Licensing long-distance *wh*-in-situ in Malayalam*

Athulya Aravind, MIT

May 30, 2016

Abstract

It is generally thought that *wh*-in-situ, like overt movement, is potentially unbounded. At the same time, certain languages have been argued to disallow long-distance *wh*-in-situ. This paper argues that even in languages that show apparent clause-boundedness effects, *wh*-in-situ, like *wh*-movement, can in principle cross an arbitrary number of clauses. Failure to license a *wh*-phrase across a clause boundary, when it occurs, can be shown to result from the interaction between *wh*-agreement and independent operations affecting embedded clauses. Evidence will be drawn primarily from Malayalam (Dravidian), which has been argued to disallow long-distance *wh*-in-situ with finite embedded clauses. I will show that the relevant factor for *wh*-licensing is not finiteness, but Ā-movement of embedded clauses, an operation that is common with finite CPs. The core of the problem lies in the fact that interrogative C is a generalized [Ā]-probe that can interact with a number of featurally more specific goals, including the [Ā]-features on the head of the moving clause. It will be shown that this approach can account for a number of facts about Malayalam *wh*-question formation, including selective transparency of certain finite clauses for long-distance *wh*-licensing.

1 Introduction

Languages can be broadly categorized into two types with respect to *wh*-question formation. There are those that overtly move one or all *wh*-phrases to the front of the clause and those that leave them in-situ. In principle, both *wh*-strategies can be unbounded. Consider the long-distance *wh*-questions in English (1a) and Japanese (1b). Though the *wh*-expressions differ in surface position, the interpretation is the same — the *wh*-phrase originating in the embedded clause takes scope over the entire sentence.

(1) a. **What** did John say [Mary bought ___]?

b. Hideki-ga [Kyoko-ga **nani-o** kat-ta to] it-ta no
Hideki-NOM [Kyoko-NOM what-ACC buy-PAST] say-PAST Q
'What did Hideki say Kyoko bought?'

*Thanks to Adam Albright, Kenyon Branan, Danny Fox, Sabine Iatridou, Norvin Richards, Roger Schwarzchild, Coppe van Urk, Michelle Yuan, as well as audiences at MIT's Syntax Square, MIT's Ling-Lunch and NELS 46 for comments and discussion. I am especially grateful to David Pesetsky for his generous feedback. All errors are my own.

The basic generalization is that this type of parametric variation with respect to *wh*-movement does not create deep differences in the kind of questions speakers of the different languages can form. Island environments aside, both movement and in-situ strategies are in principle unbounded and allow speakers to ask long-distance questions of arbitrary length.

However, it has been observed that in a number of languages, *wh*-in-situ appears to be subject to locality restrictions that are unique to this question-formation strategy, an observation which, if true, raises doubts about the generalization above. In languages like Hindi (Dayal, 1996; Mahajan, 1990), Bangla (Bayer, 1997; Simpson and Bhattacharya, 2003) and Iraqi Arabic (Ouhalla, 1996; Simpson, 2000; Wahba, 1991), in-situ *wh*-expressions inside finite complements seem restricted to clause-bound scope. On the basis of the apparent strict locality of *wh*-in-situ in these languages, a number of researchers have suggested that languages can *parametrically vary* with respect to the locality conditions on *wh*-licensing (Nakamura, 1998; Ouhalla, 1996; Simpson, 2000).

Simpson (2000), for example, argues that the domain for *wh*-feature checking may vary across languages, proposing the following three-way classification (2).

(2) *Wh-licensing domains* (Simpson 2000, p. 109, ex. (58))

a.	**Type A:** Romanian, Bulgarian	Spec, CP
b.	**Type B:** Iraqi Arabic, Hindi	Immediate tense domain of the +Q C^{o}
c.	**Type C:** English	The sentence

Of particular relevance for our purposes is the distinction between Type B and Type C languages. On Simpson's proposal, Type B languages like Iraqi Arabic and Hindi disallow long-distance *wh*-in-situ because a *wh*-phrase that remains inside a finite embedded clause cannot be accessed for *wh*-feature-checking, since the interrogative C cannot see past its immediately local tense domain. Long-distance *wh*-movement in these languages, then, is in-part motivated by the need to get the embedded *wh*-phrase into an accessible *wh*-licensing domain.

A similar proposal is put forth in Ouhalla (1996), who attributes the locus of parametric variation to properties of the *wh*-phrases themselves. More specifically, he argues that *wh*-phrases in languages like Iraqi Arabic and Hindi are anaphors that cannot be bound across a finite clause.

I will refer to those languages claimed to have clause-boundedness restrictions on *wh*-in-situ (whether it be restrictions on *wh*-licensing domains or binding requirements on the *wh*-phrase itself) as **restricted-scope** languages, and languages like English, Japanese, Mandarin and Korean as **unrestricted-scope** languages. The proposals outlined above, though presenting straightforward explanations for why *wh*-in-situ may be more constrained in restricted scope languages than *wh*-movement, presuppose that finiteness is the relevant factor in these asymmetries. One of the main claims in this paper is that finiteness is not directly correlated with clause-boundedness effects in restricted scope languages. I will argue instead that *wh*-in-situ is uniformly non-clause-bound, and

that apparent clause-boundedness effects are epiphenominal. I draw primarily on evidence from Malayalam (Dravidian), a language that has been argued in the past to have only clause-bound *wh*-in-situ (Hany Babu, 1997; Madhavan, 2013; Srikumar, 1992, 2007). I will show that the restricted scope of in-situ *wh*-phrases is not due to a general clause-boundedness constraint, but is instead the result of a pernicious interaction between *wh*-Agreement and Ā-operations affecting embedded clauses. Specifically, licensing of a *wh*-phrase is disrupted by Ā-movement of the clause that contains it, an operation that is highly favored and sometimes obligatory with finite CPs. I develop a Minimality-based analysis that explains this interaction. Licensing in-situ *wh*-phrases requires establishing an Agree relationship between the interrogative C and the *wh*-phrase[1] (Simpson, 2000; Watanabe, 2006). The features on the head of the moving clause, however, are sufficiently similar to [*wh*]-features and intervene for Agree between the higher C and the embedded *wh*, an intervention that has cascading effects on the derivation as a whole. Crucially, when this illicit configuration can be avoided, long-distance *wh*-in-situ is possible in Malayalam, demonstrating that potential differences in the scope-taking abilities of embedded *wh*-phrases in certain environments need not reflect a radically different syntax of *wh*-in-situ in this language.

The remainder of this paper is organized as follows. I begin by presenting a general picture of *wh*-question formation in Malayalam, including the fact that in certain multi-clausal constructions, long-distance *wh*-in-situ is disallowed. I will then show that the ungrammaticality of long-distance *wh*-questions reflects a general inability of *wh*-phrases to take scope out of embedded clauses — finite or non-finite — that undergo clausal fronting, an Ā-operation common in the language. I then present the core of my analysis. I argue that interrogative C in Malayalam is a generalized probe that can interact with any number of featurally more specific goals, including those on the head of the fronting clause. As a result, in configurations where the clause dominating the *wh*-phrase bears [Ā]-features relevant for fronting, the interrogative C erroneously makes contact with it. Next, I turn to cleft questions, which provide further evidence that restricted scope behavior of *wh*-in-situ is the result of fatal interactions among Ā-operations. In cleft configurations, unlike simple clauses, in-situ *wh*-phrases can take scope out of finite clauses, a pattern I will argue is due to the fact that the illicit interaction configuration is circumvented in clefts. I conclude with a a discussion of possible extensions of the present account to other restricted scope languages, as well as unrestricted scope languages like Japanese.

[1]The syntax of in-situ *wh*-expressions is much-debated. At least three camps of analysis can be found in the literature: (*i*) *wh*-phrases covertly move to C (e.g. Huang 1982), (*ii*) some Q-related operator, not the *wh*-element itself, undergoes movement (Cable, 2010; Hagstrom, 1998) and (*iii*) there is no movement at all (Baker, 1970; Reinhart, 1998). As will be shown in §2, a covert-movement analysis is not supported by the Malayalam data. For present purposes, I will assume that Malayalam *wh*-phrases remain in their base position. The syntactic link between the *wh*-expression and its scope position is taken to be established via Agree.

2 In-situ *wh*-expressions in Malayalam

The main question formation strategy in Malayalam involves *wh*-in-situ. In simplex sentences, *wh*-phrases can be left in their base position and yield a question interpretation. We see this in (3).

(3) a. **Eethu kuTTi** ee pustakam vaayi-chu?
Which child this book read-PERF
'Which boy read this book?'

b. Raman **eethu pustakam** vaayi-chu?
Raman which book read-PERF
'Which book did Raman read?'

In-situ *wh*-phrases in Malayalam do not appear to be island-sensitive, as illustrated by the fact that they are acceptable inside relative clauses (4a) and adjuncts (4b), both islands for overt movement:

(4) a. Raman [$_{\text{RelC}}$ **aaru** thann-a] pustakam vaayichu?
Raman [who gave-REL] book read?
'Which x is such that x gave Raman a book and Raman read that book?'

b. Nee Raman-e [$_{\text{TempC}}$ **enthu** paranja-ppol] shakaarichu?
you Raman-ACC [what said-while] scolded?
'What is such that you scolded Raman when he said it?'

This section will present a fuller picture of *wh*-question formation in Malayalam as necessary background. I will start out by providing evidence showing that *wh*-phrases in Malayalam remain in their base position and do not covertly move to C. I will then show that despite not involving movement, in-situ *wh*-phrases appear to show puzzling locality restrictions that appear unique to this strategy.

2.1 Malayalam *wh*-phrases as genuinely in-situ

Focus Intervention effects can establish that in-situ *wh*-phrases in Malayalam remain in their base position at the time of interpretation. *Focus intervention* is the name given to the phenomenon where in-situ *wh*-phrases are forbidden from appearing in the scope of certain scopal or focus-related elements, like negation or *only*. In Beck (2006), focus intervention is argued to be at the source of the ungrammaticality of the Korean sentence in (5), where the *wh*-phrase *nuku-lul* 'who-ACC' occurs in the scope of the focus sensitive operator *man* 'only'. Observe that scrambling the *wh*-phrase past the offending element to a higher position, as in (6), seems to repair the otherwise illicit structure.

(5) *Minsu-man nuku-lul po-ss-ni?
Minsu-only who-ACC see-PAST-Q
'Who did only Minsu see?'

(6) Nuku-lul Minsu-man po-ss-ni?
who-ACC Minsu-only see-PAST-Q
'Who did only Minsu see?' (Beck and Kim 2006)

In Beck's (2006) analysis of such effects, which builds on earlier work by Rooth (1985) and Hamblin (1973), both *wh*-phrases and focus phrases introduce focus-alternatives into the computation. Intervention occurs when the alternatives introduced by the *wh*-phrase are "picked up" by another operator before they can reach the appropriate interrogative operator. It is straightforward why movement, overt or covert, should repair the illicit configuration — the alternatives generated by the *wh*-phrase are longer be in the scope of the offending operator, and thus cannot be accidentally accessed by it.

Focus Intervention effects have been argued to distinguish between covert movement and in-situ composition of *wh*-phrases, sometimes within the very same language (e.g. Pesetsky 2000, Cable 2010, Kotek 2014). For instance, observing that in-situ *wh*-phrases in English Superiority-violating questions (7a), but not Superiority-obeying (7b) questions, were subject to intervention effects, Pesetsky (2000) argued that the two types of questions differed in where the in-situ *wh*-phrases were interpreted.

(7) a. **Which professor* did **only Fred** introduce *which student* to ___?
b. *Which student* did **only Fred** introduce to ___ *which professor*?

Though in both cases, the second *wh*-phrase is pronounced in its base position, in cases like (7b), it was argued to undergo covert movement, such that at the point of interpretation, it is no longer in the scope of the potential intervener.

If these arguments are on the right track, focus intervention effects can serve as a useful tool for determining the position of a *wh*-phrase at the time of interpretation, even when this is not transparent from the actual pronunciation site. When we apply the diagnostic to Malayalam, we find that the language patterns with Korean and Superiority-violating English questions in displaying the intervention effect. Whenever a focus-related element occurs to the left of an in-situ *wh*-phrase, the resulting structure is ungrammatical, as illustrated by the contrast between (8a) and (8b). As in Korean, scrambling the *wh*-phrase to a higher position repairs the violation (8c).[2]

[2]Mathew (2014) argues that scrambling is not a repair option for Focus-Intervention configurations in Malayalam. However, all of her examples involve scrambling the *wh*-phrase to a sentence-initial position, which may be dispreferred in Malayalam for independent reasons, like a strong preference in the language to construe sentence-initial constituents as topical (cf. Jayaseelan 2001).

(8) a. Rajan pustakangal aar-kke koduthu?
Rajan books-ACC who-DAT gave
'Who all did Rajan give books to?'

b. *Rajan pustakangal maatram aar-kke koduthu?
Rajan books only who-DAT gave
'Who all did Rajan give only books to?'

c. Rajan aar-kke pustakangal maatram koduthu?
Rajan who-DAT books only gave
'Who all did Rajan give only books to?'

If the research tradition of Pesetsky (2000). Beck (2006) and their successors is correct, the susceptability of Malayalam *wh*-phrases to intervention effects suggests that they do not raise to C, overtly or covertly.

2.2 *Wh*-asymmetries

The *wh*-in-situ strategy, which we saw to be unproblematic in mono-clausal constructions, nevertheless fails in certain bi-clausal constructions. In particular, *wh*-phrases inside finite clausal complements appear to be highly restricted. In (9), where the *wh*-phrase is embedded in the complement of *know*, which can take both interrogative and declarative complements, only the embedded question interpretation is available.

(9) [Sita **eethu pustakam** vaayich-ennu] Raman-u ariyaam
[Sita which book read-that] Raman-DAT know
✓Embedded Q: 'Raman knows which book Sita read.'
✗Matrix Q: 'For which book x does Raman know that Sita read x?'

A verb like *vicaarikk-* 'think', on the other hand, cannot take question complements. When an embedded *wh*-phrase occurs inside its complement, the resulting structure is ungrammatical, as neither a narrow scope nor wide scope reading seems to be available.

(10) *[Sita **eethu pustakam** vaayich-ennu] Raman vicaarichu?
[Sita which book read-that] Raman thought
✗Embedded Q: 'Raman thought which book Sita read.'
✗Matrix Q: 'Which book did Raman think Sita read?'

This restriction is puzzling in light of the fact that the same clause is transparent for overt extraction, at least in the case of bridge verbs like *say* or *think*. Example (11) shows that long-distance relativization from finite complements is possible. In (12), we see that clefting[3] from a finite embedded clause is possible.

[3]See §5 for evidence that Malayalam clefts involve overt movement.

(11) [[Sita ___ vaayich-ennu] Raman paranj-a] pustakam ivide unde
[[Sita read-that] Raman said-REL] book here COP
'The book that Raman said Sita read is here.'

(12) War and Peace aane [Sita ___ vaayich-ennu] Raman vicaaarich-athe
War and Peace COP [Sita read-that] Raman thought-NOMNL
'It's War and Peace that Raman thought Sita read.'

In addition, as we see below in (33), embedded *wh*-phrases seem to be able to take matrix scope out of non-finite clauses.

(13) Raman [**eethu pustakam** vaayikk-aan] shramichu?
Raman [which book read-INF] tried
'Which book did Raman try to read?'

One might jump to the conclusion at this point that it is a core fact about Malayalam that finite clauses are scope islands for *wh*-expressions. However, examples like (22) are special in another way: in addition to being finite, the embedded clauses in such examples appear in a preposed position rather than the canonical object position. A critical question that is now raised is which of the two factors — finiteness or fronted clause position — is responsible for the unavailability of wide scope. I will argue that it is clause position and not finiteness that is directly relevant for *wh*-scope. We find that when clauses appear in an immediately preverbal object position, *wh*-phrases inside them can take matrix scope; when they appear in a fronted position, this is not possible. The next section examines in more detail the correlation between clausal-fronting and *wh*-scope.

3 Clausal fronting and *wh*-in-situ

Malayalam is an SOV language, but clausal complements can and sometimes must appear sentence-initially, resulting in an OSV order, as already seen in (9) and (22).[4] It is often thought there is a strict correspondence between obligatoriness of fronting and finiteness — finite clauses must front, whereas fronting is optional for non-finite clauses (e.g. Srikumar 2007, Menon 2011). If this were the case, then it would be particularly difficult to ascertain which factor is directly responsible for the disruption of *wh*-scope. I will show, however, that fronting does not directly correlate with finiteness: not all finite clauses must front and not all non-finite clauses have the option to remain in a non-fronted position. What is relevant for fronting, instead, is the prosodic heaviness of the clause in question. In particular, I will suggest that prosodic constraints operative in the language

[4]Clauses can also appear sentence-finally. Preference for leftward versus rightward position of clausal complements seems to be subject to dialectal variation. Directionality of clausal displacement does not make a difference for the argumentation in this paper.

interact to prohibit certain heavy clauses from appearing in a medial position.[5] Later in this section, I will offer a precise characterization of heaviness relevant to fronting. What is crucial for us is that we can now construct examples in which finiteness and clausal fronting can be divorced. Once we do so, we are able to establish that it is fronting and not finiteness that is directly responsible for the *wh*-scope patterns we find in this language.

This section is organized as follows. I will first demonstrate that fronting is feature-driven syntactic movement, specifically Ā-movement. Next, I discuss the question of why certain clauses should possess the features relevant for fronting. I will suggest that the obligatoriness of fronting correlates with prosodic considerations, in particular, the weight of the embedded subject. Embedded clauses with overt subjects cannot remain in their base position, for reasons that I will argue relate to prosodic ill-formedness of such structures. But clauses with unpronounced subjects, finite or non-finite, can remain in-situ, and in such cases, embedded in-situ *wh*-expressions can freely take matrix scope.

3.1 Syntactic properties of fronted clauses

The fronted clause is generally the leftmost element in a given sentence. Previous analyses of fronted clauses generally take them to occupy a specific designated left-peripheral position (Hany Babu, 1997; Srikumar, 2007), an analysis I will adopt in this paper. I will argue, furthermore, that clauses undergo overt Ā-movement to this position. Evidence comes from obligatory reconstruction, syntactic locality effects and parasitic gap licensing, which I detail below.

3.1.1 Fronted clauses obligatorily reconstruct for binding

Consider (14) below; a pronoun inside the embedded clause can be bound by a quantifier in the matrix, even though the pronoun linearly precedes its binder[6]. The grammaticality of (14), then, shows us that fronted clauses *can* in principle reconstruct for binding and thus argues for movement.

(14) [avarude$_i$ kuTTi aane class-il onnaaman ennu] oroo sthree$_i$-um vicaarichu
[her$_i$ child COP class-in first that] each woman$_i$-UM thought
'Each woman thought that her child is first in class.'

Evidence that this movement is of the Ā-type comes from Principle C effects. It is well-known that reconstruction for Principle C is required for Ā-movement (Chomsky, 1995; Lebeaux, 1988). The ungrammaticality of (15a) illustrates that fronted clauses in Malayalam *must* reconstruct.

[5] See Dryer (1991) for a typological survey showing that movement of medial clauses to a peripheral position is commonplace for SOV languages.

[6] Mohanan (1982) argues that what matters for binding in Malayalam is linear precedence. This does not seem to be the case for the dialect spoken by my informants (from the Pathanamthitta region of Kerala, IN).

(15) a. *[Raman$_i$ aane class-il onnaaman ennu] avan$_i$ vicaarichu
[Raman$_i$ COP class-in first that] he$_i$ thought
'He thought that Raman is first in class.'

b. ?[avan$_i$ aane class-il onnaaman ennu] Raman$_i$ vicaarichu
[he$_i$ COP class-in first that] Raman thought

The examples in (15) do not contrast perfectly, as the embedded coindexed pronoun in (15b) is somewhat degraded. This degradedness is due to a blocking effect. Malayalam has a long-distance reflexive, *taan*, which takes human antecedents and is subject-oriented (e.g. Jayaseelan 1998). When the conditions for using *taan* are met, as it does in the subject position of the complement of an attitude verb, speakers of Malayalam prefer to use the reflexive instead of a bound pronoun. Thus, the counterpart of (15b) with *taan*, as in (16), is judged more acceptable.

(16) [taan$_i$ aane class-il onnaaman ennu] Raman$_i$ vicaarichu
[SELF$_i$ COP class-in first that] Raman thought

Cases like (15b) are nevertheless judged grammatical by speakers, so I take it to be sufficient evidence that the embedded clause reconstructs for binding. The example in (16) gives further support for a reconstruction analysis, since the anaphor can be bound by a matrix element, even when it is contained within a fronted clause.

3.1.2 Argument 2: Clause-fronting is island-sensitive

If clausal fronting is Ā-movement, we expect it to obey locality constraints on movement. We find that this is indeed the case. Relative clauses and temporal adjuncts are islands for Ā-extraction in Malayalam, as illustrated by the impossibility of clefting out of them in (17).

(17) a. *Sita aane Raman [___ thann-a]$_{\text{RelC}}$ pustakam vaayich-athe
Sita COP Raman [gave-REL] book read-NOMNL
Intended: 'It's Sita that Raman read a book that she gave him.'

b. *Sita aane Raman [___ wann-appol]$_{\text{TempC}}$ santhoshich-athe
Sita COP Raman [came-when] become.happy-NOMNL
Intended: 'It's Sita that Raman became happy when she came.'

Clausal fronting is similarly restricted when the clause is inside an island. As shown below, fronting out of relative clauses (18a) and temporal adjuncts (18b) is ungrammatical.

(18) *[Sita *War and Peace* vaayichu ennu]$_{\text{CP}}$ njaan [t$_i$ t$_{\text{CP}}$ paranj-a] aaL-e$_i$ kandu
[Sita War and Peace read that] I [said-REL] person-ACC saw
Intended: 'I saw the person who said that Sita read *War and Peace*.'

a. *[Sita war-um ennu]$_{CP}$ Raman [Amma t_{CP} paranj-appol]$_{TempC}$ santhoshichu
[Sita come-FUT that] Raman [mother said-when] became.happy

Intended: 'Raman became happy when mother said that Sita will come.'

3.1.3 Argument 3: Fronted clauses license parasitic gaps

Parasitic gap constructions, exemplified in (19), feature an ordinary gap left by movement and in addition, a gap inside e.g. an adjunct, which is understood as anaphoric to the original gap.

(19) Which article did you file ___ [without reading ___]?

As Engdahl (1983) observed, parasitic gaps are licensed only when there is an antecedent gap created by Ā-movement.[7] If clausal movement involves Ā-movement, then it, too, should in principle license parasitic gaps. This expectation is borne out.[8]

When we compare pairs of sentences containing gaps, which differ minimally with respect to the presence of clausal fronting, we find a contrast in acceptability. In (20a), the non-finite clause has fronted, and the parasitic gap is licensed; in (20b), where the embedded clause remains in-situ, a similar gap is quite odd.

[7] The precise reason for the restriction to Ā-movement is debated. Nissenbaum (2000), for instance, proposes that parasitic gaps involve the composition, by way of Predicate Modification (Heim and Kratzer, 1998), of two predicates of type $\langle e, t\rangle$, one derived by null-operator movement and the other by overt movement. The Ā-movement constraint follows if only this type of movement leaves the sort of variable that would result in the requisite type $\langle e, t\rangle$ predicate.

[8] A word of caution is in order, however, since certain properties of Malayalam make parasitic gap licensing a less-than-perfect diagnostic for Ā-movement in this language. Malayalam is a topic-drop language, so it is possible to ameliorate the ill-formedness of unlicensed gaps in the right contexts by imagining a dropped topic in that position. But such a strategy is generally difficult where no prior discourse exists to license topic-drop. Thus, in the absence of a rich context, there is a contrast between truly grammatically licensed gaps and those "rescued" by a topic-drop strategy. We see this when comparing a parasitic gap in a cleft construction (8) versus a passive (8); in the absence of a rich context, the cleft sentence with a gap is grammatical, but the passive sentence is quite odd.

(i) **A-movement: passive**

??War and Peace Raman-aal [___ vaayikk-aathe] verukka-ppeTTu
War and Peace Raman-BY [read-without] hate-PASS

'War and Peace was hated by Raman without having read.'

(ii) **A-bar movement: cleft**

War and Peace aane Raman [___ vaayikk-aathe] verukkunn-athe
War and Peace COP Raman [read-without] hates-NOMNL

'It's War and Peace that Raman hates without having read.'

(20) a. War and Peace vaayikk-aan Raman ___ ishtapped-aathe shramichu
War and Peace read-INF Raman enjoy-without tried
'Raman tried to read War and Peace without enjoying the activity.'

b. ??Raman War and Peace vaayikk-aan ishtapped-aathe shramichu
Raman War and Peace read-INF enjoy-without tried

I take these data to serve as evidence that parasitic gaps are genuinely licensed by the clause fronting operation, which in turn supports the view that fronting involves Ā-movement.

3.2 Why do clauses front?

The evidence considered in the previous section establishes that clausal fronting is an Ā-movement operation. I will posit a triggering feature [FR] (to evoke the fronted position of the clause), which will be responsible for fronting. The unvalued variant of this feature, i.e. the probe, is taken to be located on a left-peripheral functional head, H, which attracts the relevant embedded clause to its specifier. This [FR]-feature will not have any particular semantics or interpretive function associated with it, the reasons that will become clearer later in this section. Rather, it should be thought of as an EPP-type feature. Despite the lack of an obvious function, the existence of such a feature is well-motivated given the evidence that fronting is narrow-syntactic movement and the assumption that all syntactic movement is feature-driven.

A question that arises at this point is what determines whether or not an embedded clause possesses an [FR] feature and when the configuration will involve the movement-triggering head, H. Clausal fronting does not have obvious semantic or information-structural correlates, making it difficult to subsume fronting under more familiar operations like topicalization or focalization. It is more commonly assumed that fronting is simply a language-specific requirement on all finite clauses. I will argue here that this generalization is at best incomplete. Upon closer examination of frontable clauses, we find that finiteness is neither sufficient nor necessary to make fronting obligatory. In what follows, I will outline the environments in which clausal fronting must happen, and offer a prosodic explanation for these patterns.

3.2.1 The relevance of embedded subjects

There is a strong correlation between the size of the embedded subject and the ability of its clause to remain in-situ. We saw earlier in (9) and (22), repeated below, that certain finite complements are required to front. In both of these examples, the embedded subject is an overt lexical noun. I will argue below the obligatoriness of fronting correlates with the presence of an overt embedded subject.

(21) [Sita **eethu pustakam** vaayich-ennu] Raman-u ariyaam
[Sita which book read-that] Raman-DAT know
✓Embedded Q: 'Raman knows which book Sita read.'
✗Matrix Q: 'For which book x does Raman know that Sita read x?'

(22) *[Sita **eethu pustakam** vaayich-ennu] Raman vicaarichu?
[Sita which book read-that] Raman thought
✗Embedded Q: 'Raman thought which book Sita read.'
✗Matrix Q: 'Which book did Raman think Sita read?'

Crucially, the presence of an overt lexical subject also makes fronting obligatory for non-finite clauses. Consider the contrast between (23) and (24). The embedded clauses in (23) has a null subject and may appear in either a preposed or a medial position; on the other hand, the embedded subjects in (24) are overt, and only the fronted position is licit.

(23) a. [veett-il var-aan] Raman aagrahichu
[home-LOC come-INF] Raman wished
'Raman wished to come home.'

b. Raman [veett-il var-aan] aagrahichu
Raman [home-LOC come-INF] wished

(24) a. [Sita veett-il var-aan] Raman aagrahichu
[Sita home-LOC come-INF] Raman wished
'Raman wished for Sita to come home.'

b. ??Raman [Sita veett-il var-aan] aagrahichu
Raman [Sita home-LOC come-INF] wished

Conversely, non-overtness of the embedded subject makes fronting optional for finite clauses. The sentences in (25) and (26) differ minimally with respect to the overtness of the embedded subject. We find that only in (25) is the embedded clause allowed to remain in-situ.[9]

[9]Embedded clauses with pronominal subjects are more acceptable in a medial position than those with lexical subjects, though dispreferred in comparison to embedded clauses with null subjects.

(i) a. ?Raman [nee Sita-ye kalyaanamkazhikk-anam ennu] paranju
Raman [you Sita-ACC marry-MOD that] said
'Raman said that you should marry Sita.'

b. [nee Sita-ye kalyaanamkazhikk-anam ennu] Raman paranju
[you Sita-ACC marry-MOD that] Raman said

This gives preliminary evidence that the acceptability of medial clauses is not categorical, but gradient, a property that would be difficult to capture under analyses on which fronting is an operation that is obligatory for finite clauses. For ease of exposition, however, I will not discuss clauses with pronominal subjects further, focusing instead on clear-cut cases of acceptability (medial clauses with null subjects) and unacceptability (medial clauses with overt lexical subjects).

(25) a. Raman [*pro* Sita-ye kalyaanamkazhikk-um ennu] paranju
Raman [Sita-ACC marry-FUT that] said
'Raman said that he will marry Sita.'

b. [*pro* Sita-ye kalyaanamkazhikk-um ennu] Raman paranju
[Sita-ACC marry-FUT that] Raman said

(26) a. ??Raman [Manu Sita-ye kalyaanamkazhikk-um ennu] paranju
Raman [Manu Sita-ACC marry-FUT that] said
'Raman said that Manu will marry Sita.'

b. [Manu Sita-ye kalyaanamkazhikk-um ennu] Raman paranju
[Manu Sita-ACC marry-FUT that] Raman said

In contrast, the weight of the other constituents in the clause does not have a similar effect. Both examples in (27) have null subjects, but they differ with respect to the weight of the direct object. We find that in both cases, the embedded clause may remain in-situ.

(27) a. Raman [pustakam vaayikk-aan] shramichu
Raman [book read-INF] tried
'Raman tried to read a book.'

b. Raman [skul-il ninnu kitti-a pustakam vaayikk-aan] shramichu
Raman [school-ABL from get-REL book read-INF] tried
'Raman tried to read the book he got from school.'

What is the reason behind the special prosodic status of subjects in Malayalam? The link between prosodic heaviness and overtness of subjects may lie in the fact that subjects receive main stress and are perceived to be prominent in Malayalam[10], which, in turn, affects the way a given sentence is parsed prosodically. Fery (2009), in her prosodic analysis of Malayalam, notes that there is a strong prosodic boundary at the right edge of the subject in canonical sentences. I take this as indication that Malayalam intonationally distinguishes the subject from the predicate. Formalizing this as a prosodic constraint active in the language, the next section offers an explanation for why fronting is rendered obligatory for a given clause.

3.2.2 Syntax-prosody mapping and bi-clausal sentences

In this section, I will show that the clause-fronting patterns that we find in Malayalam fall out from general principles of prosodic parsing. Specifically, I will argue that when a heavy clause remains in its base position, the resulting structure is *prosodically ineffable*: it cannot receive a prosodic parse that does not violate at least one prosodic well-formedness principle.

[10] This could be related to the fact that canonically subjects are construed as topical in Malayalam (see e.g. Mathew 2014).

I assume that prosodic structure is distinct from syntactic structure and consists of the following, hierarchically ordered prosodic categories above the foot: ω (prosodic word), ϕ (phonological phrase), and ι (intonational phrase). Furthermore, for any prosodic category, a sentence is exhaustively parsed into a sequence of such categories (Selkirk 1984). I take there to be (rough) correspondence between the syntactic structure and prosodic structure, as long as this mapping does not conflict with prosodic well-formedness (Hayes, 1990; Nespor and Vogel, 1986; Selkirk, 2011, 1984). One such well-formedness principle is Selkirk's (1984) Strict Layer Hypothesis, a variant of which is given in (28), which mandates that a non-terminal prosodic category of one level may not be dominated by a prosodic category of a lower level in the hierarchy.

(28) Strict Layering: A prosodic category of level n (e.g. an ι-phrase) cannot dominate a category of level $n-1$ (e.g. a ϕ-phrase)

Another relevant principle that has been shown to create mismatches in the syntax-prosody mapping (e.g. Elfner 2012), involves phonology-specific binarity restrictions on phrase sizes. For Malayalam, I will propose that intonational phrases must be composed of maximally two phonological-phrases (29).

(29) MaxBin(ι,ϕ): An ι-phrase must contain maximally two ϕ-phrases

An idea that goes back to at least Selkirk (1984) is that in addition to syntax, certain extraneous features like *prominence* may also yield amendments to prosodic structure. We can formalize this notion with a constraint as in (30), which states that a prominent element must head a phonological phrase of its own.

(30) ProsProm: Prosodically prominent constituents must head a minimal ϕ-phrase

In canonical sentences in Malayalam, the subject receives prominence,[11] and the requirement in (30) would account for the observation that the language intonationally distinguishes the subject from the predicate Fery (2009).

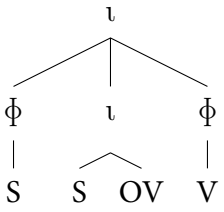
With these assumptions, let us now turn to the prosodic parsing of multi-clausal sentences in Malayalam. Medial clauses with an overt subject cannot be parsed in a way that does not violate at least one of the three principles laid out above, as will be explicated below. Suppose we were to parse each embedded clause as heading its own intonational phrase, as in (31a): this would result in the embedded clause ι-phrase being dominated by the ϕ-phrase headed by the embedded verb, in violation of Strict Layering. Parsing each major constituent as heading its own ϕ-phrase, which is then dominated by a single ι-phrase is in keeping with Strict Layering, but goes against the restriction

[11] Note that in scrambling configurations, it is the scrambled element that receives prominence and heads its own phonological phrase (Swenson et al. 2015).

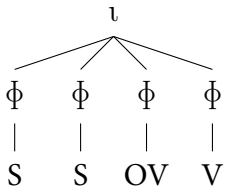
against long strings of ϕ-phrases (31b). There is a way in which we could satisfy Strict Layering and MinBin, as demonstrated in (31c). However, observe that this would fail to meet the requirement that the prosodically prominent subject head its own ϕ-phrase. When the embedded subject is unpronounced, ProsProm is no longer relevant, allowing for the possibility of medial clauses with null subjects.[12]

(31) *Structure:* S [S O V] V

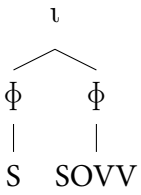
a. *Potential parse 1:*
Violates Strict Layering

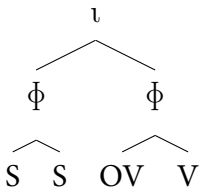


b. *Potential parse 2:*
Violates MaxBin



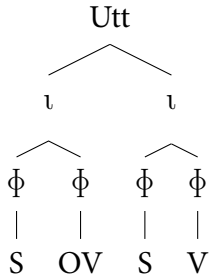
c. *Potential parse 3:*
Violates ProsProm





The availability of clause fronting as a syntactic mechanism in Malayalam makes a multi-clausal sentence utterable. A structure in which the clause has moved to a peripheral position can easily receive a well-formed prosodic parse, as illustrated in (32).

(32) *Structure:* [[SOV] SV]



[12] Though I have presented the prosodic requirements at play as obligatory rules for ease of exposition, they can be reformulated as violable constraints and incorporated into an Optimality Theoretic Model and the same conclusions should follow. However, doing so involves making claims about the syntax-prosody interface (e.g. what constitutes the input that defines the competing candidates) that go beyond the scope of this paper.

Crucially, in certain circumstances, bi-clausal sentences fail to have any acceptable prosodic parse *unless* the clause undergoes movement to a peripheral position. The idea that clausal movement to peripheral positions is driven by the need to circumvent "prosodic monsters" is now commonplace within accounts of extraposition (Féry, 2011; Hartmann, 2013; Manetta, 2012; Proceedings of WECOL, 2007; Truckenbrodt, 1995).[13] Manetta (2012), for instance, argues that finite clauses in Hindi, a restricted scope language (see §6.1), are prosodically aligned to the right edge. Similar patterns of clausal movement have also been observed by Potsdam and Edmiston (2016) for Malagasy, which extraposes embedded clauses when the subject is overt. As Potsdam and Edmiston (2016) argue, the relevant prosodic property of Malagasy, as in Malayalam, is that subjects are prosodically set apart from the rest of the clause, which in turn creates a prosodic parsing problem when the clause is left in-situ.

Let me remark on one final issue before moving on. Evidence in the previous section showed clausal fronting to be run-of-the-mill Ā-movement, and thus run-of-the-mill *syntactic* movement. How should we understand the elements driving clausal-fronting — the head H and the feature [FR] — if ultimately, the driving force behind-the-scenes is prosody? The logic is the same as the distinction advanced by Fanselow (2007) between "triggering" and "exploitation". The prosodic parser can *exploit* the result of a syntactic operations, even when the elements *triggering* the operations are purely syntactic. In line with this logic, I will suggest that [FR]-features do not have a prosodic interpretation, per se, as prosodic structure is built on the *output* of syntactic derivation. Instead, [FR] should be seen simply as a word-order or EPP-type feature. The presence of H and [FR] may very well be syntactically *optional*, for both finite and non-finite complements. However, a derivation without [FR] and the resulting clausal fronting would leave the prosodic parser with an irredeemable structure, at least in the case of prosodically heavy clauses.

3.3 Interaction with *wh*-scope

Having established that clause fronting is triggered by an optional FR feature, which is rendered de-facto obligatory for prosodically heavy clauses, we can now turn to the question of how this operation interacts with *wh*-in-situ. The last section demonstrated that fronting need not dovetail with finiteness. Non-finite clauses with null subjects and finite clauses with null subjects both front optionally. This allows us to construct test-environments where finiteness and clause position are disentangled, to identify which of the two factors has an effect on the scope of embedded *wh*-phrases.

Suppose the *wh*-licensing domain for Malayalam were the immediate finite clause, as proposed for Iraqi Arabic and Hindi by Ouhalla (1996); Simpson (2000) and others. We would then expect

[13] Not every one of these analyses take clausal movement to take place in the syntax, as I have argued is the case in Malayalam.

that an embedded *wh*-phrase inside a finite clause may not to take matrix scope, irrespective of clause position. Conversely, we expect matrix scope to be possible for *wh*-expressions inside non-finite complements, irrespective of position. If on the other hand, clause-position determines scope possibilities, we predict that fronted clauses block wide scope and in-situ clauses permit it across-the-board. We see below that the second set of predictions is borne out. When the clause appears in its base position, embedded *wh*-phrases can take matrix scope, irrespective of finiteness of the clause:

(33) **Non-finite clauses**

Raman [**eethu pustakam** vaayikk-aan] shramichu?
Raman [which book read-INF] tried
'Which book did Raman try to read?'

(34) **Light finite clauses**

Nee [**enthu patti** ennu] vicaarichu
You [what happened that] thought
'What did you think happened?'

A minimal modification of the examples above in terms of clause position results in a change in grammaticality, as we see in (35 - 36). The (b) sentences show that the corresponding declaratives are licit. This, too, is irrespective of finiteness of the clause.

(35) a. ∗[**eethu pustakam** vaayikk-aan] Raman shramichu?
[which book read-INF] Raman tried
'Which book did Raman try to read?'

b. [*War and Peace* vaayikk-aan] Raman shramichu.
[War and Peace read-INF] Raman tried
'Raman tried to read *War and Peace*.'

(36) a. ∗[**enthu patti** ennu] nee vicaarichu
[what happened that] you thought
'What did you think happened?'

b. [abhatham patti ennu] njaan vicaarichu
[mistake happened that] I thought
'I thought a mistake happened.'

On the basis of the above examples, we can form the following generalization about *wh*-licensing in Malayalam:

(37) **Clause position – *Wh* Correlation**
A *wh*-phrase cannot take scope outside of a fronted embedded clause.

As a consequence, for *wh*-containing clauses that *obligatorily* front, neither the fronted nor the in-situ option is available. A fronted heavy clause with an embedded *wh* will be impossible in the matrix-question reading, as fronting appears to restrict *wh*-scope. An unfronted heavy clause is illicit because of the resulting prosodic illformedness of medial heavy clauses. The following section offers an explanation for the correlation in (37).

3.4 Section summary

This section focused on the phenomenon of clausal fronting, which I argued was directly responsible for restricting the scope of *wh*-in-situ. The role of fronting for *wh*-licensing is often obscured by the fact that fronting is common with finite clauses. It was shown that obligatory fronting is conditioned by prosodic factors and that finite clauses that are sufficiently light have the option of remaining in their base position. Such clauses, when they remain in-situ, allow for embedded *wh*-phrases to take scope out of them, suggesting that *wh*-in-situ in Malayalam cannot be accurately described as having clause-bound scope across-the-board.

4 A locality-based approach to restricted scope

The main question that arises from this revised picture is why fronting of a clause keeps a *wh*-expression from taking scope outside of that clause. We saw in the previous section that fronted clauses obligatorily reconstruct, so it is not the case that they are outside of the domain of the interrogative C at the point of interpretation. Furthermore, since overt extraction is possible out of fronted clauses, as we previously saw in (11) and (12), fronted clauses are not altogether opaque for syntactic operations. When, then, does fronting creates a scope island for *wh*-expressions? I will argue as follows: Because the Agree relations triggered by the clause-fronting feature [FR] and by [*wh*] are both Ā-operations, the two features are similar enough to interact. Crucially, this interaction has fatal consequences. The [FR]-features on the head of the embedded clause intervene for Agree between a higher interrogative C and the embedded *wh*-phrase, leading C to erroneously Agree with the head of the clause. This, in turn, uses up features relevant for fronting, leaving the higher head H without a suitable goal.

4.1 Crossing Ā-paths

A core idea within accounts of syntactic locality is that syntactic relations must involve the closest eligible elements (e.g. Rizzi 1990, Chomsky 1995) . The contrast between the sentences in (38) is attributable to the fact that in the (b) example, a lower *wh*-phrase is the target for movement by the

lower C, skipping over a higher *wh*-expression.

(38) a. [What subject]$_i$ do you know who$_j$ PRO to talk to t$_j$ about t$_i$

b. *Who$_j$ do you know [what subject]$_i$ PRO to talk to t$_j$ about t$_i$

However, as observed by Pesetsky (1982), similar ungrammaticality can obtain even when the interacting operations are not the same type. Examples (39) and (40) are parallel to (38) above, except that in these examples, *wh*-movement interacts with relativization and topicalization respectively.

(39) **Interaction 1: *wh...Relativization***

a. chess, which$_i$ I wonder who$_j$ you believe t$_j$ to play t$_i$ well

b. *John, who$_j$ I wonder [what game]$_i$ you believe t$_j$ to play t$_i$ well

(40) **Interaction 2: *wh...Topicalization***

a. This problem$_i$, Mary knows who$_j$ PRO to consult $_j$ about t$_i$

b. *This specialist$_j$, Mary knows [what problems]$_i$ to consult t$_j$ about t$_i$

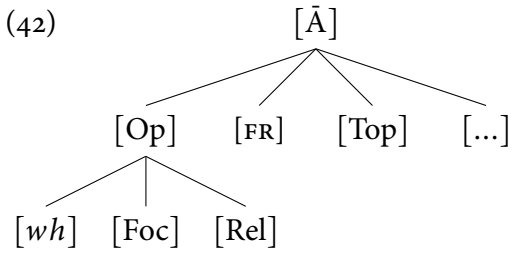
Thus, features like [Topic] and [Rel] are interveners, in English at least, for an operation typically taken to target [*wh*]-features. This behavior can be captured under the classical conception of Relativized Minimality (Rizzi, 1990), where an intervening element of the same class (A, Ā, Head) blocks operations belonging to the same class. However, such an approach would overgenerate, as these interactions seem to be subject to cross-linguistic variation. A case in point is Topicalization, which, as we saw in (40), interacts with *wh*-movement in English, but not in Italian or German (Müller and Sternefeld, 1993; Rizzi, 2004). An Italian example is given in (41) for illustration.

(41) Mi domando, il premio Nobel, a chi lo potrebbero dare
‘ I wonder, the Nobel Prize, to whom they could give.’ (Rizzi, 1997)

The more restrictive conception of Minimality in Chomsky's (1995) Attract Closest, which requires featural identity between the target and the intervener, is too selective, as it fails to account for the interactions in (39) and (40).

What seems to be necessary is a more articulated feature geometry for Ā-elements, as has been developed for the class of φ-elements (by Harley and Ritter 2002, Béjar and Rezac 2009 and Preminger 2012, among others). These approaches posit hierarchies of features, with subclasses and superclasses and entailment relations among features within the hierarchy. Furthermore, it is assumed that probes may be relativized to more or less specific features, and this relativization may vary across languages. Building on Starke (2001), Rizzi (2004), Abels (2012) and others, I will extend the feature-geometric approach to the Ā-domain and make use of a hierarchy as in (42).

(42)



Under the model sketched above, the feature [*wh*] entails higher-level features like [Op] and [Ā]. I will use the symbol '→' to mark an entailment relation between features. Thus, the notation [*wh*] → [Ā] would indicate that the head bearing a [*wh*]-feature also, by entailment, bears an [Ā]-feature.

Let us see now how these tools help us capture the interactions in, for example, the ungrammatical topicalization example from above. Suppose C in English is a relatively flat probe that looks for [Ā]-features. Let us also take for granted that the topicalized phrase *the specialist* bears a [Topic]-feature and that *what problems* bears a [*wh*]-feature. Both [Topic] and [*wh*] will entail [Ā] under to our system. Therefore, a relationship between C and the *wh*, as schematized in (43), is impossible, since *the specialist* also bears the relevant [Ā]-feature.

(43) [...[**C**$_{[\bar{A}]}$ [to consult **the specialist**$_{[\textbf{Top}] \rightarrow [\bar{A}]}$ about **which problem**$_{[\textit{wh}] \rightarrow [\bar{A}]}$]]]

✗

The lack of similar interactions in Italian would be due to a different feature-specification on the probe. So, the Italian C may probe at the [Op] level, which would mean that Focus and Relative operators would intervene, but not Topics.

4.2 Ā-interactions in Malayalam

The feature geometric approach outlined above allows us to capture intervention effects that occur even when the features involved are not identical. Note, however, that the cases that we discussed in the previous section all involved two movement steps. This sets them apart from the Malayalam cases we are interested in, which involve only one instance of movement (the *wh*-phrase remains in-situ). However, the Malayalam configurations in question do involve two Agree steps. Given that Agree is a prerequisite for movement, a recasting of Minimality as a locality constraint on Agree allows us to capture both cases under the same umbrella. I will therefore be adopting a view of Minimality as in (44). Closeness is defined in terms of c-command, along the lines outlined in (45).

(44) **Generalized Minimality**
A probe must Agree with the closest possible goal bearing the relevant features.

(45) **Closeness**
A probe α which c-commands two goals β and γ is closer to β than to γ if β c-commands γ

The ungrammaticality of certain Malayalam long-distance questions, I will argue, is in part due to the fact that any derivation that associates the *wh*-element with interrogative C when the *wh* is inside a fronted clause also violates the constraint in (44). This is so because the interrogative C in Malayalam is relativized to [Ā], a feature entailed by both [*wh*] and [FR], the feature responsible for clausal fronting. Thus, both are viable goals for C. In a configuration where the frontable clause dominates the *wh*-phrase, the head of the clause will necessarily be structurally higher and, by the definition in (45), a closer goal for C than *wh*. Thus, as long as the [FR]-feature is present on the clause head, *wh*-Agreement between C and a clause-internal *wh*-phrase cannot take place.

Before going into the details of this account, I want to take note of some necessary auxiliary assumptions:

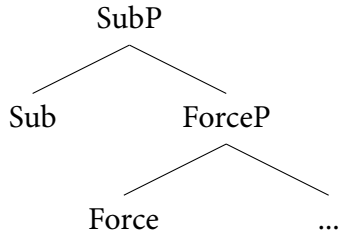
1. The Malayalam left periphery

I will take the Malayalam left-periphery to involve into (at least) two distinct projections, in the spirit of Rizzi (1997). This is because subordination markers like *ennu* 'that' co-occur with the Q-morpheme *-oo*, as illustrated in (46).[14]

(46) [Raman War and Peace vayich-oo ennu] Amma chodichu
[Raman War and Peace read-Q that] mother asked
'Mother asked if Raman read War and Peace.'

Concretely, I will take the C-domain in Malayalam to be split into two projections: Sub, responsible for hosting subordination markers like *ennu* 'that', and Force, responsible for hosting Q-morphemes. I take interrogative C to instantiate Force, even when an overt Q-morpheme is absent. The structure is schematized in (47).

(47)



For convenience, I will continue throughout the paper to refer to the element responsible for question-licensing as the interrogative C.

[14] Though the question morpheme is not overtly present in *wh*-questions, I will follow Hagstrom (1998) and Cable (2010), among others, in assuming that a phonologically null variant is nevertheless present. Additional support for this comes from the fact that historically, the question particle *-oo* was pronounced even in *wh*-questions (Jayaseelan, 2001).

2. Locality of Agree

The Phase Impenetrability Condition (PIC) (Chomsky, 2001) prohibits syntactic operations from occuring across phase boundaries; only phase edges are accessible to higher operations. I will assume that Agree is subject to locality and that the PIC restricts its application (see e.g. Adger and Ramchand 2005. For Malayalam, the phase in question is CP. As we saw previously, long-distance *wh*-questions *are* possible in certain cases. We therefore need a mechanism for implementing Long Distance Agreement in a way that doesn't violate the PIC. For present purposes, I will adopt a version of Cyclic Agree, proposed in Legate (2005). The main idea behind this approach is that Long Distance Agree proceeds via a number of local Agree steps.[15] Implementing this idea requires an additional assumption that intermediate probe features, once valued, may act as a goal for further instances of Agree by higher probes.

For the relevant cross-phasal dependencies in Malayalam, I will take the embedded C to serve as an intermediary. The role of the intermediate C in the derivation of long-distance *wh*-movement is well-established: it is often proposed that the intermediate C, while not an interrogative complementizer in and of itself, possesses a semantically inert [*wh*]-feature that triggers movement of the embedded *wh* to its specifier (Abels, 2012; McCloskey, 2002). Similarly, I take the intermediate C in Malayalam to be equipped with a [*wh*]-feature, which, once valued via Agree with the *wh*-phrase itself, can serve as the goal for the higher C (see Munataka 2006 for a similar proposal for Japanese long-distance questions).

3. Properties of Agree and features

Following Preminger (2012), the operation Agree is obligatory, but if a probe does not find a matching goal in its domain, this failure does not result in ungrammaticality. The Activity Condition (Chomsky 2000, 2001), defined in (48), ensures that once a set of features have been Agreed with, they are ineligible for further operations.

(48) **Activity Condition**
A feature on a given head may serve as a goal for Agree only once.

[15] An alternative would be to posit that finite clauses in Malayalam are not phasal. Some researchers have argued in favor of a more contextual approach to phasehood (see e.g. Bošković 2005, 2014. Adjudicating between the two alternatives is not important for us at present, as the current proposal is compatible with either one.

The feature-specifications on the relevant probes in Malayalam are provided in (49).

(49) **Feature specifications in Malayalam**

Interrogative C:	$[u\bar{A}]$
Embedded C:	$[wh] \rightarrow [\bar{A}]$
H:	$[u\text{FR}]$

The interrogative C is taken to be a generalized [Ā]-probe. The *wh*-expression bears [*wh*]-features, which entail [Ā]. Assuming Cyclic Agree, in the case of finite CPs, the embedded C will serve as an intermediary and bear [*wh*]-features. Clauses that undergo fronting — SubP when finite, or a smaller structure, arguably AspP, when non-finite — will bear the requisite [FR]-features, the unvalued variant of which will be on the head H triggering fronting in the first place.

4.3 Deriving the patterns

Let us consider the two main patterns that need to be accounted for:

1. A *wh*-phrase inside an embedded clause cannot take matrix scope if the container has fronted.
2. A *wh*-phrase inside an embedded clause can take matrix scope if the container clause remains in-situ.

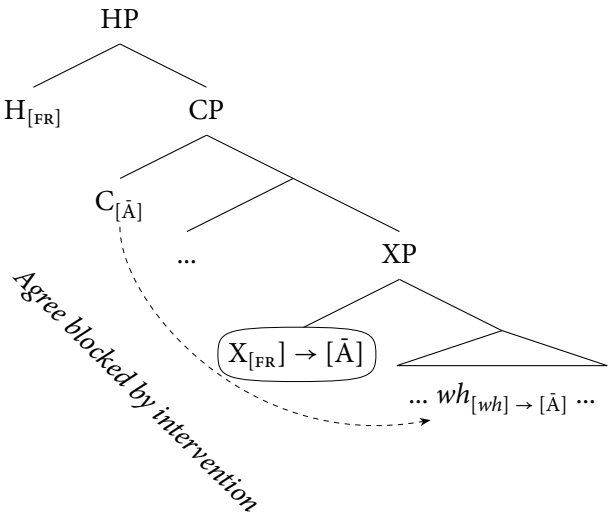
I discuss pattern 1 first.

4.3.1 Explaining ungrammaticality

In order for a construction to receive a question interpretation, an interrogative C must be present in the derivation and it must establish a syntactic dependency with the *wh*-element (see e.g. Adger and Ramchand 2005; Simpson 2000; Watanabe 2006. Clauses that undergo fronting must bear the requisite [FR]-feature. According to the feature-specifications in (49), the interrogative C in Malayalam is a generalized [Ā]-probe. It can successfully make contact with the appropriate [*wh*]-feature-bearing element as long as no other [Ā]-feature intervenes.

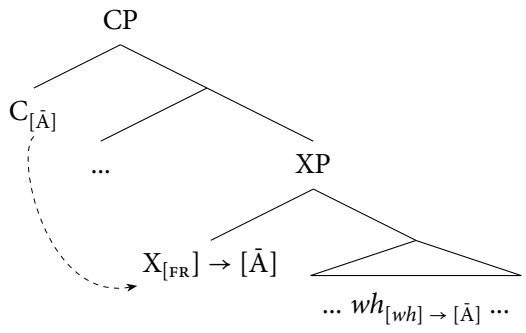
Now, consider what will happen when the head of the clause dominating the *wh*-expression bears an [FR]-feature. Because [FR] entails [Ā], the head of the clause bearing this feature will necessarily be a closer goal for C than the [*wh*]-feature bearing element in these configuration. The result is that Agree between C and *wh* itself is blocked. The intervention configuration is schematized in (50).

(50) Intervention Configuration

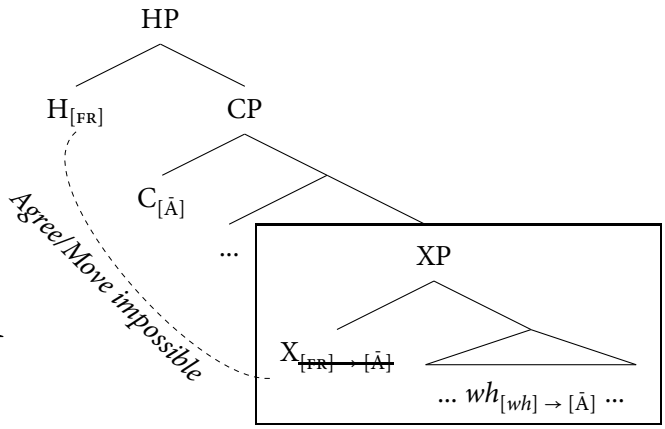


Instead, the generalized probe on C will make erroneous contact with [FR]. By the Activity Condition stated in (48), the Agree relation between C and the head of the clause renders [FR] ineligible for further Agree operations, with the result that when the higher head H merges, it can no longer make contact with the frontable embedded clause. This is schematized in (51).

(51) a. C Agrees with X



b. H cannot Agree with X



If H cannot establish an Agree-relation with the embedded clause, it is impossible for the embedded clause to front. We now have a way of explaining the ungrammaticality of examples like (**??**), which involve both a matrix-scope-taking *wh*-phrase (and thus a matrix interrogative C) and a fronted clause: these structures can never be generated.

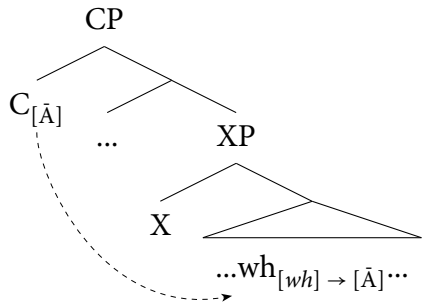
Given our assumptions about the nature of Agree, H's failure to find a goal need not force the derivation to crash. However, it does result in a structure in which the embedded clause remains in-situ. Depending on the size of that clause, this could result in a structure that cannot be pronounced, given the arguments in §3.2 concerning Malayalam prosody. Thus, for derivations involving a heavy

embedded clause, *wh*-Agreement and clausal fronting, we are faced with two alternatives that are both untenable: we can establish the relevant syntactic dependencies — between interrogative C and the *wh*-element and between H and the head of the frontable clause — at the cost of violating locality constraints on Agree, or we can obey locality and be left with a prosodically ill-formed structure.[16]

4.3.2 Explaining grammaticality

Recall that the presence of [FR] on the head of the clause is formally optional, though sometimes required for prosodic effability, as discussed above. The acceptable examples, in which *wh*-expressions can take wide scope, all involve *wh*-expressions embedded inside prosodically light clauses that remain in their base positions. Thus, such clauses need not possess the [FR]-features necessary for fronting. In such a derivation, there is no intervening [Ā]-feature to potentially interrupt Agree between C and *wh*. This is schematized in (52).

(52)



Because the generalized probe that is susceptible to erroneous interactions is located on the *interrogative C*, we do not expect clausal fronting to create problems in the absence of the [Ā]-feature-bearing C. In declarative sentences, from which the intervening [Ā]-probe is absent, the left-peripheral head H will be the first probe to make contact with the [FR]-features.

4.4 Section summary

This section showed that the restricted scope of certain embedded *wh*-phrases in Malayalam is the result of a pernicious interaction between two Ā-operations. A feature-geometric approach, along with a generalized view of minimality as a constraint on Agree, derived the patterns. Interrogative C, a flat [Ā]-probe, erroneously makes contact with [FR]-features on the head of the *wh*-containing embedded clause, in turn blocking both *wh*-Agreement and a higher head's attempt to Attract the clause.

[16] One of these alternatives involves violating a syntactic locality constraint generally taken to be fundamental, whereas the other involves illicit prosody. We might expect that derivations involving the prosodic violation is more tolerable than those involving the former. Testing this prediction is difficult, as it involves comparing ill-formedness of different types, but my informants do find *wh*-in-situ with medial heavy clauses to be somewhat better than *wh*-in-situ inside fronted clauses.

As a result of these interactions, derivations in which both *wh*-Agreement *and* clausal-fronting take place simply cannot be generated in this language.

5 Order of operations

The previous section developed an analysis of long-distance *wh*-in-situ in Malayalam in which apparent clause-boundedness is reduced to locality considerations: *wh*-Agreement fails if an intervening head bears features relevant for the generalized probe on C. In the ungrammatical cases considered above, Agree between C and the *wh* was attempted before clausal fronting, which meant that the [FR] features on the to-be-fronted clause was still active and an eligible goal for the [Ā]-probe on C. However, the Activity Condition defined in (48) predicts that the Agree between C and *wh* should be possible in the same configurations if the intervening Ā-features has already been targeted for Agree by some other probe. In other words, *the order of the two Ā-operations* is predicted to make a difference for whether or not long-distance *wh*-in-situ is possible. As it happens, cleft questions in Malayalam involve configurations where *wh*-Agreement is attempted *after* an Ā-operation affecting the embedded clause, namely clefting. As predicted, embedded *wh*-expressions in clefts can always take matrix scope.

Recall that long-distance *wh*-question formation is not possible when the *wh*-phrase is inside a clause that is required to front.

(53) *[Sita **eethu pustakam** vaayich-ennu] Raman vicaarichu? (=(22))
[Sita which book read-that] Raman thought
Intended: 'Which book did Raman think Sita read?'

To form the intended question in (53), one would use a cleft question, as in (54):

(54) [Sita **eethu pustakam** vaayichu ennu] aane Raman vicaarich-athe?
[Sita which book read that] COP Raman thought-NOMNL
'Which book was it that Raman thought Sita read?'

The cleft question above shares a number of properties with the ungrammatical non-cleft question in (53). In both, the *wh*-phrase remains in its base-position within the finite complement clause and in both, the container clause undergoes leftward Ā-movement. Yet, the *wh*-phrase can scope out of the embedded clause only in the cleft question. I will show below that the present account readily accounts for this pattern.

5.1 Background on Malayalam clefts

Clefts in Malayalam are biclausal constructions, involving the copula *aane*, a discourse-prominent or focused constituent (the *pivot*) and a cleft-clause headed by the nominalizer *athe*. The nominalized clause has a reduced structure and does not seem to support projections higher than TP. An example of a Malayalam cleft is given in (55):

(55) Raman aane Sita-kke pustakam koduth-athe
Raman COP Sita-DAT book gave-NOMNL
PIVOT COPULA CLEFT CLAUSE
'It's Raman that gave a book to Sita.'

As mentioned in previous sections, cleft-constructions in Malayalam show properties associated with Ā-movement. Furthermore, they display semantic effects taken to characterize clefts across languages, which include: (*i*) an existence presupposition that some entity satisfies the cleft clause and (*ii*) an inference that the entity denoted by the pivot exhaustively satisfies the predicate.

Analyses of clefts diverge on whether the pivot is base-generated in its surface position (e.g. Percus 1997), or whether the clefted constituent originates inside the cleft-clause and overtly moves to its surface position (Chomsky, 1977; Frascarelli and Ramaglia, 2013; Kiss, 1998). Malayalam clefts display case and idiom connectivity between the pivot and the cleft-clause, both explained straightforwardly under an overt movement analysis. Consider (56), in which the pivot *aana-ye* 'elephant-ACC' behaves as though case-marked by the verb inside the cleft-clause. (57) shows that the case connectivity persists in quirky case-marked experiencer subjects.

(56) aana-ye/*aana aane kutti *t* null-athe
elephant-ACC/*elephant.NOM COP child pinch-NOMNL
'It's the elephant that the child pinched.'

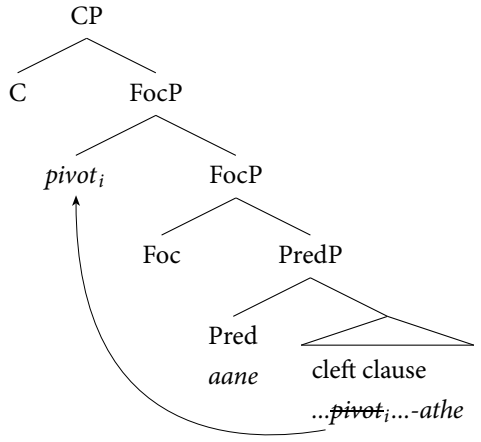
(57) Sita-kkə aane Raman-odə *t* deshyam thonni-athe
Sita-DAT COP Raman-SOC anger feel-NOMNL
'It's Sita that felt anger towards Raman.'

The sentence in (58) illustrates idiom-connectivity in clefts: a phrase that is part of a larger idiom chunk can serve as the clefted constituent, while retaining the idiomatic interpretation. The availability of the idiomatic interpretation for phrases separated on the surface gives evidence for reconstruction, which in turn, is evidence for movement.

(58) Annaan kunj-ine aane ningal maramkeettaan padippikk-unn-athe
Squirrel baby-ACC COP you.PL tree.climbing teach-PROG-ATHE
Lit. 'It's a baby squirrel that you are teaching how to climb.'
'It's an expert that you are trying to teach; this is unnecessary.'

I will therefore assume a movement analysis and the following syntactic structure for Malayalam clefts:

(59)



The crucial difference between clefting and clause-fronting, for our purposes, lies in the relative ordering of the heads triggering the Ā-operations — the interrogative C in clefts is structurally higher than the Focus head responsible for clefting. For cases as in (54) above, this means that wh-Agreement takes place *after* the container clause has undergone clefting. Furthermore, movement of the clause to a peripheral position via clefting should also solve any prosodic parsing issues that would otherwise arise with medial clauses. In other words, clausal clefting would carry out precisely the kind of movement that yields a prosodically parsable structure. Thus, neither FR-features nor the head H driving clausal fronting would then be necessary in these constructions.[17]

5.2 In-situ composition of *wh*-phrases in clefts

In cleft questions, the *wh*-phrase itself or a larger pied-piped constituent occupies the pivot position, as we see in (60).

(60) aare aane Sita-kke pustakam koduth-athe?
who COP Sita-DAT book gave-NOMNL
'Who is it that gave a book to Sita?'

Though the construction in (60) involves movement, I will argue that it is not an instance of *wh*-movement. Rather, I suggest that clefting of the *wh*-phrase is driven by the same mechanism driving non-*wh*-clefting, a [Foc]-feature on the clefted constituent. As discussed in §2, both *wh*-phrases and Focus-phrases have been argued to project Focus-alternatives (e.g. Beck 2006). Furthermore,

[17] It is also the case that clefts are prosodically distinct from canonical sentences, with main stress falling on the clefted constituent (e.g. Swenson et al. 2015). It could be the case that prosodic ill-formedness that could lead to clausal movement in canonical sentences simply do not arise in cleft configurations.

wh-phrases and Focus phrases compete for the same position in languages that have a designated syntactic focus position, e.g. Hungarian (Brody, 1990; Kiss, 1998; Szabolcsi, 1981). It does not seem unreasonable, then, to say that *wh*-phrases also bear [Focus]-features (see Lipták 2001 and Haida 2007 for similar claims).

From the clefted position, licensing of the *wh*-phrase proceeds in the same way as ordinary in-situ *wh*-phrases, via in-situ composition. Evidence for this once again comes from intervention effects. Recall that ordinary in-situ *wh*-questions in Malayalam are subject to intervention from a focus-sensitive element that occurs between the *wh*-phrase and the interrogative C. We find similar intervention effects in cleft questions. It is possible to scramble material from the cleft-clause to a position to the left of the clefted constituent. In keeping with the literature on scrambling, I will take the landing site for scrambled constituents to be a position below C. In (61), *Rajan* is the pivot, but *pustakangal* 'books' has scrambled to its left.

(61) Pustakangal Rajan aane Sita-kke koduth-athe
Books Rajan COP Sita-DAT gave-NOMNL
'It's Rajan that gave books to Sita.'

In a cleft question, a focus-sensitive element scrambled to the left of the pivot can lead to an intervention effect.

(62) a. Eethu pustakam aaNe Lily maatram vaayich-athe?
which book COP Lily only read-NOMNL
'Which book was it that only Lily read?'

b. ??Lily maatram eethu pustakam aaNe vaayich-athe?
Lily only which book COP read-NOMNL (Kim 2002)

The parallel effects in cleft and non-cleft questions suggest that *wh*-interpretation in both cases happens the same way, although the *wh*-phrase has undergone movement in a cleft question, this movement is not itself *wh*-movement, i.e. movement-to-C. The *wh*-phrases in cleft and canonical questions are licensed in an analogous fashion.

It is worth noting that the presence of intervention effects in turn give support for the structure proposed in (59), where the head C c-commands Foc. Focus-alternatives are taken to propagate up the derivation (Beck, 2006; Kratzer and Shimoyama, 2002), so an intervention problem occurs between the operator that interprets these alternatives, namely C, and the source of the alternatives, the *wh*-expression, only if the both the *wh* and the intervener are *below* C.

5.3 Explaining the grammaticality of long-distance cleft questions

I now turn to the question of interest: why are in-situ embedded *wh*-phrases licensed in cleft questions, when they are not in canonical *wh*-questions? The relevant contrast between canonical ques-

tions (63) and cleft questions (64) is repeated from above:

(63) *[Sita **eethu pustakam** vaayich-ennu] Raman vicaarichu?
[Sita which book read-that] Raman thought
Intended: 'Which book did Raman think Sita read?'

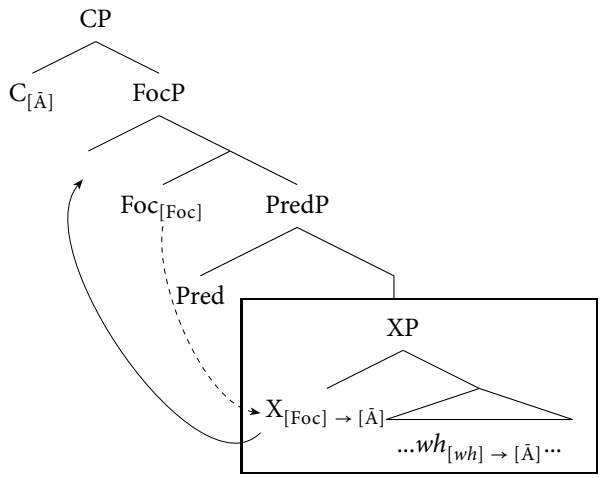
(64) [Sita **eethu pustakam** vaayichu ennu] aane Raman vicaarich-athe?
[Sita which book read that] COP Raman thought-NOMNL
'Which book was it that Raman thought Sita read?'

As I suggested above, the crucial difference between the two configurations has to do with the relative order of the probes and thus the order of operations. In canonical questions, the head H responsible for clausal fronting is structurally higher than C. In clefts, the head Focus triggering clausal clefting is below C. Therefore, clefting takes place prior to *wh*-Agreement. Consequently, the potentially intervening [Ā] features on the clefted clauses have already been checked and deleted.

Let us consider the derivation in more detail. Since the entire embedded clause is to undergo clefting, the head of this clause, Sub, would bear the requisite [Foc]-feature. The features on the Focus head probe first, agreeing with and attracting the entire embedded clause to its specifier. Conditions on Activity demand that the features on Sub are now deleted and no longer accessible for further Agree relations. Assuming Cyclic Agree as we did in the previous section, the embedded C would bear [*wh*] and, by entailment, [Ā] features. By the time the matrix C is merged, these are the only relevant *active* features in the derivations and *wh*-Agreement can freely take place. The derivation is schematized in (65):

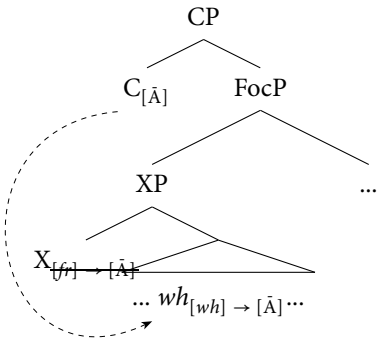
(65) a. STEP 1:

Foc Agrees w/+Attracts XP



b. STEP 2:

C Agrees with *wh*



The contrast in acceptability between canonical questions and cleft-questions is mysterious under approaches that take the scope of *wh*-in-situ in certain languages to be clause-bound. We see

in the case of Malayalam that the same finite clause may be selectively transparent for *wh*-scope depending on the configuration. Because restricted-scope, under the present account, is the result of a particular intervention configuration, it predicts the selective availability of long-distance *wh*-in-situ in clefts configurations.

6 Beyond Malayalam

This paper has argued that the claimed correlation of finiteness with specially restricted scope for embedded wh-in-situ is false for Malayalam, and offered an entirely different account of the relevant phenomena. Beyond Malayalam, however, apparent finite-clause-boundedness of *wh*-in-situ has been observed in a number of unrelated languages. In this section, I discuss two languages that are well-known to have this property — Hindi and Iraqi Arabic. I will suggest that in both of these languages, the restricted behavior of in-situ *wh*-phrases can be seen as epiphenomenal, but whereas Hindi is essentially the mirror image of Malayalam, properties of *wh*-in-situ in Iraqi Arabic may not represent the same phenomenon at all.

I turn first to Hindi and show that matrix scope of *wh*-in-situ in this language is blocked by the movement of the complement clause, which happens to be *rightward*. Iraqi Arabic, though often discussed alongside Hindi as being a case of restricted *wh*-scope, is different from both Hindi and Malayalam in that clauses in Iraqi Arabic are opaque not just for scope of *wh*-in-situ, but also for overt extraction. I will suggest that in this language, finite clauses are in fact be islands. A closer inspection of the data from the two languages suggests that a variety of language-specific factors affecting embedded clauses may give rise to the the appearance of restricted scope for embedded *wh*-expressions. It is possible that once we can understand and give an adequate analysis for embedded clauses in these restricted-scope languages, the need for a radically different syntactic analysis for *wh*-in-situ in these languages is eliminated. I close this section with a brief discussion of unrestricted scope languages like Japanese, which differ from Malayalam and Hindi in that licensing of *wh*-in-situ does not seem to interact with operations targetting embedded clauses. I will suggest that we can understand this pattern of variation if the feature-specification on interrogative probes is different across these two classes of languages.

6.1 Hindi

Mahajan (1987) observed that *wh*-in-situ inside embedded finite clauses in Hindi cannot take matrix scope; this is illustrated by the ungrammaticality of (66).

(66) *tum soch-te ho [ki **kaun** aa-egaa]?
you.PL think-HAB.MPL be.PRS.2PL [that who come-FUT.3MSG]
'Who do you think will come?' (Dayal 1996)

It is also well-known that *wh*-scope in Hindi correlates with another phenomenon that takes place in the language, namely the obligatory post-verbal positioning of finite clausal complements. Though Hindi an SOV language, finite clausal complements must appear post-verbally, resulting in an SVO order for bi-clausal sentences.[18] Based on evidence from non-finite clauses, which can optionally appear post-verbally, Dayal (1996) argues that clause-position is the determining factor for *wh*-scope. As shown in (67), embedded *wh*-phrases fail to take matrix scope when the container clause is post-verbal, even when the clause is non-finite.

(67) a. **Preverbal complement clause: wide-scope possible**

tum [**kyaa** paRhnaa] caahte ho
you [what read.INF] want PR
'What do you want to read?'

b. **Postverbal complement clause: wide-scope blocked**

*tum caahte ho [**kyaa** paRhnaa]
you want PR [what read.INF]
Intended: 'What do you want to read?' (Dayal 1996)

Dayal (1996) argues that the scope-islandhood of post-verbal clauses is due to the fact that they are base-generated adjuncts and therefore islands for covert *wh*-movement. Such an analysis faces a number of challenges, chief among which is the fact that post-verbal clauses are generally not barriers for overt extraction (Bayer, 1997; Mahajan, 1997). Furthermore, evidence from binding phenomena suggests that the post-verbal position of clauses is derived by rightward movement, rather than being base-generated. Post-verbal clauses can contain elements bound by a matrix element (68a) and display Principle C effects (68b).[19]

[18] Like clausal fronting in Malayalam, the post-verbal clause position does not have obvious semantic or information-structural correlates.

[19] A number of authors have attempted to account for the binding facts by adopting an anti-symmetric approach (Mahajan, 1997; Simpson and Bhattacharya, 2003; Simpson and Choudhury, 2015). These authors take SVO to be the "default" word-order and argue that nominals get to a pre-verbal position via leftward movement, likely for Case reasons. I will not adopt this analysis here for a number of reasons. First, such an approach fails to explain the optional post-verbal positioning of non-finite clauses. Second, anti-symmetric approaches cannot explain the correlation between clause position and *wh*-scope and take them to be spurious (Simson and Choudhury 2015). If patterns of restricted *wh*-scope across languages reflect the same underlying phenomenon, then the Malayalam data we saw in previous sections provide a compelling argument against such a view. The Malayalam fronted clauses we examined here do not appear in a post-verbal position to begin with and therefore the clausal movement patterns cannot be explained by resorting to

(68) a. Raam-ne har laRke$_i$-se kahaa [ki vo$_i$ jiitegaa]
Raam-ERG every boy-SOC said [tht he win.FUT]
'Raam told every boy$_i$ that he$_i$ will win.'

b. *Raam-ne us$_i$-se kahaa [ki Mohan$_i$ jiitegaa]
Raam-ERG he-SOC said [that Mohan win.FUT]
'Raam told him$_i$ that Mohan$_i$ will win.' (Kidwai 2013)

Principle C effects can be taken to be a diagnostic of Ā-movement, in which case clausal extraposition in Hindi can be thought of as an Ā-operation, much like clause fronting in Malayalam (see also Büring and Hartmann 1997; Müller 1997. The account we proposed for Malayalam can then be extended straightforwardly to Hindi, with the minimal modification that the head triggering rightward movement has a right-branching specifier. Suppose the interrogative C in Hindi is also a flat probe looking for [Ā]-features. For concreteness, let us also say that clausal extraposition is driven by the feature [EX] on a left-peripheral head (for concreteness, I will refer to it once again as H). Clauses that undergo movement will bear [EX], which in turn entails [Ā]. The [Ā]-features on the clause would intervene for Agree between the matrix C and the clause-internal *wh*, thus blocking the syntactic dependency necessary for *wh*-licensing.

The puzzle in Hindi is complicated by the fact that wide-scope for *wh*-phrases can be obtained by overtly fronting the *wh*-phrase, as in (69).

(69) **kaun**$_i$ tum soch-te ho [ki t$_i$ aa-egaa]?
who you.PL think-HAB.MPL be.PRS.2PL [that come-FUT.3MSG]
'Who do you think will come?' (Dayal 1996)

Mahajan (1987) and Dayal (1996), however, have presented a number of arguments for treating this instance of *wh*-fronting as an instance of long-distance scrambling, rather than overt *wh*-movement. Fronted *wh*-phrases can be preceded by other scrambled elements. Moreover, the position of moved *wh*-phrases in relation to the complementizer, at least in Hindi, is *below* C, suggesting that these elements are not moving to Spec, CP, where *wh*-moved elements are typically assumed to go.

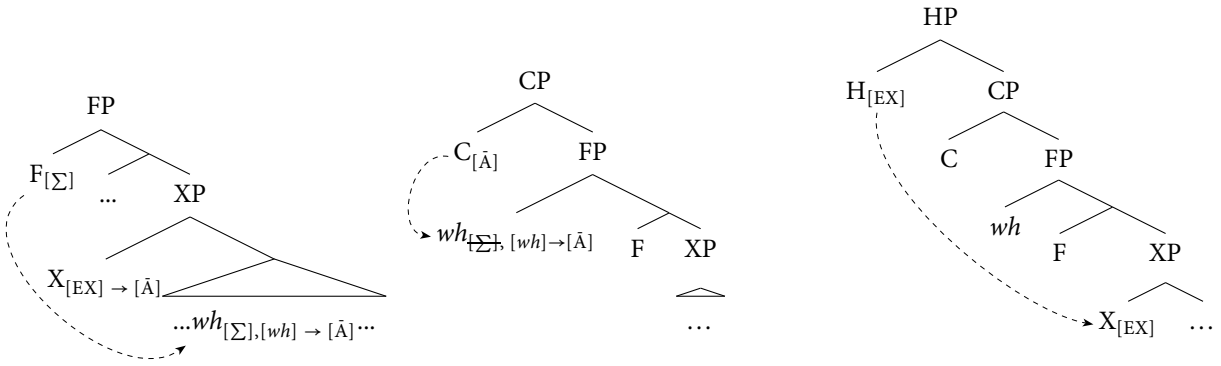
(70) Ravi jaan-taa hai [ki kis-ko$_i$ tum maan-tii ho [ki
Ravi know-HAB.MSG be.PRS.SG [that who-ACC you believe.HAB.F be.PRES.2PL [that
Billu-ne t$_i$ maar-aa]]
Billu-ERG hit-PFV]]
'Ravi knows who you think that Billu hit.' (Mahajan 1987)

anti-symmetry. On the other hand, an overt movement approach can capture the patterns in both Hindi and Malayalam in a uniform fashion. For further arguments against anti-symmetric approaches, I refer the reader to Bhatt and Dayal (2007).

I will argue that *wh*-scrambling is one strategy available in Hindi to circumvent intervention. The logic is fully parallel to clefting in Malayalam. The probe responsible for scrambling, which is lower than interrogative C, can trigger movement of *wh*-phrase out of the c-command domain of the clause-head bearing [Ā]-features and closer to interrogative C. Thus, when interrogative C probes, the *wh*-phrase is the first [Ā]-feature-bearing element it encounters.

I follow Grewendorf and Sabel (1999); Miyagawa (1997); Sauerland (1996), and Müller (2000) in taking scrambling to be feature-driven movement (driven by a feature [∑]) to the specifier of a functional head below C. Given that long-distance scrambling in Hindi has Ā-properties (see e.g. Mahajan 1990), I will assume that [∑] entails [Ā]. In scrambling configurations, the first relevant element to probe will be the functional head seeking [∑]-features. The *wh*-phrase bearing [∑]-features will be attracted to the specifier of this head, with the result that it will now be closer to the interrogative C than the Ā-features on the embedded clause. The head of the embedded clause is no longer in a position to interact with the flat probe on C. This is schematized in (71).

(71) **1: Scrambling** **2: C-*wh* Agreement** **3: Extraposition**



Thus, long-distance question formation in Hindi can straightforwardly be given an analysis where the limited scope of *wh*-in-situ is a side-effect of how interrogative C interacts with [Ā]-features on the embedded clause. The account also succeeds in capturing the differences in scope-possibilities between overt scrambling and in-situ licensing, a difference which, on the present view, has to do with the relative positioning of the two probes at work.

6.2 Iraqi Arabic

Iraqi Arabic is frequently discussed alongside Hindi as having clause-bound scope for *wh*-in-situ. The previous section showed that the clause-boundedness effects in Hindi can be thought of as being the result of how licensing of embedded *wh*-phrases interacts with properties of the embedded clause itself. It is more challenging to demonstrate for Iraqi Arabic that the restricted scope effects are also the result of language-specific constraints that are independent of *wh*-licensing. Relevant

data on Iraqi Arabic is sparse, with much of the literature relying on a single source (Wahba, 1991), whose examples are consistent with a number of different analyses. The proposal that follows should therefore be taken as tentative. My goal in this section is to show that restrictions on long-distance wh-in-situ in Iraqi Arabic can, in principle, be given an explanation that does not involve positing language-specific locality constraints on *wh*-in-situ.

The *wh*-patterns in Iraqi Arabic, as first described in Wahba (1984), are as follows. Monoclausal questions may be formed using an in-situ strategy or overt movement.[20]

(72) a. Mona shaafat meno?
Mona saw who
'Who did Mona see?'

b. Meno Mona shaafat?
Who Mona saw
'Who did Mona see?'

Long-distance *wh*-in-situ is possible with non-finite clauses, as illustrated in (73). However, long-distance question formation is more restricted when the embedded clause in question is finite: *wh*-in-situ can only have strictly local scope when embedded inside a finite complement, as illustrated by the impossibility of (74).

(73) Mona hawlat tishtiri sheno?
Mona tried to-buy what
'What did Mona try t buy?'

(74) *Mona tsawwarat [Ali ishtara sheno] ?
Mona thought [Ali bought what]
'What did Mona think Ali bought?'

Unlike *wh*-in-situ, however, overt long-distance movement of the *wh*-phrase seems possible and gives wide scope to the embedded *wh*-phrase, as we see in (75).

(75) Sheno tsawwarat Mona [Ali ishtara] ?
what thought Mona [Ali bought]
'What did Mona think Ali bought?'

This apparent asymmetry between *wh*-in-situ and *wh*-movement has been the primary argument for claiming that Iraqi Arabic varies parametrically from languages like English in imposing special locality restrictions on *wh*-phrases. For instance, Ouhalla (1996) argued that *wh*-phrases in Iraqi Arabic are anaphors that need to be in a sufficiently local relationship with its antecedent, which he takes to be the interrogative C. While this would account for the difference between *wh*-in-situ and *wh*-movement in (74) and (75) respectively, it fails to account for another crucial property of Iraqi

[20] As pointed out in Ouhalla (1996), Iraqi Arabic does not allow scrambling, which suggests that this is genuine *wh*-movement.

Arabic, namely that overt long-distance movement of non-nominal *wh*-expressions is also blocked (76).

(76) *leesh tsawwarit Mona [Ali masha] ?
why thought Mona [Ali left]
'Why did Mona think Ali left?'

Thus, restrictions on *wh*-question formation in Iraqi Arabic are not limited to embedded *wh*-in-situ. A full account of Iraqi Arabic long-distance question-formation needs to provide explanations for both (i) the apparent ban on long-distance *wh*-in-situ and (ii) the asymmetry between nominal and non-nominal *wh*-expressions when it comes to overt *wh*-fronting.

The idea that I will pursue here builds on insights in Basilico (1998) and takes all Iraqi Arabic finite clauses to be islands. Iraqi Arabic prohibits *wh*-in-situ within islands generally, as illustrated by the ungrammaticality of in-situ *wh*-phrases inside *wh*-islands (77a) and relative clauses (77b).

(77) a. * Mona nasat li-meno tinti sheno?
Mona forgot to-whom to give what
'What did Mona forget to give to whom?'

b. * Mona ʿurfit il-bint illu ishtarat sheno?
Mona know the-girl who bought what
'What did Mona know the girl who bought?'

If embedded clauses are also islands, then the restriction on long-distance *wh*-in-situ follows straightforwardly. The impossibility of overt long-distance extraction of non-nominal *wh*-expressions also falls out from the islandhood of finite embedded clauses, as we expect that they are opaque for all Ā-extraction operations.

The explanandum, on the view that finite complement clauses are islands, is the apparent *availability* of overt movement for nominal *wh*-expressions from embedded positions. One possible line of analysis is that what looks like long-movement is not movement at all, but base-generation of a null resumptive pronoun in the embedded position.[21] In fact, this pronoun can sometimes be overt in Iraqi Arabic, as shown in (78), and is obligatorily overt in the same position in Egyptian Arabic, a closely related language which shows similar *wh*-in-situ restrictions (Wahba 1984).

(78) Suha minnu: ta?atagid [ra:H ya?zim **hu** Ahmad] ?
Suha who think.3FS [will invite.3MS 3MS Ahmad]
'Who does Suha think that Ahmad will invite him?'

Association with a resumptive pronoun seems to be a strategy that is generally available in Iraqi Arabic where extraction is blocked. Thus, resumption is possible with other types of islands, as we

[21] See Georgopoulos (1985) for a similar analysis for Palauan *wh*-questions.

see with the *wh*-island in (79) and the adjunct island in (80).[22]

(79) minnu ytsasa:ʔil Ragheb le:S **hyi** ba:sit Behjet?
who wonder.3MS Ragheb why 3FS kissed.3FS Behjet
'Who is Ragheb wondering why she kissed Behjet at the party?'

(80) minnu ga:l Samer li-Ragheb Sw:aget **hyi** ra:H timSi: li-Baghdad?
who said.3MS Samer to-Radheb when 3FS will go.3FS to-Baghdad
'Who did Samer say to Ragheb when she will go to Baghdad?'

(Sterian 2011)

The resumption strategy is unavailable for non-nominal *wh*-phrases because Iraqi Arabic simply does not have non-nominal pronominals (Sterian, 2011).

Because Iraqi Arabic imposes restrictions not just on long-distance *wh*-in-situ, but also certain kinds of *wh*-movement, I suggested that finite embedded clauses are islands in this language. What looks like overt extraction is base-generation and association with a resumptive pronoun. On this view, long-distance questions involving nominal versus non-nominal *wh*-phrases do not differ with respect to the grammaticality of extraction — *wh*-extraction is impossible in both cases. Rather, the two types of questions vary with respect to the availability of the right type of pronominal associate, which in turn has an effect on whether the resumption strategy is possible at all for long-distance questions.

I have attempted to show that restrictions on long-distance *wh*-in-situ in Iraqi Arabic can, in principle, be given an explanation that is consistent with maintaining that the scope of *wh*-in-situ is not necessarily clause-bound. It makes a number of predictions, testing which is presently difficult due to the lack of crucial data. First, I have suggested that finite complement clauses in the language are islands for Ā-operations. The impossibility of non-nominal *wh*-extraction provides preliminary support for this hypothesis. However, we need to show that this is a general property of the embedded clause domain and not specific to *wh*-question-formation. Thus, we need to show that similar patterns obtain with other Ā-phenomena, like topicalization and relativization. Second, I have suggested that what looks like long-distance movement actually involves a resumption strategy. It is important to test this claim using diagnostics for movement (or lackthereof) like idiomatic meaning retention across the clause-boundary. Finally, my suggestion that the language can sometimes employ null resumption as a strategy for forming questions across islands calls for a more thorough investigation into the distribution of null and overt resumptive pronouns in Iraqi Arabic.

[22] The story is complicated by the fact that overt resumptives are required in these environments. Null resumptives, as argued to be possible inside finite complement clauses, are not available. Thus, a full picture of Iraqi Arabic extraction and resumption strategies for *wh*-question formation needs to be supplemented with an account of the distribution of null resumptives.

Though many relevant aspects of Iraqi Arabic *wh*-question formation remain open questions at this stage, I believe that the account here is promising in a number of ways. It is, to begin with, more parsimonious in that it does not necessitate the positing of language-specific constraints on *wh*-licensing. Second, it offers a potential explanation for the otherwise puzzling asymmetry between nominal and non-nominal *wh*-extraction.

6.3 Unrestricted scope languages

We began with the observation that not all in-situ languages show similar restrictions on *wh*-in-situ. In languages like Japanese and Korean, an embedded *wh*-phrase can take unbounded scope while remaining in its base position. A Japanese example is provided in ((81), repeated from §1).

(81) Hideki-ga [Kyoko-ga **nani-o** ka-tta to] i-tta no
Hideki-NOM [Kyoko-NOM what-ACC buy-PAST that] say-PAST Q
'What did Hideki say Kyoko bought?'

There are two ways in which languages like Japanese differ from Malayalam. First, clauses of all sorts in these languages are able to stay in their base position without issue, as we see in (81). This suggests that the prosodic constraints at work here are of a different sort than in Malayalam. More importantly, we find that even when the embedded clause appears in a sentence-initial position, as in (82), the embedded *wh*-phrase is able to take matrix scope.

(82) [Kyoko-ga **nani-o** ka-tta to] Hideki-ga it-ta no
[Kyoko-NOM what-ACC buy-PAST that] Hideki-NOM say-PAST Q
'What did Hideki say Kyoko bought?'

There are a number of factors that could be responsible for this contrast. One possibility is that while the clausal fronting operation in Japanese has the same properties as in Malayalam, the feature specifications on interrogative C differ. Japanese interrogative C, unlike its Malayalam counterpart, could be a maximally specified probe that only interacts with elements bearing the feature [*wh*]. Alternatively, the interrogative C might be a generalized probe, but the fronting operation is of a different sort, such that the head of the clause is never a potential intervener to begin with.

It is also possible that the fronted CP in Japanese has undergone short-scrambling, which has been argued to have A-properties in Japanese Tada (1993). If this is the case, it is difficult to adjudicate between the two alternatives posited above when there is only one level of embedding. Long-distance movement, on the other hand, is taken to be Ā-movement, even if it is a scrambling operation. Thus, if a CP undergoes long movement, we expect the head of that clause to bear [Ā]-features. Such examples can therefore serve as our testing ground.

We see in (83) that long-distance fronting of CPs is generally acceptable in Japanese.

(83) [Kyoko-ga kukki-o tabe-ta to] Hideki-wa [Taro-ga t_i omot-teiru to]
[Kyoko-NOM cookie-ACC eat-PAST that] Hideki-TOP [Taro-NOM think-PROG that]
it-ta
say-PAST
'Kyoko said that Hideki thinks that Taro ate the cookies.'

When the fronted clause contains a *wh*-phrase, the resulting sentence, though somewhat degraded, is crucially not ungrammatical. This suggests that even when the *wh*-container clause bears [Ā]-features of its own, in-situ *wh*-licensing is possible in Japanese.

(84) ?[Kyoko-ga nani-o tabe-ta to] Hideki-wa [Taro-ga t_i omot-teiru to]
[Kyoko-NOM what-ACC eat-PAST that] Hideki-TOP [Taro-NOM think-PROG that]
it-ta no
say-PAST Q
'What did Kyoko say that Hideki thinks that Taro ate?'

This important difference between Japanese and Malayalam can be captured straightforwardly within the Relativized Probing framework adopted here. The appeal of Relativized Probing is that the characteristics (i.e. the feature-structure) of goals can remain constant across languages, with differences in syntactic behavior, when they occur, can be explained by variability in the specificity of the probe itself. The differences between Japanese and Malayalam can be understood along the same lines. If interrogative C in Japanese is equipped with a probe seeking [*wh*]-features this is the case, the [Ā]-features on the fronting clause would be incapable of interacting with it, as it does not possess features relevant for C.

7 Conclusion

This paper examined apparent restrictions on *wh*-in-situ and argued that they are not inherent to *wh*-question-formation, but rather, consequences of how *wh*-licensing in a given language interacts with other language-specific operations. We considered supporting evidence from Malayalam, a language previously claimed to display clause-boundedness of *wh*-in-situ. I argued that the apparent clause-boundedness of *wh*-in-situ is not clause-boundedness at all, but the result of a pernicious interaction between features relevant for *wh*-licensing and those necessary for the fronting of the *wh*-containing clause. Specifically, I showed that Agree between the interrogative C and an embedded *wh*-element is blocked when (*i*) the interrogative C is a generalized [Ā]-probe and (*ii*) other [Ā]-feature-bearing heads occur in between C and the *wh*-phrase.

Malayalam thus provides evidence that *wh*-in-situ across languages is less diverse than initially thought. This is not to say that the account predicts no variability. The relativized probing approach

advanced in this paper predicts three ways in which languages could vary with respect to how *wh*-licensing interacts with other Ā-operations. We might find a language in which the interrogative C is a flat probe, as in Malayalam, in which case we expect it to interact with any given feature in the relevant class. We might have on C an intermediate-level probe, which interacts with a proper subset of the features in the relevant class. Also possible is a maximally-specified probe, which can "skip over" features that are not point-by-point identical to it. Crucially, the locus of parametrization is on the featural make-up of probes, as opposed to locality conditions inherent to the syntax of *wh*-in-situ. Not only is such an approach independently motivated (see e.g. Béjar and Rezac 2009; Preminger 2012), it is also more adequate on empirical grounds. Locality constraints on *wh*-in-situ are the very same as those on the operation Agree, which licenses *wh*-dependencies. Given that Agree is the fundamental structure-building operation in natural language, we expect locality conditions on the operation not to be subject to cross-linguistic variation.

References

Abels, Klaus. 2012. The Italian left periphery: A view from locality. *Linguistic Inquiry* 43: 229–254.

Adger, David, and Gillian Ramchand. 2005. Merge and move: *Wh*-dependencies revisited. *Linguistic Inquiry* 36 (2): 161–193.

Baker, C. L. 1970. Double negatives. *Linguistic Inquiry* 1 (2): 169–186.

Basilico, David. 1998. Wh-movement in Iraqi Arabic and Slave. *The Linguistic Review* 15 (4).

Bayer, Josef. 1997. CP extraposition as argument shift. In *Rightward movement*, eds. Dorothee Beerman, David LeBlanc, and Henk van Riemsdijk, 37–58. Amsterdam: John Benjamins Publishing Company.

Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14 (1): 1–56.

Béjar, Susana, and Milan Rezac. 2009. Cyclic agree. *Linguistic Inquiry* 40 (1): 35–73.

Bhatt, Rajesh, and Veneeta Dayal. 2007. Rightward scrambling as rightward remnant movement. *Linguistic Inquiry* 38 (2): 287–301.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59 (1): 1–45.

Bošković, Željko. 2014. Now i'm a phase, now i'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45: 27–89.

Brody, Michael. 1990. Some remarks on the focus field in Hungarian. In *University college london working papers in linguistics*, Vol. 2, 201–225. UCL.

Büring, Daniel, and Katharina Hartmann. 1997. The Kayne mutiny. In *Rightward movement*, eds. Dorothee Beerman, David LeBlanc, and Henk van Riemsdijk, 59–79. Amsterdam: John Benjamins Publishing Company.

Cable, Seth. 2010. *The grammar of Q: Q-particles, Wh-Movement and Pied-Piping*. Oxford University Press.

Chomsky, Noam. 1977. On wh-movement. In *Formal syntax*, eds. P. Culicover, T. Wasow, and A. Akmajian. New York, New York: Academic Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, eds. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in linguistics*, ed. Michael Kenstowicz, 1–52. Cambridge, Massachusetts: MIT Press.

Dayal, Veneeta. 1996. *Locality in Wh-quantification: Questions and relative clauses in Hindi*. Dordrecht: Kluwer Academic Press.

Dryer, Matthew S. 1991. SVO languages and the OV/VO typology. *Journal of Linguistics* 27: 443–482.

Engdahl, E. 1983. Parasitic gaps. *Linguistics and Philosophy* 6 (1): 5–34.

Fanselow, Gisbert. 2007. The restricted access of information structure to syntax: A minority report. In *The notions of information structure*, eds. Caroline Féry, Gisbert Fanselow, and Manfred Krifka. *Working papers of sfb632: Interdisciplinary studies on information structure*. Potsdam: Universitatsverlag Potsdam.

Fery, Caroline. 2009. Indian languages as intonational 'phrase languages'.

Féry, Caroline. 2011. German sentence accents and embedded prosodic phrases. *Lingua* 121: 1906–1922.

Frascarelli, Mara, and Francesca Ramaglia. 2013. Agreement, information structure and CP, ed. Camacho-Taboada et al. Benjamins.

Georgopoulos, Carol. 1985. Variables in Palauan syntax. *Natural Language and Linguistic Theory* 3 (1): 59–94.

Grewendorf, Günther, and Joachim Sabel. 1999. Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language and Linguistic Theory* 17 (1): 1–65.

Hagstrom, Paul. 1998. Decomposing questions. PhD diss, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Haida, Andreas. 2007. Reconsidering the indirect dependency approach to wh-scope marking. In *Proceedings of SALT XVII*, eds. T. Friedman and M. Gibson, Vol. 16, 130–147. Ithaca, NY: Cornell University.

Hamblin, Charles. 1973. Questions in Montague grammar. *Foundations of Language* 10 (1): 41–53.

Hany Babu, M. T. 1997. The syntax of functional categories. PhD diss, Central Institute of English and Foreign Languages.

Harley, Heidi, and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78 (3): 482–526.

Hartmann, Katharina. 2013. Rightward movement in a comparative perspective, eds. G. Webelhuth, M. Sailer, and H. Walker. John Benjamins.

Hayes, Bruce. 1990. Precompiled phrasal phonology. In *The phonology-syntax connection*, eds. Sharon Inkelas and Draga Zec, 85–108. Chicago: University of Chicago Press.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.

Huang, C. T. James. 1982. Logical relations in Chinese and the theory of grammar. PhD diss, Massachusetts Institute of Technology.

Jayaseelan, K. A. 2001. Questions and question-word incorporating quantifiers in Malayalam. *Syntax* 4 (2): 63–93.

Kiss, Katalin E. 1998. Identificational focus versus information focus. *Language* 74: 245–273.

Kotek, Hadas. 2014. Composing questions. PhD diss, Massachusetts Institute of Technology.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In *Proceedings of the Tokyo conference on psycholinguistics*, ed. Yukio Otsu, Vol. 3, 1–25. Tokyo: Hituzi Syobo.

Lebeaux, David. 1988. Language acquisition and the form of the grammar. PhD diss, University of Massachusetts, Amherst.

Legate, Julie Anne. 2005. Phases and cyclic agreement. In *Perspectives on phases*, eds. Martha McGinnis and Norvin Richards. MITWPL.

Lipták, Anikó. 2001. On the syntax of wh-items in Hungarian. PhD diss, University of Leiden.

Madhavan, Punnapurtath. 2013. Cleft structures, ed. Tonjes Hartmann Katharina; Veenstra. John Benjamins. Chap. 10.

Mahajan, Anoop. 1990. The A/A-bar distinction and movement theory. PhD diss, Massachusetts Institute of Technology.

Mahajan, Anoop. 1997. Rightward scrambling. In *Rightward movement*, eds. Dorothee Beerman, David LeBlanc, and Henk van Riemsdijk, 185–213. Amsterdam: John Benjamins Publishing Company.

Mahajan, Anoop K. 1987. Notes on wh-questions in Hindi. Paper presented at Sala, Cornell and Syracuse Universities.

Manetta, Emily. 2012. Reconsidering rightward scrambling: Postverbal constituents in Hindi-Urdu. *Syntax* 43 (1): 43–74.

Mathew, Rosmin. 2014. The syntactic effect of head movement: Wh and verb movement in Malayalam. PhD diss, University of Tromso.

McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and explanation in the minimalist program*, eds. Samuel David Epstein and T. Daniel Seely, 184–226. Malden, Massachusetts: Blackwell Publishing.

Menon, Mythili. 2011. Revisiting infinitives: The need for tense in Malayalam. In *Finiteness in south asian languages*. Norway. University of Tromso. University of Tromso.

Miyagawa, Shigeru. 1997. Against optional scrambling. *Linguistic Inquiry* 28 (1): 1–25.

Mohanan, K. P. 1982. Infinitival subjects, government, and the abstract case. *Linguistic Inquiry* 13 (2): 323–327.

Müller, Gereon. 1997. Extraposition as remnant movement. In *Rightward movement*, eds. Dorothee Beerman, David LeBlanc, and Henk van Riemsdijk, 215–246. Amsterdam: John Benjamins Publishing Company.

Müller, Gereon. 2000. Shape conservation and remnant movement. In *Proceedings of the north east linguistic society*, eds. Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-yung Kim, 525–540. University of Massachusetts at Amherst: Graduate Linguistic Student Association.

Müller, Gereon, and Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24 (3): 461–508.

Munataka, Takashi. 2006. Intermediate Agree: Complementizer as a bridge. In *Proceedings of the 3rd wafl*.

Nakamura, Masanori. 1998. Reference set, Minimal Link Condition, and parameterization. In *Is the best good enough?*, eds. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, 291–314. Cambridge, Massachusetts: MIT Press.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht, Holland ; Riverton, N.J.: Foris.

Nissenbaum, Jon. 2000. Covert movement and parasitic gaps. In *Proceedings of the north east linguistic society*, eds. Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-yung Kim, 541–556. Rutgers University: Graduate Linguistic Student Association.

Ouhalla, Jamal. 1996. Remarks on the binding properties of wh-pronouns. *Linguistic Inquiry* 27 (4): 676–707.

Percus, Orin. 1997. Prying open the cleft. In *Proceedings of north east linguistic society 27*, ed. Kiyomi Kusumoto, 337–351. McGill University: Graduate Linguistic Student Association.

Pesetsky, David. 1982. Paths and categories. PhD diss, Massachusetts Institute of Technology.

Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge: MIT Press.

Potsdam, Eric, and Daniel Edmiston. 2016. Extraposition in Malagasy. In *Proceedings of the 22nd meeting of the austronesian formal linguistics association*.

Preminger, Omer. 2012. Agreement as a fallible operation. PhD diss, Massachusetts Institute of Technology.

Proceedings of WECOL. 2007. Extraposition as PF movement.

Reinhart, Tanya. 1998. Wh-in-situ in the framework of the Minimalist Program. *Natural Language Semantics* 6 (1): 29–56.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Massachusetts: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.

Rizzi, Luigi. 2004. Locality and left periphery. In *Structures and beyond: The cartography of syntactic structures*, ed. Adriana Belletti, Vol. 3, 223–251. Oxford: Oxford University Press.

Rooth, Mats. 1985. Association with focus. PhD diss, University of Massachusetts, Amherst.

Sauerland, Uli. 1996. The interpretability of scrambling. In *Formal approaches to japanese linguistics*, eds. Masatoshi Koizumi, Masayuki Oishi, and Uli Sauerland, Vol. 2.

Selkirk, Elisabeth. 2011. The syntax-phonology interface, eds. J. A. Goldsmith, J. Riggie, and A Yu. Blackwell Publishers.

Selkirk, Elisabeth O. 1984. *Phonology and syntax : the relation between sound and structure*. Cambridge, Mass.: MIT Press.

Simpson, Andrew. 2000. *Wh-movement and the theory of feature-checking*. Amsterdam: John Benjamins Publishing Company.

Simpson, Andrew, and Tanmoy Bhattacharya. 2003. Obligatory overt wh-movement in a wh-in-situ language. *Linguistic Inquiry* 34 (1): 127–142.

Simpson, Andrew, and Arunima Choudhury. 2015. The nonuniform syntax of postverbal elements in SOV languages: Hindi, Bangla, and the rightward scrambling debate. *Linguistic Inquiry* 46 (3): 533–551.

Srikumar, K. 1992. Question-word movement in Malayalam and GB theory. PhD diss, Osmania University, Hyderabad.

Srikumar, K. 2007. Clausal pied-piping and subjacency. In *Linguistic theory and south asian languages/essays in honour of k.a. jayaseelan*, eds. Joseph Bayer, Tanmoy Bhattacharya, and MT Hany Babu, 53–69. John Benjamins Publishing Company.

Starke, Michal. 2001. Move dissolves into merge: A theory of locality. PhD diss, University of Geneva.

Sterian, Laura. 2011. The syntax and semantics of gap and resumptive strategies in Iraqi Arabic. PhD diss, University of British Columbia.

Swenson, Amanda, Athulya Aravind, and Paul Marty. 2015. The expression of information structure in Malayalam.

Szabolcsi, Anna. 1981. Compositionality in focus. *Acta Linguistica Societatis Linguistice Europaeae*.

Tada, Hiroaki. 1993. A/A' partition in derivation. PhD diss, Massachusetts Institute of Technology.

Truckenbrodt, Hubert. 1995. Extraposition from np and prosodic structure. In *Proceedings of the north east linguistic society 25*, ed. Jill Beckman, 503–518. University of Pennsylvania: Graduate Linguistic Student Association.

Wahba, Wafaa. 1984. Wh-constructions in Egyptian Arabic. PhD diss, University of Illinois at Urbana-Champaign.

Wahba, Wafaa. 1991. Lf movement in Iraqi Arabic. In *Logical structure and linguistic structure: Cross-linguistic perspectives*, ed. C. T. James Huang, 253–276. Dordrecht.

Watanabe, Akira. 2006. Wh-movement: Moving on, ed. Norbert Cheng Lisa Lai Shen; Corver. MIT Press.