1c) that features scrambling of the object to the left permits the intended binding relation. These patterns suggest that only leftward scrambling allows reordering for interpretational causations, implying that rightward scrambling does not reconstruct at LF for scope or variable binding.

Similar insights arise from the data with reciprocal binding in (2) where only leftward scrambling allows the object to bind into the subject, demonstrating direction sensitive asymmetry in the syntactic licenser of reciprocal dependencies.

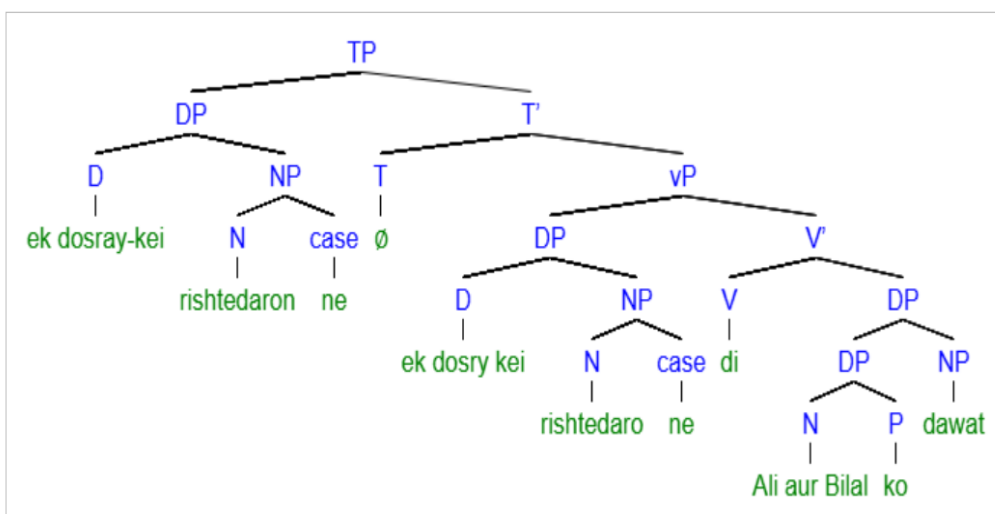
a. ???[Ek dosray-kei rishtedaron]-ne [Ali aur Bilal]-koi dawat di.

each.other-GEN relatives-ERG Ali and Bilal-ACC invitation gave

‘???Each other’s relatives invited Ali and Bilal.’

[TP [DP [D ek dosray-kei][NP[N rishtedaron][case ne]]][T'[T Ø][vP[DP[D ek dosry kei][NP[N rishtedaro][case ne]]][V'[V di][DP [DP[N Ali aur Bilal][D ko]]][NP dawat]

Figure 4.9. *S DO V*



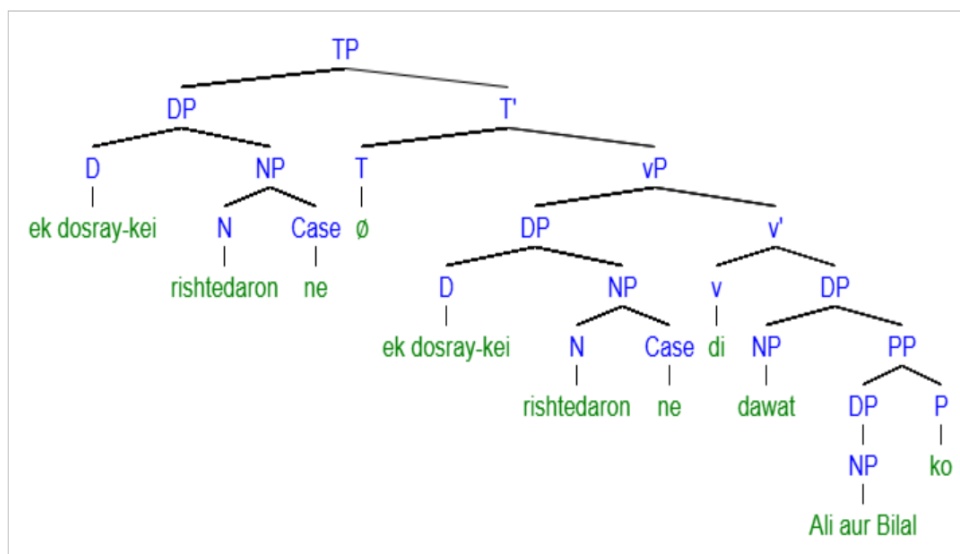
b.???[Ek dosray-kei rishtedaron]-ne dawat di [Ali aur Bilal]-koi.

each. other-GEN relatives-ERG invitation gave Ali and Bilal-ACC

‘???Each other’s relatives invited Ali and Bilal.’

[TP [DP [D ek dosray-kei] [NP [N rishtedaron] [Case ne]]] [T' [T \emptyset] [vP [DP [D ek dosray-kei] [NP [N rishtedaron] [Case ne]]] [v' [v di] [DP [NP dawat] [PP [DP [NP Ali aur Bilal]] [P ko]]]]]]]

Figure 4.10. *S V DO Rightward scrambling*



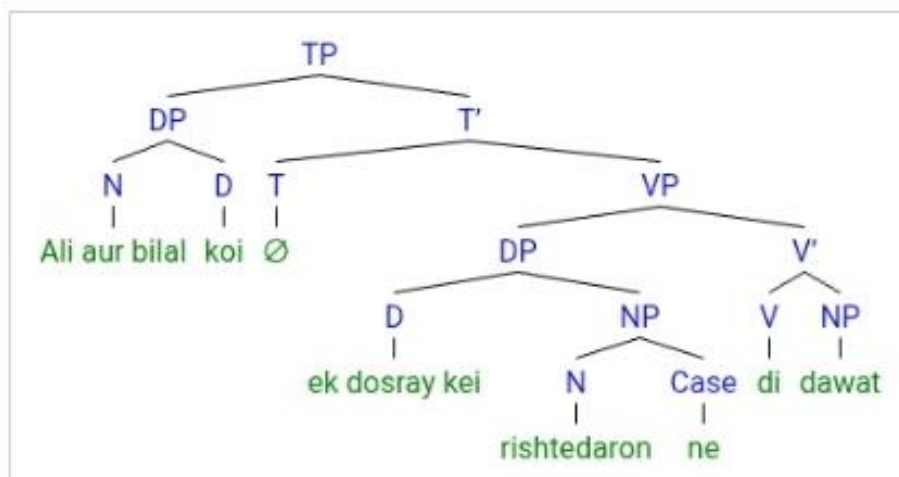
c.[Ali aur Bilal]-koi [ek dosray-kei rishtedaron]-ne dawat di.

Ali and Bilal-ACC each. other-GEN relatives-ERG invitation gave

‘Each other’s relatives invited Ali and Bilal.’

[TP [DP [N Ali aur bilal] [D koi]] [T' [T \emptyset] [VP [DP [D ek dosray kei] [NP [N rishtedaron] [Case ne]]] [V' [V di] [NP dawat]]]]]

Figure 4.11. *DO S V leftward scrambling*



Again, only the leftward-scrambled configuration in (2c) yields a grammatical reciprocal interpretation. The unacceptability of reciprocal binding in (2a) and (2b) confirms that the linear position of the binder relative to the bindee is interpretively significant.

Further support for this generalization comes from cases involving multiple rightward scrambled arguments. In such configurations, the binding potential of co-arguments is strictly governed by linear precedence:

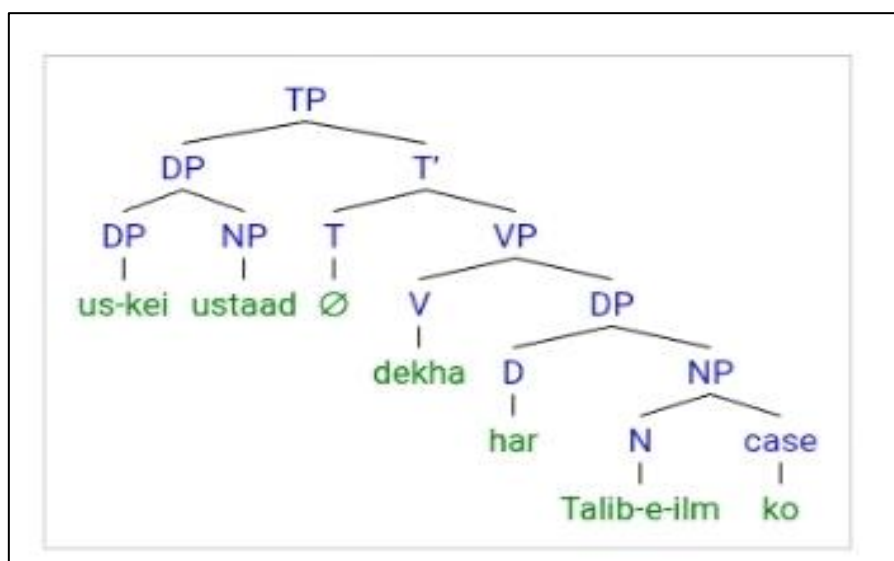
(2) a. *Dekha [us-kei ustaad]-ne [har talib-e-ilm]-koi.*

saw he-GEN teacher-ERG every student-ACC

‘*His teacher saw every student.’

[TP[DP [DP us-kei] [NP ustaad]] [T' [T Ø] [VP [V dekha] [DP [D har] [NP [N Talib-e-ilm] [case ko]]]]]

Figure 4.12. *V S DO Rightward scrambling*



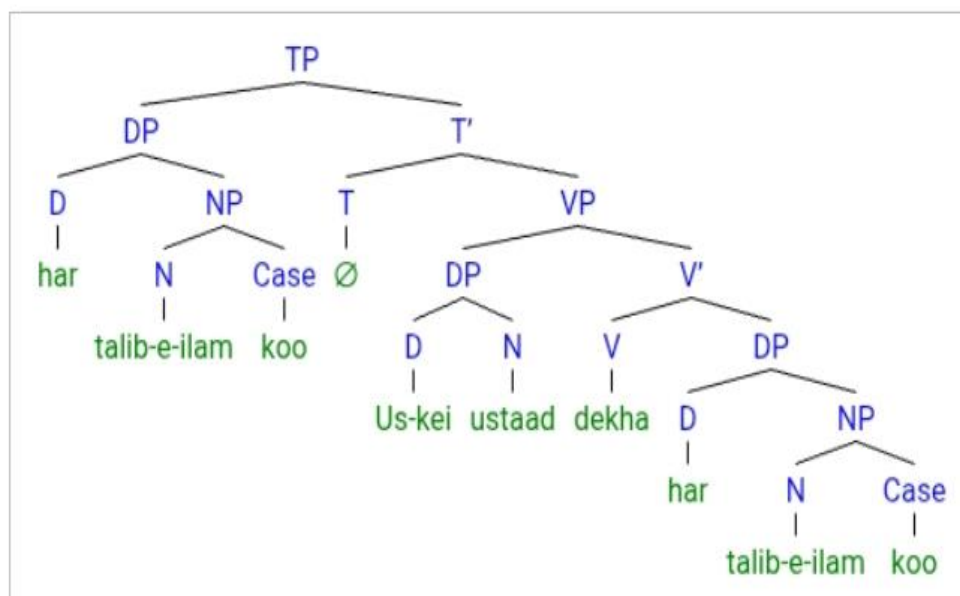
b. *Dekha [har talib-e-ilm]-koi [us-kei ustaad]-ne*

saw every student-ACC he-GEN teacher-ERG

‘Hisi teacher saw every student.’

[TP[DP[D har][NP[N talib-e-ilam][Case koo]]][T'[T Ø][VP[DP[D Us-kei][N ustaad]]][V'[V dekha][DP[D har][NP[N talib-e-ilam][Case koo]]]

Figure 4.13. *V DO S leftward scrambling*



In (3b), the initial Corel Argument successfully connects to the second one, but in (3a), the other arrangement gives a WCO violation. This demonstrates that rightward scrambling does not generate symmetrical environments for binding; instead, it upholds a dominant order among post-verbal constituents that limits their interpretive relations.

In light of WCO, reciprocal binding, and multiple scrambling, we suggest this maxim concerning scope and binding in Urdu:

4.2 Linear Order Generalization for Urdu

In Urdu, the presence of variable binding and pronominal coreference features between co-arguments is dependent on the surface linear order. If XP_1 and XP_2 are co-arguments and XP_1 precedes XP_2 in the linear string, then in Logical Form XP_1 c-commands and has scope over XP_2 .

The claim that contrastive grammar has not been leant on too heavily because there exists scrambling in the language is still valid. Urdu uses constructs that are dependent on order of elements in a quasi-greedy, surface-sensitive manner. As a result, obstruction is produced when scope and binding are configured significantly in a manner that scope is still available but exceed acceptable bounds. These results are patterned on typological strategies to handle scrambling in South Asian languages showing that, unlike many languages, Urdu is not interpretively vacuous and lacks delimiters bound to directionality and shape restrictions.

4.3 Restricted Scope of Rightward-Scrambled Wh-Expressions in Urdu

Although rightward scrambling of wh-movements in Urdu seems to be syntactically harmless in several cases, it does impose a severe restriction on the scope of wh-expressions. Let's examine the differences in objectives with respect to the wh-phrases within the object's slot.

(1) a. S Owh V Aux

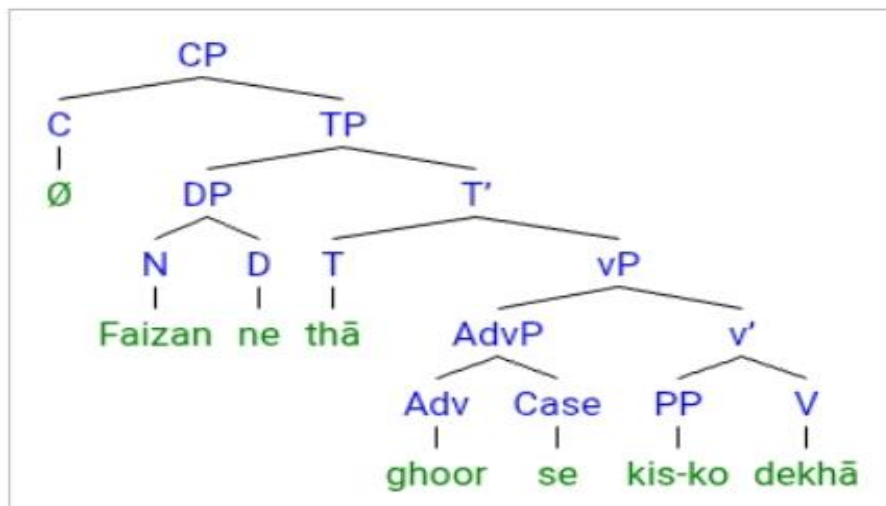
Faizan-ne ghoor se kis-ko dekhā thā?

Faizan-ERG attentively who-ACC see-PFV be.PST

'Who did Faizan look at attentively?'

[CP[C Ø] [TP[DP[N Faizan][D ne]][T' [T thā] [vP [AdvP[Adv ghoor][Case se]][v' [PP kis-ko][V dekhā]]]

Figure 4.14. a. *S Owh V Aux*

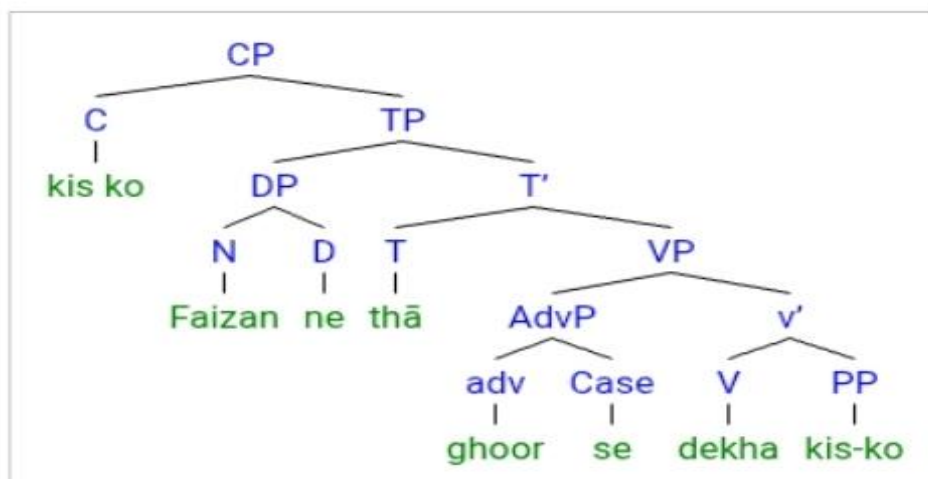


b. Owh S V Aux

Kis-ko Faizan-ne ghoor se dekhā thā?
 who-ACC Faizan-ERG attentively see-PFV be.PST
 'Who did Faizan look at attentively?'

[CP[C kis ko][TP[DP[N Faizan][D ne]][T'[T thā] [VP [AdvP[adv ghoor][Case se]][v'[V dekha][PP kis-ko]]]

Figure 4.15. *Owh S V Aux*

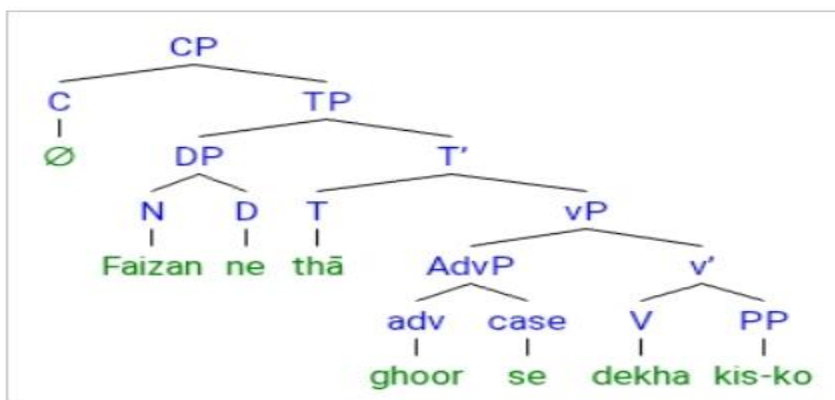


b. S V Aux Owh

Faizan-ne ghoor se dekhā thā kis-ko?
 Faizan-ERG attentively see-PFV be.PST who-ACC
 #'Who did Faizan look at attentively?' (✓ only as echo)

[CP [C Ø][TP[DP Faizan-ne][T' [T thā][vP[AdvP[adv ghoor][Case se]][V'[V dekha] [PP kis-ko]]]

Figure 4.16. *Owh S V Aux*



In (1a) and (1b), the wh-expression obtains standard interrogative scope, either in-situ or leftward scrambled. But in (1c), where the wh-expression is scrambled to the post verbal position beyond both the verb and auxiliary, it does not yield a true information seeking question. It is rather interpretable only as an echo question. This indicates that rightward scrambling across both the verb and auxiliary does not enable conventional wh-scope licensing, and thus standard wh-scope contextualization is disrupted in this scenario. Importantly, the unacceptability in (1c) is not simply due to rightward scrambling. It is due to the combination of rightward scrambling with auxiliary placement. This is evident from the minimally contrasting (2), where the wh-phrase is post-verbal but precedes the auxiliary:

(2) S V Owh Aux

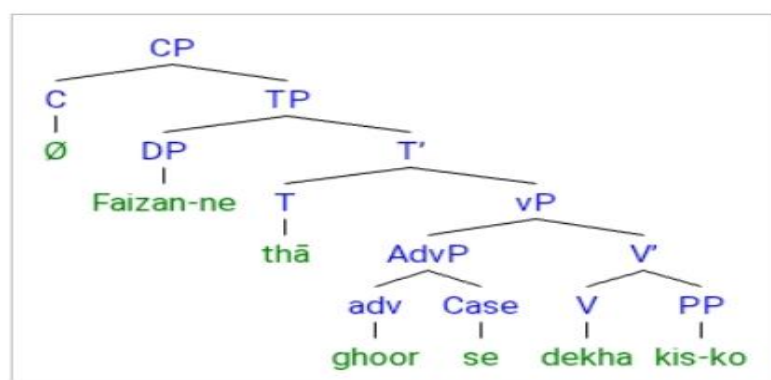
Faizan-ne ghoor se dekhā kis-ko thā?

Faizan-ERG attentively see-PFV who-ACC be.PST

‘Who had Faizan looked at attentively?’

[CP [C Ø][TP[DP Faizan-ne][T' [T thā][vP[AdvP[adv ghoor][Case se]][V'[V dekha] [PP kis-ko]]]]]

Figure 4. 17. S V Owh Aux



Unlike (1c), the configuration in (2) allows for the normal interpretation of wh-questions, suggesting that scrambling over the verb only does not block scope interpretation. It is the cross over of both the verb and. The auxiliary is the one that causes interpretive degradation. In Urdu, however, the position of the auxiliary with respect to the wh-phrase triggers scope licensing makes bounding the scope possible within the Permissible regions for control External clauses The wh-phrase enables the sentence. This restriction corresponds only to wh-dependencies,

and does not apply to other scope related matters like variable binding or pronominal coreference (as shown in §1.2), which remain unaffected regardless of the *wh*-phrase's position in relation to the auxiliary.

These facts suggest that the *wh*-expressions in clauses obey a locality constraint enforced by the grammar of Urdu, which prohibits rightward-scrambled *wh*-phrases from taking scope over clausal domains they spatially appear to have vacated. This observation corresponds to the more general Urdu observations that the scope of dislocated *wh*-clauses is controlled by the placement of the subordinate clause-auxiliaries-surface syntax. Embedded clauses support this further. For example, in finite complement clauses, *wh*-expressions have to move mandatorily to the matrix domain to achieve wide scope. When embedded *wh*-elements remain in situ within a post-verbal finite clause, they are interpreted with narrow (embedded) scope only:

(3) a. S V Aux [FiniteCP ... *wh* ...]

Ali jantā thā [ke kaun āyā thā]?

Ali.M know.IMPF be.PST that who.M come.PFV be.PST

'Ali knew who had come.' (✓ embedded reading only)

b. *whi* S V Aux [FiniteCP ... *ti* ...]

Kaun Ali jantā thā [ke *ti* āyā thā]?

who.M Ali.M know.IMPF be.PST that come.PFV be.PST

'Who did Ali know had come?' (✓ matrix reading)

Similarly, for nonfinite complements—typically gerundive constructions in Urdu—the scope of embedded *wh*-elements is sensitive to preverbal versus postverbal placement. A preverbal nonfinite clause permits matrix scope for its *wh*-elements:

(4) a. S [Nonfinite ... *wh* ...] V Aux

Asma-ne [kis-ko milnā] chāhā thā?

Asma-ERG who-ACC meet.INF want.PFV be.PST

'Who did Asma want to meet?'

b. S V Aux [Nonfinite ... *wh* ...]

?Asma-ne chāhā thā [kis-ko milnā]?

Asma-ERG want.PFV be.PST who-ACC meet.INF

'Who did Asma want to meet?' (✓ marginally acceptable or echo)

Thus, the bounding interpretive constraints due to rightward scrambling in Urdu are not confined to matrix clauses, but extend to embedded structures too. In all these instances, the surface order resembles the figurative scope quite closely, which further reinforces the conclusion that rightward scrambling does limit the interpretability of *wh*-elements when they traverse clause boundary markers like auxiliaries.

4.4 A Rightward Remnant Movement Approach

Remnant-VP Movement in Urdu

With the primary empirical data about scope and word order in Urdu having been provided, we now proceed to a theoretical explanation that does not incorporate anti-symmetry or assume a base-generated SVO ordering. Though some earlier attempts have used an anti-symmetric account for Indo-Aryan languages, this is by no means necessary to understand rightward scrambling behavior in Urdu. We instead suggest that the appropriate orders result from remnant VP movement to the right after the main exhaust of the verb phrase has already been moved.

We posit the widely accepted hypothesis that places Urdu among the SOV languages; that is, the base position of the arguments occurs prior to the verb, and constituents can be scrambled leftward into higher specifier and adjunct positions. We further maintain that the heads of

verbal phrases may optionally lift to an Aspect Phrase (AspP) above vP as has been argued due to word order relations involving negation and auxiliary aspectual verbs.

The main focus of this paper is that the rightward argument scrambling in Urdu is most reasonably captured as right-moving verbal projections (VPs), in this case remnant VPs where the verb has been extracted by head movement. This accounts for the dependency of word order and scope in a more coherent way.

Consider the following pair of examples illustrating the canonical and scrambled orders in Urdu:

(1) a. Us-ne [hamaari baatein] dhyān se sunnī thī.

he-ERG our talks attentively hear. PFV.F be. PST.F

‘He had listened attentively to our conversation.’

b. [Us-ne thī] [[hamaari baatein] dhyān se sunnī].

he-ERG be. PST.F our talks attentively hear. PFV.F

‘He had listened attentively to our conversation.’

In (1a) the object appears before the verb and auxiliary, indicating the default SOV word order.

In (1b), object and verb are placed after the verb as a result of rightward VP movement, leaving the auxiliary in its higher functional projection. This kind of movement accounts for the observed string without reference to post-syntactic reordering or linearization restrictions. Think about cases with ditransitive constructions that contain a direct and indirect object. Such structures are easily explained by employing a nested VP structure through the remnant movement approach:

(2) a. Canonical S IO DO V Aux

Zaid-ne Amina-ko kitaab dī thī.

Zaid-ERG Amina-DAT book.F give.PFV.F be.PST.F

‘Zaid had given a book to Amina.’

b Rightward remnant VP2 movement:

[Zaid-ne Amina-ko dī thī] [kitaab].

Zaid-ERG Amina-DAT give.PFV.F be.PST.F book.F

‘Zaid had given a book to Amina.’

2c Full remnant VP1 movement:

[Zaid-ne dī thī] [Amina-ko kitaab].

Zaid-ERG give.PFV.F be.PST.F Amina-DAT book.F

‘Zaid had given a book to Amina.’

In (2b), only the inner VP2 (with the direct object) has been operation rightward, resulting in an auxiliary pre-posed structure where the direct object is post-auxiliary. In (2c), entire VP1 (containing both the IO and DO) undergoes rightward movement, leaving the auxiliary and subject in their original positions. These changes show different levels of remnant movement and account for the flexibility in Urdu word order. More importantly, the rightward movement is restricted to verbal projections. For some proposed surface orders, this is problematic. For instance, the order [S DO V Aux IO] not directly stem from (2a), because there is not a VP containing the indirect object while excluding the direct object. In order to explain this, we have to start with the assumption that there is a leftward scrambling of the DO over the IO within the VP before any remnant movement occurs:

(5) a. Pre-scrambling:

Zaid-ne [kitaab] [Amina-ko] dī thī.

Zaid-ERG book. F Amina-DAT give. PFV.F be. PST.F

‘Zaid had given a book to Amina.’

b. Remnant VP movement after scrambling:

[Zaid-ne kitaab dī thī] [Amina-ko].

Zaid-ERG book.F give. PFV.F be. PST.F Amina-DAT
'Zaid had given a book to Amina.'

The analysis of the remnant movement predicts accurately that these surface orders can only result functionally from previous scrambling actions which are independently attested in Urdu. This combination of scrambling and remnant VP movement enables us to account for a number of attested word orders without having to appeal to claimed base-generated variants or antisymmetric structures. Also, this explains certain structural scope rigidity of post-verbal elements. As argued in §1.3, whatever-expressions that are scrambled past auxiliary verbs to the right lose interrogative scope. With the remnant movement explanation, this is justified by the fact that scope taking elements must c-command the scope domain at LF. An embedded wh-phrase in a remnant VP subjected to rightward movement cannot reconstruct above the auxiliary, so loses the ability to command the appropriate scope site.

4.5 Deriving Scope Relationships in Urdu Syntax

One of the most important aspects of Urdu syntax is that the different orders of arguments seem to correspond to scope interpretation hierarchy. This portion explains such correlations using remnant VP movement along with its interpretational consequences, specifically reconstruction. We adopt, as in Huang (1993), the constitutionalist view that remnant movements are always subject to obligatory reconstruction (Soh, 1998). The movement in Urdu occurs because the VP internal constituents have been displaced via scrambling precede the remnant VP rightward movement which is headless because of head movement. Assume constituents order base structure in Urdu the subject S indirect object IO and direct object DO comes after S IO DO V Aux give the verb phrase order in which S IO DO precedes verb phrase in which DO is in VP2 and IO is contained in VP1 simultaneously. In rightward remnant movement the order on the surface changes, but the meaning hierarchy still correlates with base positions as shown in 18b.

(18) a. [XP1 [VP1 XP2 [VP2 XP3 t_j]] V_j Aux]

b. Scope: XP1 > XP2 > XP3

In this arrangement, XP1 normally the subject c-commands XP2 the IO, who c-commands also XP3 the DO, setting up their relative scope relations. When a remnant VP consisting of the traces of the scrambled constituents is rightward shifted, reconstruction achieves the original c-command relation at LF. This is shown schematically in the derivations below which demonstrate the different surface orders while preserving scope relations via reconstruction:

(19) a. Surface order: S V Aux IO DO

i. Structure: [[XP1 t_i V_j Aux] [VP1 XP2 [VP2 XP3 t_j]]]_i

ii. After reconstruction: [XP1 [VP1 XP2 [VP2 XP3 t_j]] V_j Aux]

b. Surface order: S DO V Aux IO

i. Structure: [[XP1 [XP3 t_i] V_j Aux] [VP1 XP2 [VP2 t₃ t_j]]]_i

ii. After reconstruction: [XP1 [XP3 [VP1 XP2 [VP2 t₃ t_j]]] V_j Aux]

In (19a), the subject also takes wide scope with respect to both the IO and DO because the canonical base structure is maintained and the rightward-moved VP is reconstructed to one where it undergoes subject scoping. In (19b), where the DO appears pre-verbally and the IO post-verbally, reconstruction of order takes care of the fact that the DO appears to scope over the IO. This must be the case because the DO has scrambled leftward before remnant movement. This explains why in Urdu, postverbal arguments cannot scope over preverbal arguments unless they have undergone such scrambling beforehand. Most importantly, this approach strengthens a phase view concerning the argument's position in the hierarchy of the displacement in Urdu, where scope remnant VP movement is found to combine with motivated scrambles and head V to Asp movements.

4.6 Deriving the Wh-Scope Effects in Urdu

This part examines the proposed remnant movement analysis of rightward scrambling in Urdu and how it interfaces with the interpretation of *wh*-expressions. Following customary views found in generative syntax, we assume that *wh*-phrases in Urdu are required to be in the *c*-command domain vis-a-vis a complementizer $C^0[+wh]$ and, therefore, are able to interpret the question as an interrogative. Upward movement to this position could be covert, occurring at the Logical Form stage, as seen in *wh*-in-situ languages. It has been noted in the literature that remnant VP movement creates scope islands which capture constituents inside the remnant and, hence, forbid them from taking scope outside (Lechner, 2003; Sauerland, 1998). Applying this to Urdu, we account for the ungrammaticality of questions in which the *wh*-element is embedded in a postverbal VP remnant, making the sentence uninterpretable due to the inability of the *wh*-phrase to move covertly to $C^0[+wh]$. Consider the following example:

(21) a. *[VP *kis-ko dekhaa*] Aisha-ne thaa?

who-ACC see.PFV Aisha-ERG be.PST

Intended: 'Who had Aisha seen?'

In (21a), *kis-ko* ('whom') is located within a VP which has been rightward moved. This remnant creates a scope island which blocks covert *wh* movement to the interrogative complementizer. The outcome is an incomprehensible structure as a *wh* question. By contrast, in an acceptable instance where the *wh* expression is placed between the verb and the auxiliary, there is no escaping the remnant *wh* extraction:

(22) Aisha-ne ghawr-se dekhaa kis-ko thaa?

Aisha-ERG carefully see.PFV who-ACC be.PST

'Who had Aisha looked at carefully?'

In (22), *kis-ko* is not inside the VP remnant, and the question is grammatical. We propose the following derivation, which involves **leftward scrambling of the object**, followed by **topicalization of the participial verb**, and ultimately **covert *wh*-movement**:

(23) a. Base structure:

[S [AdvP [VP D_{Owh} V]] Aux]

b. Scrambling of DO:

[S [AdvP D_{Owh_i} [VP t_i V]] Aux]

c. Fronting of V:

[S [AdvP [V_j D_{Owh_i} [t_j t_i]]] Aux]

d. Covert *wh*-movement to $C^0[+wh]$:

[D_{Owh_i} [S [AdvP [V_j t_i [t_j]]] Aux]]

Urdu syntax is articulated above as a set of independently well-formed structures corresponding to each stage. The *wh*-phrase is moved out prior to the remnant VP undergoing rightward movement, so it can still be interpreted under the $[+wh]$ operator. Interrogative scope and well-formedness is retained and thus can be accounted for. In other words, the scope effects of *wh*-phrases in Urdu provide evidence in support of rightward scrambling being remnant VP movement instead of the movement of single arguments. This account is able to capture both the scope island restrictions and the boundedness of questions that allow remnant movement which is carefully planned through preemptive scrambling. In the following sections, we set out contrasts with antisymmetric approaches like Mahajan (1997) or Simpson & Bhattacharya (2003) and show how these approaches do not adequately explain the greater scope-related constraints specific to Urdu.

4.7 Antisymmetric Accounts Linearity and Scope

Integrating scope with the previously discussed linear order within Urdu, we now focus on the antisymmetric approach by Mahajan (1997). Mahajan suggests that the '[S IO DO V Aux]' order in Hindi-Urdu is derived from a head-initial structure- which is base generated- through a series of leftward movements. Rightward-scrambled constituents also posited to arise from remnant movement where all but the rightmost element move to higher positions. As we shall show, such reasoning does not explain the scope relations in Urdu that we will present."

Consider the derivation of the [S IO V Aux DO] order, simplified here for clarity:

- a. [IO [V DO]]
- b. [IOi [DOj [ti V tj]]]
- c. [S [IOi [DOj [ti V tj]]]]
- d. [Aux [S [IO [DO V]]]]
- e. [Si [Aux [ti [IO [DO V]]]]]
- f. [DOj [Si [Aux [ti [IO [tj V]]]]]]]
- g. [Auxk [DOj [Si [tk [ti [IO [tj V]]]]]]]]]
- h. [[Si [tk [ti [IO [tj V]]]]]]! [Auxk [DOj tl]]]

In (h), the indirect object (IO) does not c-command the direct object (DO) as they are arranged in a particular order, which would otherwise suggest that in the hierarchy of command, an IO commands a DO. To resolve this, one has to blend c-command with the notion that XPs in a fronted remnant can c-command outside the phrase they are contained within. This is both suspect and literally doesn't exist in practice. c-command domains are very wide in scope do not exist for constituents embedded in fronted phrases in Urdu. Take a look at the following examples involving control and binding that are opposed to the assumption of antisymmetry, meaning that, paired with scope, one invokes linear order:

(22a) [Zayd-ka us-sei chupke milna]j [Amina-kii maa~-ko tj bilkul pasand nahĩ: hai.

Zayd-GEN her-INST secretly meeting Amina-GEN.F mother-DAT at.all like NEG be.PRS
'Amina'si mother does not at all like Zayd meeting with heri in secret.'

(22b) [Usei da~:tne]-koj [Amina-kii maa~-ne [us-kiii ustaad]-se ti kehā.

him.DAT scold-INF.OBL-ACC Amina-GEN.F mother-ERG he-GEN.F teacher-INST
say.PFV

'Amina'si mother told hisi teacher to scold himi.'

These examples demonstrate that scope and binding are not reliably determined by linear order. Further evidence from result clauses in Urdu reinforces this point:

(23) [Itne zyada logo~-ne usei tohfe diye [ke Zayd maalāmāl ho gayā].

so. many people-ERG him.DAT gifts give. PFV that Zayd rich become go. PFV

'So many people gave himi gifts that Zaydi became rich.'

In such cases, even though *usei* precedes the clause containing *Zayd*, there is no c-command or scope relation between the two. Thus, appealing to reconstruction of the remnant does not salvage the antisymmetric approach it in fact leads to incorrect scope predictions, such as:

(24) DO > S > IO

(actual: S > IO > DO)

This misalignment between derived structure and observed interpretation undermines the antisymmetric account. Moreover, the antisymmetric models of Simpson and Bhattacharya (2003), which share Mahajan's assumptions, also inherit this shortcoming.

4.8 The Restriction on Wh-Scope

A further challenge for antisymmetric theories of Urdu is the restriction on matrix-scope interpretations of rightward-scrambled wh-expressions. While Mahajan (1997) and Simpson and Bhattacharya (2003) both note this restriction in passing, neither provides an adequate analysis. Simpson and Bhattacharya, focusing on Bangla, suggest that overt wh-movement is

masked by movement to a Topic position above CP. Their derivation for [S Owh V] proceeds as:

- a. Base: [S [V Owh]]
- b. Case: [S [Owh [V]]]
- c. Wh-movement: [Owh-i [C [S [ti V]]]]
- d. Topic-movement: [S [Top [Owh-i [C [tj [ti V]]]]]]

This analysis predicts that [S V DOwh] is ungrammatical unless the wh-element moves to Spec,CP. However, it allows the possibility of a well-formed [S V DOwh] order if the verb moves to a Topic head. To block such derivations, they stipulate that the finite verb cannot move past C0.

This faces difficulties with Urdu sentences like:

(26) Amina-ne ghoor-se dekha kis-ko thā?

Amina-ERG carefully see.PFV who-ACC be.PST

‘Who had Amina looked at carefully?’

In (26), *kis-ko* comes after the participial verb and before *thā*, yet maintains a matrix-scope interpretation. This indicates that the verb can move to an upper position more than what is deemed necessary, hinting towards the need for more specific and detailed partitions of the verbal components—an unwelcome complication.

In contrast, the approach done using remnant movement like the one we proposed explain the restriction effortlessly: wh-elements left behind in rightward-scrambled locations are still captured in non-scope-taking phrases. Therefore, the reason they don’t obtain matrix scope is not because of bad movement, but because the position that the phrasing is in—scope-embedding.

4.9 Wh-Expressions Inside Finite Complements in Urdu

This portion assesses the movement of wh-expressions in finite complement clauses in Urdu with particular attention paid to the effects of remnant movement on the interpretation. Urdu is known for its consistent SOV word order. In cases where a Finite CP is pronounced post-verbally (CP-in-situ) as in [S V [CP ... wh ...]], the wh-pronoun inside the CP cannot escape matrix scope. More specifically, such configurations fail to allow matrix wh-questions.

Most likely, the lack of overt wh-movement within the wh-phrase leads to a suspension of matrix scope. Because the wh-phrase is not overtly moved, it can only be construed in situ where the CP is embedded. The CP must therefore be captured in an upper position like Spec, CP, necessitating it to be inverted to serve matrix scope (Rizzi, 1993). In Urdu, we suggest that such postverbal CPs do not occupy a base-generated position but instead are obtained through rightward extraposition stranding them within a remnant verbal projection. As argued earlier, verbal remnants in Urdu are scope islands: once constituents get embedded in a rightward-turned remnant, they lose accessibility to dominant control seeking operations. Hence we cannot have a wh-element that is captured within a CP which itself sits inside a remnant, covertly wh-movement to the higher level scope to take matrix position scope as surface syntax would suggest. This scope opacity persists even after reconstruction of the remnant at LF since reconstruction does not reestablish the critical c-command relations that enabled crossing of island borders.

This analysis implies that *wh*-in-situ inside an extraposed CP is syntactically licensed but semantically inert with respect to matrix scope. For example:

[S V [CP ke Aisha kis se milī]]

‘Who did Aisha meet?’ (Intended matrix reading unavailable)

This type of embedded reading is the only one available and it incorporates matrix clauses with declarative complements. The wh-phrase *kis se* is kept inside the scope domain of the CP and because the postverbal clause is a remnant, it cannot scope outside of CP. Earlier accounts, like

Dayal 1996, suggest that there are extraposed CPs that may get back to the basic position which is under the *wh*-movement that is allowed afterwards. In reconstructing these positions, however, other discrepancies, both empirical and theoretical, arise. With my approach, the lack of visibility of extraposed CPs comes from their nature and structure as remnants, rather than from movement domain stipulations. Notably, these constraints on *wh*-scope in Urdu are more complex than in the language's simple scenarios. A comprehensive treatment must address:

- **Partial *wh*-movement / scope marking**, where a matrix *wh*-expression appears to correlate with a lower clause *wh*-phrase:
[*S wh V [CP ... wh ...]*]
- **Pair-list readings**, involving multiple *wh*-phrases in nested clauses:
[*CP ... wh ... V [CP ... wh ... wh ...]*]
- **Pronominal resumption with extraposed CPs**, where an overt matrix pronoun is coreferential with an embedded clause containing a *wh*-phrase:
[*S pronouni V [CP ... wh ...]i*]

These configurations allow empirical assessment of the predictions associated with the remnant-based analysis. If we assume that the rightward movement of CPs in Urdu invariably traps embedded *wh*-elements, we should predict that the scope possibilities in these constructions are equally constrained unless particular scope-marking techniques are used. This way, the remnant movement approach not only explains why there are no matrix *wh*-extrapolated readings in *wh*-extrapolated CP's but provides a "Urdu-way" explanation in terms of movement typology. Further research is needed on how the typological features of Urdu, especially its verb-final sentence structure, interact with *wh*-scope and extraposition across other languages.

Conclusion

This investigation has greatly advanced our knowledge regarding the structural flexibility of rightward scrambling and its scope related effects in Urdu by providing compelling empirical evidence on its syntactic and interpretive features. We have shown through a naturalistic study of speech that Urdu allows scrambling of various elements, such as noun phrases (NPs), prepositional phrases (PPs), and adjuncts, and still maintains grammaticality and acceptability. Importantly, the data suggests that scrambling is not an isolated or merely stylistic choice, but a fundamental syntactic operation furrowed within the grammar of Urdu. Our results demonstrate that scrambling in Urdu is not devoid of meaning; on the contrary, it is crucial in defining the scope of some *wh*-expressions. In particular, rightward scrambling of arguments causes scope freezing, where a *wh*- element is confined to a narrowly defined lower sense range; however, scrambling of adjuncts may allow wider scope interpretation depending on context. This relationship between scrambling and scope reinforces the argument that the language features tightly knit syntax and semantics, where order of utterance is constrained by the need to control the interpretation.

The research also demonstrates how information structure is affected, revealing that rightward displacement typically marks backgrounded elements of lower discourse prominence, while preverbal position signals focus or topicality. These considerations are in keeping with recent cartographic and minimalist frameworks that argue scrambling arises from a feature-driven movement within the left or right edge of the clause. In any case, this work integrates Urdu into a more comprehensive cross-linguistic analysis concerning the rights and lefts bounds of scrambling and as such enhances typological and theoretical works on Urdu scrambling. It also validates that rightward scrambling is not simply extraposition, but rather a syntactically valid operation within discourse, scope, and interface conditions. Comparative work on these questions is particularly needed, for example, with Japanese or Persian languages, to deepen understandings of parametric differences in scrambling and its interpretation.

References

- Ali, A., & Malik, N. A. (2023). Split tense projection in Urdu: An illusion. *Pakistan Journal of Language Studies*, 7(1), 16-31.
- Bhatt, R., & Dayal, V. (2007). Rightward scrambling as rightward remnant movement. *Linguistic inquiry*, 38(2), 287-301. <https://doi.org/10.1162/ling.2007.38.2.287>
- Chomsky, N. (1995). Language and nature. *Mind*, 104(413), 1-61. <https://doi.org/10.1093/mind/104.413.1>
- Chomsky, N. (2004). Beyond explanatory adequacy. *Structures and beyond: The cartography of syntactic structures*, 3, 104-131. <https://doi.org/10.1093/oso/9780195171976.003.0004>
- Chomsky, N., Belletti, A., & Rizzi, L. (2002). An interview on minimalism. *N. Chomsky, On Nature and Language*, 92-161.
- Dayal, V. S. (1994). Binding facts in Hindi and the scrambling phenomenon. Theoretical perspectives on word order in South Asian languages, ed. by Miriam Butt, Tracy Holloway King & Gillian Ramchand, 237-61. In: Stanford: CSLI.
- Haider, H. (2021). A null theory of scrambling. *Zeitschrift für Sprachwissenschaft*, 39(3), 375-405. <https://doi.org/10.1515/zfs-2020-2019>
- Kareem, R. A., & Yaseen, A. S. (2023). A Minimalist Analysis of Scrambling in Central Kurdish. *Journal of University of Human Development*, 9(2), 106-123. <https://doi.org/10.21928/juhd.v9n2y2023.pp106-123>
- Kidwai, A. (2000). *XP-adjunction in Universal Grammar: Scrambling and binding in Hindi-Urdu*. Oxford University Press, USA.
- Lechner, W. (1998). Two kinds of reconstruction. *Studia Linguistica*, 52(3), 276-310.
- Lechner, W. (2003). Phrase structure paradoxes, movement and ellipsis. In *The interfaces: Deriving and interpreting omitted structures* (pp. 177-203). John Benjamins Publishing Company.
- Mahajan, A. K. (1990). *The A/A-bar distinction and movement theory* Massachusetts Institute of Technology].
- Mahajan, A. K. (2011). Rightward scrambling. In *Rightward movement* (pp. 185-214). John Benjamins Publishing Company.
- Raza, G., Ahmed, T., Butt, M., & King, T. H. (2011). Argument Scrambling within Urdu NPs. *Proceedings of LFG11*, 461, 1-13.
- Rizzi, L. (1993). Some notes on linguistic theory and language development: The case of root infinitives. *Language acquisition*, 3(4), 371-393.
- Saito, M. (1989). Scrambling as Semantically. *Alternative conceptions of phrase structure*, 182.
- Sauerland, U. (1998). *The meaning of chains*. MIT Press Cambridge, MA.
- Simpson, A., & Bhattacharya, T. (2003). Obligatory overt wh-movement in a wh-in-situ language. *Linguistic inquiry*, 34(1), 127-142.
- Soh, H. L. (1998). *Object scrambling in Chinese* Massachusetts Institute of Technology].