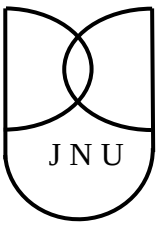
FOCUS IN MALAYALAM: AN ALTERNATIVE APPROACH

A dissertation submitted to Jawaharlal Nehru University in partial fulfillment of the requirements for the award of the degree of

Master of Philosophy

by

Hima S.



Centre for Linguistics
School of Language, Literature & Cultural Studies
Jawaharlal Nehru University
New Delhi-110067, India
2012

Centre for Linguistics
School of Language, Literature & Culture Studies
Jawaharlal Nehru University
New Delhi-110067, India

Dated: 27 July, 2012

CERTIFICATE

This thesis titled **"Focus in Malayalam: an Alternative Approach"** submitted by **Ms. Hima S.**, Centre for Linguistics, School of Language, Literature and Culture Studies, Jawaharlal Nehru University, New Delhi, for the award of the degree of **Master of Philosophy**, is an original work and has not been submitted so far in part or in full, for any other degree or diploma of any University or Institution.

This may be placed before the examiners for evaluation for the award of the degree of Master of Philosophy.

Dr. Ayesha Kidwai
Associate Professor
Centre for Linguistics
School of Language, Literature & Culture Studies
Jawaharlal Nehru University, New Delhi-110067

Dr. Ayesha Kidwai
(Supervisor)

Professor Vaishna Narang
Chairperson
Centre for Linguistics
School of Language, Literature and Culture Studies
Jawaharlal Nehru University
New Delhi-110067

Prof. Vaishna Narang
(Chairperson)

Tel: +91-11-26704199 Fax: +91-11-26741586

Dated: 27 July, 2012

DECLARATION BY THE CANDIDATE

This thesis titled **"Focus in Malayalam: an Alternative Approach"** submitted by me, for the award of the degree of **Master of Philosophy**, is an original work and has not been submitted so far in part or in full, for any other degree or diploma of any University or Institute.

Hima S.
M.Phil Student
Centre for Linguistics
SLL&CS
JNU

Acknowledgements

I would like to thank everyone who have directly or indirectly played a part in the 'coming into shape' of my dissertation.

First and foremost, my heartfelt thanks go to my Supervisor Dr. Ayesha Kidwai, for her untiring support and help. Next in line is of course Pritha, without whose constant encouragement, motivation and inspiration I could never even have dreamed about finishing this dissertation.

There always needs to be a support system of friends to sail through the difficult waters of dissertation writing. My love to Deepechi, Anindita, Hameeda and Siddik for their unconditional love. And of course, who can understand the blues better than people sailing on the same ship! Gayathiri, Jan and Ankush, ahoy there! Thanks to all others on similar ships too – especially Sujoy, Tresa and Jyotirupa.

I also wish to place on record my gratitude to all my informants, among them Deepechi, Sam, Amma, Ashok, Sudheesh, Tresa, Subin, Divya, Sawmiya, Padmam, Shilpa and Sarulakshmi especially. Also to be mentioned here are all those people whose name I have intentionally or otherwise omitted, who have played even a small part in this journey.

Of course, the support of one's own family cannot be replaced by anything else. Thank you, Amma, Achan and Sarath!

And finally my heartfelt gratitude to my constant source of joy, inspiration and support, my "friend, philosopher and guide", Sam.

,

CONTENTS

Abbreviations

1, 2, 3	First/ Second/ Third person
ACC	Accusative (Case)
DAT	Dative (Case)
F	Feminine (Gender)
FM	Focus Marker
M	Masculine (Gender)
N	Neuter (Gender)
NMLR	Nominalizer
OBJ	Object
OV	Object Verb (Word order)
P	Plural
PRES	Present (Tense)
PST	Past (Tense)
S	Singular
SOC	Sociative (Case)
SOV	Subject Object Verb (Word order)
SUB	Subject
VO	Verb Object (Word Order)

CHAPTER 1

INTRODUCTION

This dissertation aims to investigate into the syntactic manifestations of the phenomenon of Focus in the South Dravidian language Malayalam, from the current theoretical framework of generative syntax.

1.1 Defining focus

The term FOCUS has been used in linguistics to denote various phenomena. Hence it becomes imperative, before advancing any theory of focus, to define it. Focus can primarily be understood as an Information Structural or discourse notion that refers to a device that languages commonly use to indicate emphasis, or that the information being presented by the speaker is to be understood as being more important or newer in some sense, etc., compared to other parts of the utterance. It may also be explained, from a discourse point of view, as a device to mark what the speaker wishes to invite the attention of the hearer to, probably because she/he thinks the hearer is not likely to consider it to enter the common ground of the ongoing discourse. However, the most accepted definition of Focus seems to be the one from Alternative Semantics, proposed by Rooth (1985, et seq): "Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions."

It is also generally assumed that focusing divides the sentence into two parts, Focus and Background. This distinction is made based on the Givenness/Newness distinction. Thus if a constituent is new information compared to the rest of the sentence, it is the focus, and the rest of the sentence, which is 'familiar' or 'presupposed', is called the background.

There are different types of focus that are distinguished between, depending on its semantic and discursive functions. In the words of É. Kiss (1995),

> The term *focus* is used in generative linguistics in at least two different senses: it can denote the sentence part carrying new information, and it can also mean an operator expressing identification. The two meanings are often referred to as wide focus and narrow focus respectively. (É. Kiss (1995), 15)

Other differentiations made are in terms of Presentational vs. Contrastive focus, Identificational vs. Information focus, corrective focus, etc. Presentational focus is almost equivalent to information focus; it marks new (or non-presupposed) information. This is clear from the explanation Lee, Gordon & Büring (2008) give; for them "A commonly used diagnostic for presentational focus is questioning: the focus constituent is the part of the sentence that corresponds to the answer to a question, either overt or implied." It is usually used as opposed to Contrastive focus, which is used to pick out one member of a salient set as opposed to the others. Simlarly corrective focus is used for correction of information that was given in immediate previous discourse.

Interest in focus phenomena in generative syntax comes from the claims for an operator-like behaviour of focused elements. Focused constituents sometimes tend to move to a scope taking position, much like QR. Thus, possibilities like a separate functional projection Foc(us)P(hrase) has been posited in order to check the [+FOCUS] feature on the focused constituent. Also of interest is how the different types of focus are syntactically differentiated. In semantics, the truth conditional and quantificational effects of focus are sought to be explained. However, though semantic accounts usually assume that information structural notions like focus are syntactically represented in languages, they do not distinguish between the various types of focus.

Different languages use different methods, or their combinations, to mark focus: for instance by assigning prosodic stress, by marking it morphologically, by word order, etc. It is generally considered an interface phenomenon, since all modules of the language faculty seem to have some role to play in it. Grohmann (2009) argues that there are three ways in which languages mark focus syntactically, i.e., in situ, by movement or by clefting, and that languages generally use at most two of them. In most theories in (minimalist) syntax, focus is assumed to be syntactically associated or represented by a formal feature.

1.2 Theoretical framework

The theoretical framework that I adopt here is basically that set forth by Chomsky (1992, 1995, 1998, 1999, 2001).

The mind/brain is assumed to be modular, with a separate faculty of language that interacts with other cognitive systems (modules). Performance systems that access the language faculty are assumed to be external to the language faculty. The unitary and distinct systems called Articulatory-Perceptual (A-P) system and the Conceptual-Intentional (C-I) systems that relate to the sound and meaning part respectively of language are the only systems that access language, according to the Minimalist framework of Chomsky (1992, 1995, 1998, 1999, 2001). These systems interface with the language faculty through the interface levels of P(honetic) F(orm) (the A-P interface) and the L(ogical) F(orm) (the C-I interface). There is also a computational component C_{HL}, where the expressions of the language, which are essentially a pair (π, γ) of sound and meaning information, respectively. The generative procedure of grammar is assumed to be derivational. In Chomsky (1992, 1995, 1998) the PF and LF are two levels of representation, which weakens the assumption that the computational procedure is strictly derivational. The theory proposed in Chomsky (1999) however makes it stronger (to which we shall come back later).

The fundamental question or object of linguistic enquiry is to find out to what extent the human faculty of language FL, being a biological system, is an optimal solution to the minimal design specifications imposed by the other systems that may use it. Also known as 'legibility conditions' or bare output conditions, these are the constraints imposed by such systems on the FL such that the output of FL be legible or usable to them. These conditions require that nothing uninterpretable to the external systems within the derivation when it reaches the interface levels .

The constraints imposed on the C_{HL} are then the specifications of "good design" – something that is only empirically verifiable. The most optimal or perfect solution would be one where just the legibility conditions are satisfied. Thus it calls for formal and substantial minimalism, in the form of economy conditions. Thus, a derivation in one step is better than one in two steps, and having only two categories is better than four. If empirical evidence points towards any requirements that are "imperfect" in some sense, they should be accountable as being due to some independent external reasons.

The derivations generated by the language faculty are said to *converge* if it converges at both PF and LF. Even then, operations that form a step of a derivation are allowed only if the derivation would *crash* otherwise. Derivational steps are usually motivated by interface effects. Even in this, less 'costlier' (in terms of computing costs) steps are preferred over more 'costlier' ones. Thus simpler operations like Merge or Agree are preferred over costlier operations like Move.

These are the primary operations allowed in the C_{HL}. Universal Grammar (UG) is assumed to make available a universal set of features {F}, consisting of both uninterpretable and interpretable features. For instance, the Φ-features on DPs/NPs are interpretable features, whereas Φ-features on T, structural Case on DPs/ NPs and the EPP feature assigned to strong

phase heads etc are uninterpretable at the interfaces. For the sake of computational simplicity it is assumed that each language makes a one-time selection of features [F] from the universal set {F}, and assembles them into a Lexicon. The lexicon contains both substantive ('lexical') and functional items. The selection of the set [F] of features available in a language and their organization into the lexicon is assumed to be the locus of language variation, apart from parameter settings modifying the computational procedure.

Computation starts with a one-time selection of lexical items (LIs) from the Lexicon to form what is called a Numeration. The computational process maps this lexical array or Numeration onto LF and PF. The computational procedure is assumed to be uniform from the lexical array to the LF. However, at (a) certain stage(s), an operation Spell-Out separates information relevant to PF and 'ships it off' to PF.

Computational load is also reduced if only chunks of data have to be processed at a time. Chomsky (1998, 2001) propose that the numeration is further arranged into lexical sub-arrays which will used to build up one propositional unit at a time, which is called a *phase.* The core functional categories (CFCs) are identified to be T, C and v. When T is selected by C, T has a full complement of agreement features. Instead if selected by V, T is defective. v can only be selected by a functional category, and it can select, in addition to verbal elements, a DP/NP as its external argument. Each CFC is allowed an extra specifier beyond its s-selection, by virtue of the property of EPP. Whereas T is assumed to possess a universal EPP feature, for C and v it is optional. The Theta-theoretic Principle restricts the mode of EPP feature satisfaction:

(1) "Pure merge in theta positions is required of (and restricted to) arguments."

Thus T can satisfy its EPP by Merge of an expletive, whereas C and v have to satisfy it through the complex operation Move. The notion of phases is linked to the notion of CFCs. The lexical sub-arrays selected from the numeration is assumed to be a "natural syntactic

object SO", which is assumed to be "the closest syntactic counterpart to a proposition: either a verb phrase in which all theta roles are assigned or a full clause including the tense and force." (Chomsky (1998)). Thus a lexical sub-array should contain one occurrence of C or one occurrence of v. Phases are assumed to satisfy the strong cyclicity condition:

(2) "The head of a phase is 'inert' after the phase is complete, triggering no further operations."

Once a phase is completed its head cannot probe anymore, and its complement domain is inaccessible for further operations. This is given by the Phase Impenetrability Condition PIC:

(3) In phase α with head H, the domain of H is not accessible to operations outside α, but only H and its edge." (Chomsky (1998)).

Chomsky (1999) proposes a stronger derivational approach compared to Chomsky (1995, 1998) by proposing that operation Spell-Out can occur more than once during the course of the derivation. The domain of each strong phase head is Spelt-out or shipped off to PF at the next strong phase, and it (the Spelt-out domain) is treated as a kind of compound.

In this dissertation, I adopt and extend the basic analysis proposed by Joseph (2000) for clefts, which is based on the theory of phases and Multiple Spell-out. The derivation I propose is monoclausal, with the equative copula performing the function of a focus marker. This focus marker is what lends the semantics of exhaustivity and the existential presupposition etc. to the cleft focus. I attempt to give a unified analysis for the two focus marking strategies claimed for Malayalam, which are the preverbal focus position proposal and the cleft proposal.

1.3 Organisation of Chapters

This dissertation is organized as follows. Chapter 2 will basically be a review of the main approaches taken to analyse Focus in Malayalam so far. Chapter 3 will point out some empirical problems that cannot be resolved using the previous analyses and suggest plausible alternatives. Chapter 4 will be the conclusion.

CHAPTER 2

FOCUS IN MALAYALAM: EARLIER APPROACHES

Malayalam, spoken primarily in the southern parts of the Indian subcontinent, is an agglutinating language with relatively free word order (though basically SOV like other Dravidian languages) and no verb agreement. It is commonly described as a *wh- in situ* language, and has rich case marking. There are two ways in which focus has been claimed to be syntactically manifested in Malayalam. The most productive way of focusing a constituent in Malayalam is by clefting it. There are also claims for a preverbal focus position.

2.1 Jayaseelan (1989 *et seq*)

Jayaseelan (1989, 1995, 2002, 2008, 2010) claims that Malayalam is a positional focus language. He claims that there is a dedicated preverbal Focus position in Malayalam, similar to those attested in various other languages. Evidence for such a position, for him, comes from the *wh-* questions in the language. In accordance with popular claims that *wh-* words are inherently focused and that *wh-* movement takes place via focus positions, he claims that Malayalam uses this very same focus position for *wh-* movement too. (1) below is Jayaseelan's example of a simple *wh-* question:

(1) ninn-e aarə tall-i

2S-ACC who beat-PST

'Who beat you?'

The (linear) position just before the verb is the abovementioned focus position for him. We know the *wh-* word has moved to this focus position, because the canonical word order, which is S – IO – DO – V, has changed.

This preverbal focus position is also evidenced, according to him, by the stacking of the *wh*-words in a multiple *wh*- question:

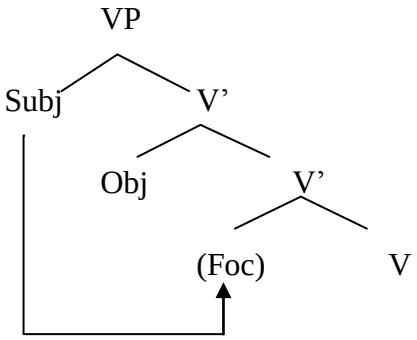
(2) ninn-e aarə eppo:L entinə tall-i

2S-ACC who when why beat-PST

'Who beat you when (and) why?'

However, postulating such a position contiguous to V would lead to downward movement:

(3)



Thus in order to overcome this problem (and to argue for the existence of such a position universally in all languages), he assumes an underlying Kaynean universal Spec-Head-Complement word order for all languages. In Jayaseelan (2010), adopting a proposal by Koopman and Szabolsci (2000), he builds a theory of generating the VO and OV surface word orders of languages by a mechanism of 'stacking' and 'stranding', which are steps of a uniform algorithm independently motivated by the need to bring together V and its inflectional suffix. This algorithm is always "morphologically generated" – T is always generated above V and they have to be brought together.

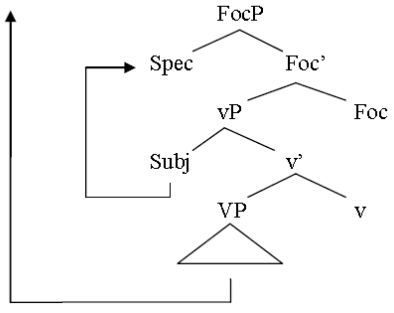
According to this system, all languages basically start out with a Spec-Head-Complement word order. After the V (actually $[_{v}$V-v]) is merged, all elements to the right of V are raised to the Spec of an XP above VP. This is known as "stacking". In the next step, the remnant VP is raised to Spec TP for getting the inflectional suffix. This is called "stranding" because the

arguments of the verb have been stranded at Spec XP. If the algorithm stops here, the language will have a surface VO order. If instead of stranding, the XP as a whole is "pied-piped" to Spec TP with the verb, the language gets a surface OV order.

Thus the OV order is derived from the same universal basic word order by VP-'vacating' movement as shown below in (4) [1]. Now the preverbal focus position can be obtained by

[1] His explanation in Jayaseelan (2001) is that the canonical word order of SOV languages like Malayalam is the mirror image of that of SVO languages like English. This can be obtained by reversing Larson's (1988) Thematic Hierarchy (AGENT > THEME > GOAL > OBLIQUES) or his principle for mapping the Thematic Hierarchy into constituent structure ("The lowest role on the Thematic Hierarchy is assigned to the lowest argument in constituent structure, the next lowest role to the next lowest argument, and so on.")

Even then, in order to account for the requirement that focused constituents occupy a position to the immediate left of V would require the postulation of a Focus Phrase above vP/VP and 'VP-vacating movement' of the arguments of V past the Spec of the Focus Phrase:

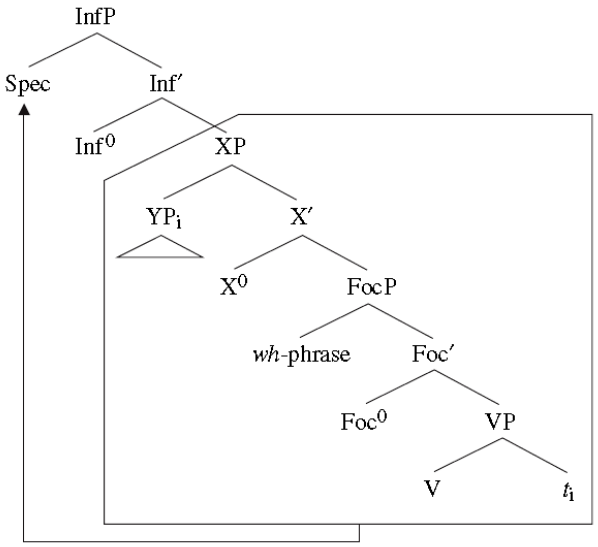


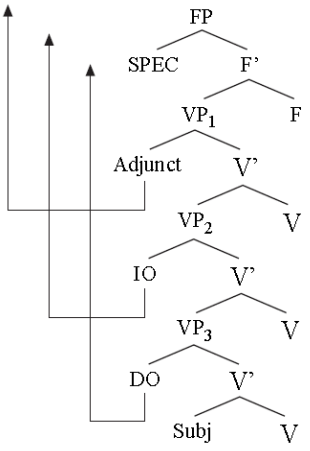
Thus the VP-vacating movement cannot be done away with. And when it is adopted, it would inevitably lead to 'crossing' movements as shown in (i) below:

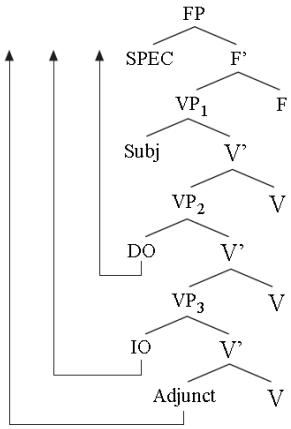
(i) (ii)

positing a Focus Phrase above vP: the required preverbal focus position is the Spec of a Foc head that selects the vP as its complement. The arguments of V move past this position when they raise for stacking. In the next step, the entire structure including the VP and the XP is pied-piped by VP, when it moves into Spec TP for the verb to pick up its inflection:

(4)







On the other hand, assuming a Spec-Head-Complement word order would enable having 'nested' movements (cf. (ii) above), which is a computationally simpler option. When it comes to the derivation of clefts, we will see that this also enables him to get the right word order.

This mechanism also explains the order of questions in multiple *wh*- questions for Jayaseelan (2001). In Malayalam multiple *wh*- questions, even though all the *wh*-s are moved to the preverbal position, the adjunct *wh*-s are required to be closer to the verb, cf. (5) below (Jayaseelan's (46a) and (b)):

(5) (a) awan a:r-e entinə konn-u

3SM who-ACC why kill-PST

'Why did he kill whom?'

(b) *awan entinə a:r-e konn-u

3SM why who-ACC kill-PST

For Jayaseelan, the nested ordering of the stacking movements causes this order restriction. Thus the higher of the *wh*-s moves out of the vP first when stacking takes place, and after pied-piping, will appear closest to the verb.

Elsewhere, he also claims that in terms of Chomsky's (1998, 2001) notion of Phases, the preverbal focus position (i.e., Spec FocP) is an escape hatch out of the vP phase. In order to do so he adopts Rizzi's (1997) articulated structure of the left periphery, i.e., the CP domain, and its extension to the vP phase.

However, Malayalam speakers generally prefer to ask their *wh*- questions in the form of a clefted sentence, cf. (6) below.

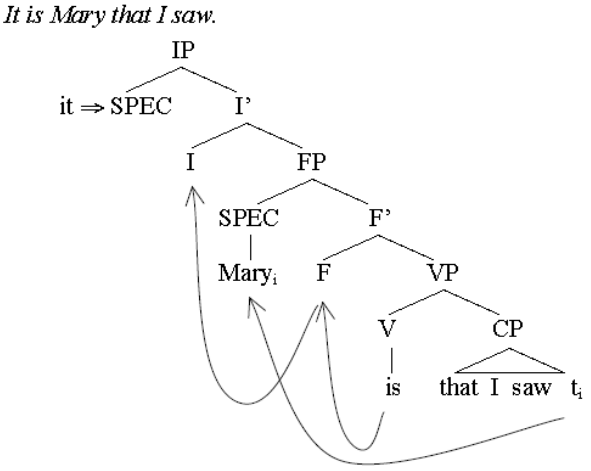
(6) aarə aaNə ninn-e tall-i-(y)atə

who be.PRES 2S-ACC beat-PST-NMLR

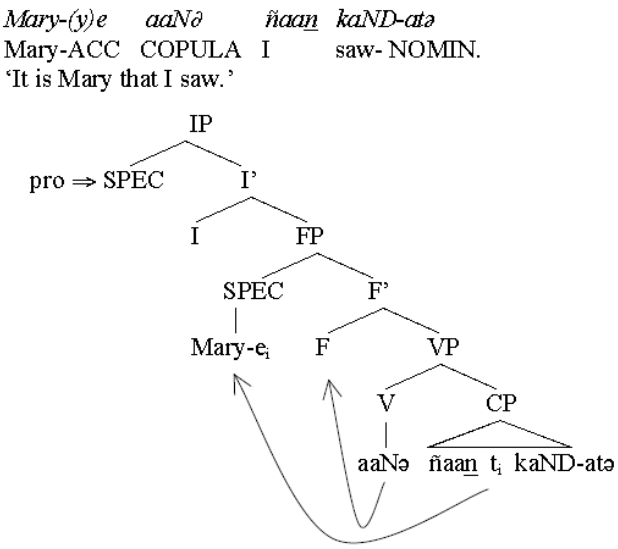
'Who is it that beat you?'

According to Jayaseelan, the matrix verb in a cleft sentence is the copula, and thus clefting is just another strategy to bring the *wh*- word to a preverbal position. The verb 'be' takes a clausal complement, and the constituent to be focused moves to the specifier position of a FocP just above the VP. Malayalam being a *pro*-drop language, the Spec of the matrix TP is filled by *pro,* just as in English the Spec TP is filled by pleonastic *it.* His structures for the English and Malayalam clefts are given below in (7) and (8):

(7) English:



(8) Malayalam:



As regards the nominalizer *-atə* suffixed to the cleft clause, Jayaseelan explains that it is the neuter agreeing form of the relativizer *–a* in Malayalam. The explanation he gives for why there is a neuter form is that either it is the default agreement form or that it does not count as agreement anymore, in accordance with the loss of verb agreement from the language.

Therefore the apparent 'floating' of the *wh*- word even into the cleft clause (cf. (9) and (10) below) does not disprove his claim that clefting also makes use of the preverbal focus position.

(9) ninn-e aarə aaNə tall-i-(y)atə
2S-ACC who be.PRES beat-PST-NMLR
'Who is it that beat you?'

(10) ninn-e tall-i-(y)atə aarə aaNə
2S-ACC beat-PST-NMLR who be.PRES
'Who is it that beat you?'

This is because the constituents to the left of the clefted phrase have, according to him, actually moved to topic position(s) above the focus projection.

It is further argued in Amritavalli and Jayaseelan (2005) that there are two different types of cleft in Malayalam, a long-distance cleft that employs relativization to extract the clefted phrase from the cleft clause, and a short-distance cleft that is derived from a cleft clause that has no C-system: it is only an IP, and since it is not a phase the clefted phrase is freely extracted into the matrix clause.[2]

[2] Additionally, he also points to the existence of another type of cleft in Malayalam – one with an infinitival cleft clause, denoting habitual action or a 'usual' state-of-affairs. This type of cleft behaves in a similar fashion as the short-distance cleft, and does not involve relativization, nor does it allow long-distance clefting:

However, there are some problems with Jayaseelan's approach to focus in Malayalam.

(a) A close inspection of data will reveal certain important differences between the cleft and the non-cleft preverbal focus construction, making it difficult to bring them both under the common umbrella of a single focus position. For instance, there are marked differences in the interpretation of both the constructions. Cleft focus induces an exhaustive reading of the focus/ focused constituent (cf. (11) below).

(11) jo:N a:Nə a:ppiL tinn-atə

John be.PRES apple ate-NMLR

'It is John who ate the apple.' (Interpreted as no one else but John)

In addition, clefts also give rise to some sort of an existential presupposition. For instance, (11) above carries a presupposition that someone ate the apple. Such presuppositions are not available in case of the non-cleft preverbal focus.

(i) **kaññi a:Nə** [John t_i kuDik'k'-uka]

rice gruel be.PRES John drink-INF

'It is rice gruel that John drinks (as a practice).'

But even a cleft sentence to indicate a future event or action will use the same infinitival cleft clause:

(ii) na:Le a:Nə jo:N ett-uka

tomorrow be.PRES John reach-INF

'It is tomorrow that John will reach.'

In the same way, (i) is ambiguous: it could also mean that it is rice gruel that John will drink (probably at the next meal).

(b) Moving any constituent to a higher position changes the c-command and other relations obtaining in the structure. Movement therefore has to be properly motivated, and is constrained by economy conditions like *Procrastinate*. Usually movement is motivated by scope considerations, which displays its consequences in the form of crossover effects, operator-variable binding relations, etc. For instance, in clefts one can generally expect a weak crossover effect due to the movement of the clefted constituent to a higher position. cf. (12) below.

(12) ?? a:r-e_i a:Nə avan-te_i amma kaND-atə

who-ACC be.PRES 3SM-GEN mother saw-NMLR

'??Who did his mother see (lit. who was it that his mother saw)?'

As I will later argue, argument *wh*-s in the preverbal position need not elicit answers, which shows that the *wh*- phrase does not get wide scope. On the other hand, clefting does induce the desired results. É. Kiss (1998) proposes various 'diagnostics' to distinguish between two types of focus, Identificational and Information focus, applying which, clefting clearly turns out to be focus movement. The same result is not got in the case of movement to a preverbal focus position.

(c) Jayaseelan basically follows a Kaynean approach. In the current Chomskyan theoretical framework, syntactic phenomena are based on the properties and interactions of syntactic features. In Chomsky (2000), phenomena like focus and topic are explained in terms of 'P-features', and do not call for the projection of separate heads. Strong phase heads like C and v* are assumed to have special properties, like EPP and the ability to have 'outer' or 'extra' specifiers. This comprises what is called the *edge* of the phase and can host the P-features, which spell-out as 'force', 'topic' and 'focus' features. The edge of the phase serves as the escape-hatch through which constituents may move out of the phase. The preverbal position

is, for Jayaseelan, the Spec of a FocP above vP, which serves as an escape hatch out of the vP. However, the escape hatch can be merely another Spec position (at the worst) adjoined to vP, and does not motivate the projection of another Foc head. Spec FocP cannot be an escape hatch since the FocP is not a strong phase.

(d) In the analysis of the cleft construction as using a preverbal focus position parallel to non-clefts, the copula is treated by Jayaseelan as a verb. This is problematic, as the copula is incapable of acting as a full verb. This is evident in its interactions with various properties like tense etc. Babu (2006) identifies two different functions of the copula *a:Nə* in Malayalam, one of which is in the cleft construction as a 'focus marker', that contributes the semantics of exhaustivity to the cleft focus. Jayaseelan's account does not consider this intuition.

(e) Speakers seem to be divided on the interpretation of non-cleft questions as valid questions[3]. Many speakers (including me) interpret the non-cleft question as an echo question rather than an information question. Again, in this too, there is distinction between argument *wh*-s and adjunct *wh*-s. Whereas most argument *wh*-s in the preverbal position would more often than not be interpreted as echo questions, adjunct *wh*-s seem to be fine as information questions. These facts cannot be accounted for by Jayaseelan's theory.

[3] I assume it to be because of the overwhelming preference for the cleft construction to ask *wh*- questions. Indeed, while giving judgement on the non-clefted *wh*- construction, explanation or elaboration by the informants reveals that they usually tend to process or re-interpret the non-clefted *wh*- question as a clefted *wh*-question. Also Malayalam speakers have a tendency of simplifying pronunciation, due to which a cleft like *ninne a:rə a:Nə kaNDatə* 'Who is it that saw you?' is pronounced as *ninne a:ra: kaNDe*. This could also have a part to play in re-interpreting the non-cleft question as a cleft question. (cf. fn. (2) in chapter 3)

2.2 Mathew (2011)

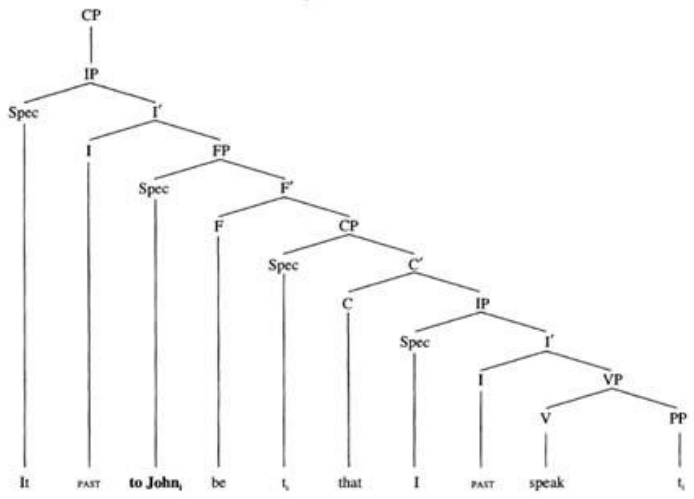
Mathew (2011) points out, based on É. Kiss' (1998) distinction between Identificational and Information focus, that the cleft and non-cleft focus constructions described by Jayaseelan (1989 *et seq*) are in fact distinct. Malayalam has not one but two focus projections, one above vP and another one in the CP domain. She investigates Chomsky's (2001 *et seq*) claim that projections like Focus appear recursively in every phase.

É. Kiss (1998) identifies the two types of Focus as distinct from each other, based on the observation that in Hungarian they are never interpretational variants but are associated with different structural positions. *Wh-* phrases (other than *miért* 'why') always occur in the preverbal identificational focus position in Hungarian.

According to É. Kiss (1998), syntactically the constituent called the identificational focus itself acts as an operator, moving into a scope position in the specifier of a functional projection and binding a variable. She posits a Focus Phrase that selects VP as its complement, and the focused constituent moves to Spec FP, accompanied by V-to-F movement. On the other hand, Information focus merely indicates new or non-presupposed information and is usually marked by one or more pitch accents.

É. Kiss (1998, 1999) also proposes the same distinction between Identificational and Information focus for English. The English counterpart of the Hungarian preverbal identificational focus construction is the cleft construction, and the English VP-internal *in situ* information focus constituent is assigned a pitch accent. Her structure for identificational focus for the English cleft construction is reproduced below in (13). The only difference in the structure of the Hungarian preverbal focus construction and the English cleft construction, according to her, is that the FP in Hungarian selects for a VP, whereas in English it selects a CP.

(13)



Information focus in English, on the other hand, is *in situ* and receives nuclear stress.

Adopting the difference between the two types of focus, Mathew (2011) claims that identificational focus is marked in Malayalam by the cleft construction, which uses the FocP in the CP domain, whereas information focus is marked by movement to the FocP above vP[4]. She uses Kiss' (1998) diagnostics of exhaustivity and distributional restrictions to make this distinction[5].

The cleft construction in Malayalam is clearly exhaustive in nature, unlike the non-clefted focus construction. This is shown using Szabolsci's (1981) tests of exhaustivity. A sentence with a co-ordinated phrase in the focus position is compared to an identical sentence with only one of the co-ordinated phrases in the focus position. If the second sentence is not a logical consequence of the first, then the focus is exhaustive. Controlling for possible

[4] In her footnote fn. (4), she says that she has only adopted the classification of Information focus from Kiss (1998), but not its analysis. Though É. Kiss (1998) analyzes Information focus to not involve any movement, Mathew adopts Jayaseelan's movement to preverbal position to be for Information focus in Malayalam.

[5] She mentions in her footnote fn. (8) that É. Kiss employs scope interactions as well in order to identify Exhaustive Focus. However, she does not reproduce those tests as Malayalam shows surface scope and makes use of scrambling, which make the use of this test less than ideal.

misinterpretation due to the distributive/collective distinction in the sense of Gryllia (2008), it is shown that the Malayalam cleft construction is exhaustive (cf. (14) and (15)), whereas Jayaseelan's non-clefted focus construction is not (cf. (16) and (17)). The following examples taken from Mathew (2011) serve to illustrate the point:

(14) de:vi-kk-um paili-kk-um a:Nə karambi o:ro: pustakam koDutt-atə

Devi-DAT-CONJ Paily-DAT-CONJ FM Karambi a book each gave-NMLR

'It is to Devi and Paily that Karambi gave a book each.'

(15) paili-kkə a:Nə karambi oru pustakam koDutt-atə

Paily-DAT FM Karambi a book gave-NMLR

'It is to Paily that Karambi gave a book.'

(16) enn-e karambi-(y)um paili-(y)um kaND-u

1S-ACC Karambi-CONJ Paily-CONJ see-PST

'Karambi and Paily saw me.'

(17) enn-e paili kaND-u

1S-ACC Paily see-PST

'Paily saw me.'

Also, unlike the cleft, the preverbal position is also amenable to adding information.

(18) ninn-e a:rə kaND-u

2S-ACC who see-PST

'Who saw you?'

(19) enn-e paili kaND-u; karambi-(y)um kaND-u

1S-ACC Paily see-PST Karambi-CONJ see-PST

'Paily saw me; and Karambi too.'

Distributional restrictions in terms of the availability of the focus position for universal quantifiers, *also*-phrases and *even*-phrases that apply to the Hungarian preverbal Identificational focus position also apply in Malayalam to the cleft construction (cf. (20)-(22)), but not the preverbal position (cf. (23)-(25)).

(20) *ella:varum a:Nə vann-atə

everyone FM came-NMLR

'It is everyone that came.'

(21) *karambi-(y)um a:Nə paili-(y)e kaND-atə

Karambi-CONJ FM Paily-ACC saw-NMLR

'It is Karambi also that saw Paily.'

(22) *karambi po:lum a:Nə paili-(y)e kaND-atə

Karambi even FM Paily-ACC saw-NMLR

'It is even Karambi who saw Paily.'

(23) paili-(y)e ella:varum kaND-u

Paily-ACC everyone see-PST

'Everyone saw Paily.'

(24) paili-(y)e avan-um kaND-u

Paily-ACC 3SM-CONJ see-PST

'He also saw Paily.'

(25) paili-(y)e avan po:lum kaND-u

Paily-ACC 3SM even see-PST

'Even he saw Paily.'

This also shows that the cleft is an instance of Identificational focus, whereas the non-cleft preverbal focus construction only demonstrates Information focus.

With regard to É. Kiss' claim as to Identificational focus being [+exhaustive] and [+contrastive] in Hungarian, Mathew points out that this is not so in Malayalam. The cleft construction is not associated with contrastivity. She quotes Horvath's (2010) definition of contrastivity in terms of sets of alternatives[6] and demonstrates that the Malayalam cleft is not associated with alternative sets. Example (26) below is taken from Mathew (2011) (her (36)):

(26) A: John-(in)e a:Nə avaR select ceyt-atə

John-ACC FM 3P select did-NMLR

'It is John whom they selected.'

B: Were there other applicants?

or There might not have been any other applicants.

The responses in B are perfectly fine to respond to the statement in A, which means the cleft does not involve contrastivity. She also shows that in fact the cleft is the preferred option in focusing unique single member sets, and cannot be associated with subsets (of alternatives) like contrastive focus:

(27) (In the context of talking about who won in the last World Cup)

Italy a:Nə jeyicc-atə

Italy FM won-NMLR

'It is Italy that won.'

(28) #Italy ma:tRam jeyicc-u

Italy only win-PST

'Only Italy won.'

[6] This account is apparently comparable to Rooth's (1985, 1992) Alternative Semantics proposal of focus being indicative of the presence of a set of alternatives to the focused constituent.

So now after showing that both the focus constructions are very different she proposes a different analysis for the cleft construction. For her the copula *a:Nə* in the Malayalam cleft construction is not a verb as Jayaseelan considers, rather she analyses it as a focus marker. She adopts the monoclausal analysis of clefts in Hiraiwa and Ishihara (2002).

Hiraiwa and Ishihara (2002) compare three similar constructions, sluicing, cleft and an *in situ* focus construction, and propose a unified syntactic account for them adopting Rizzi's (1997) fine division of the C-domain. Crucially, they explain the structure of the *in situ* focus construction using the Rizzi structure of the C-domain, and derive the cleft construction from it by focus movement and remnant topicalization.

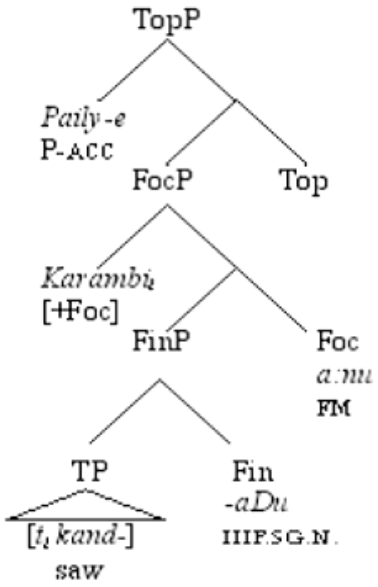
Mathew (2011) accordingly adopts their structure of the *in situ* focus construction and derives the Malayalam cleft from it by the right dislocation of the background information in the remnant FinP. Below (30a-b) is her analysis of the Malayalam cleft sentence in (29) (her (58)).

(29) Paily-e karambi a:Nə kanD-atə

Paily-ACC Karambi FM saw-3SN

'Paily, it is Karambi that saw him.'

(30) (a)



(b) Paily$_j$-e Karambi$_i$ t$_k$ a:Nə [[t$_i$ t$_j$ kanD-]atə] $_k$ (right dislocation)

However, there are some problems here too.

(a) For Mathew (2011) the Identificational focus position is in the "left periphery" of the CP. Kidwai (1999) argues for focus positions to be in the vP domain, showing that in many 'positional focus' languages adjacency to V is due to the need for PF-checking of the [+FOCUS] feature which is for her [PF-Interpretable]. Going by that argument, Spec CP is "too far off" to enable checking of the feature.

(b) In order to derive the right word order, Mathew (2011) takes recourse to various phenomena like topicalization, right dislocation, etc. Whether these are operations independently available in Malayalam is a question. For instance, Sener (2010), Kidwai (1995) etc. show how there is no need for movement to license topics in languages like Turkish and Hindi respectively.

2.3 Joseph (2000)

Joseph (2000) agrees with Jayaseelan's claim of the existence of a preverbal focus position. She analyzes the cleft construction in Malayalam and claims that the focus assigned to the clefted phrase is due to its precopular position, which is an extension of the preverbal focus position advocated by Jayaseelan (1989 *et seq*). However, she attributes the (semantic) differences between clefts and non-clefted focus sentences (like contrastive interpretation etc.) to the nature of the copula[7].

She adopts, unlike Amritavalli and Jayaseelan's (2005) biclausal analysis of clefts, a monoclausal account. Let us look at how she would derive a cleft sentence like the one in (31):

(31) ravi a:Nə si:ta-kkə pustakam koDutt-atə

Ravi be.PRES Sita-DAT book gave-NMLR

'It is Ravi who gave the book to Sita.'

The copula being a monadic predicate, selects a DP complement. However, the vP dominating the copula lacks a full complement of features and thus is a weak phase. Therefore there is no selection of external argument nor assignment of case to the complement by the copula. The derivation of the clause with the copula as the predicate is given below:

(32) [$_{CP}$ [$_{TP}$ [$_{vP}$ [$_{VP}$ [$_{DP}$ ...] aaNə]]]]

[7] Babu (2006) elaborates on the two copulas in Malayalam and the differences in their nature. He explains among other things that one of the functions of the Malayalam equational copula *a:Nə* is as a focus marker in cleft constructions. He also argues that it is what contributes the exhaustivity to the cleft focus. More about the copula will have to await further research.

With overt verb raising in Malayalam (Jayaseelan 1989, Madhavan 1987, Srikumar 1992, 1994a) the copula adjoins to T. Once the derivation reaches the strong CP, the TP can spell out.

As for the cleft clause, her derivation is as follows. The *-atə* suffix attached to the predicate of the cleft clause, traditionally glossed as the nominalizer, is a pronominal, and therefore a D[8], which refers to and selects for the CP, which is the cleft clause. This CP, however, lacks force, evidenced by its inability to take a complementizer or a question particle, and therefore is argued to be a weak phase. It is this DP (headed by the pronominal *atə*) that is selected by the copula.

The crucial part of her proposal lies in the adoption of Multiple Spell Out (MSO), as detailed in Uriagereka (1999) and Chomsky (1999). According to the proposals in MSO, the linearization of a structure in PF is sensitive to command and the setting of the Head Parameter. A Command Unit must instantiate a unique setting of the Head Parameter, which implies that a merged structure would spell out before it further merges to a head or its projection with a different setting of the Head Parameter. The spelled out structure then behaves like a single word or a compound.

Malayalam is essentially a head final language with a basic SOV order, but according to Joseph its DPs are head initial. Thus the DPs spell out separately, as shown in (21) below, and replaced by placeholder Ds. Derivations essentially take place in parallel and thus while the DPs are merged and spelled out, simultaneously the cleft clause is merged until the DP

[8] Even though Joseph considers DPs to be head-initial in Malayalam, the DP headed by this D is head-final. This is because it is different from the other Ds: it selects a CP as its complement. However, it is debatable whether Malayalam, and the Dravidian languages in general, have DPs at all.

headed by *atə* is formed, which spells out, as in (34). This DP is in turn selected by the copula and the derivation continues till the next strong phase, the matrix CP (cf. (36) below). Thus the cleft is derived and spelled out.

(33) [$_{DP}$ D [$_{NP}$ ravi]][$_{DP}$ D [$_{NP}$ si:ta-kkə]][$_{DP}$ D [$_{NP}$ pustakam]]

(34) [$_{DP}$[$_{CP}$[$_{TP}$[$_{T'}$[$_{\underline{v}P}$ D [$_{\underline{v}'}$[$_{VP}$ D [$_{V'}$ D t_V]] $\underline{v}$]] T-kodutt-u$_V$]] C] atə]

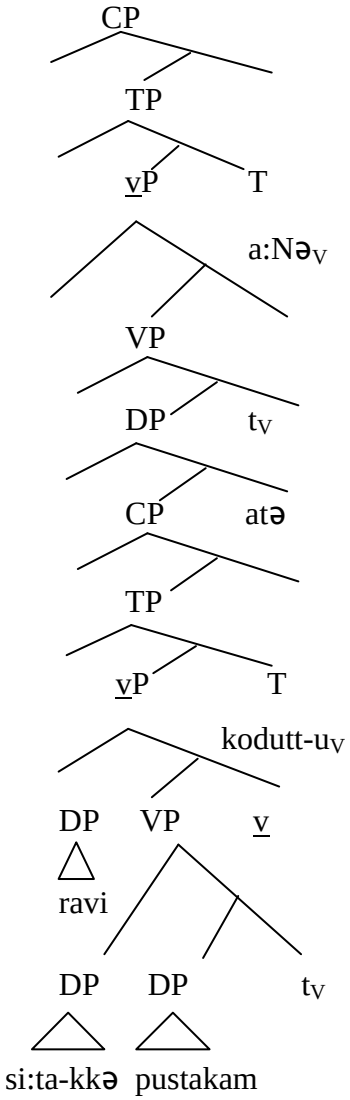
Now the required surface order of the sentence would be obtained by PF-reordering, which may look something like (35) below:

(35) ravi$_j$ t_i a:Nə [t_j sita-kkə pustakam kodutt-atə]$_i$

Ravi be.PRES Sita-DAT book gave-NMLR

'It is Ravi who gave her the book.'

(36)



2.4 Conclusion

We have looked at the main literature on Malayalam focus and seen that they leave some questions unanswered. The main questions that I will attempt to answer in this dissertation are the following:

(a) We have seen that Jayaseelan's (1989, *et seq*) account does not differentiate between the preverbal informational focus and the cleft. Mathew's (2011) account also is inadequate. There needs to be another explanation that will take care of the relevant issues.

(b) Both Jayaseelan (1989, *et seq*) and Mathew (2011) propose a separate functional head Focus and movement to its specifier position for licensing the focus feature. However these movements are not motivated independently, nor are functional projections like FocP justified by the current framework in generative syntactic enquiry. How to tackle these phenomena within the current framework will be another question.

(c) For Jayaseelan (1989, *et seq*) and Joseph (2000), the *a:Nə* in the cleft construction is a copula. On the other hand, Mathew (2011) treats it as a focus marker. Babu (2006) explains that *a:Nə* is in fact the equative copula of Malayalam, but it has another avatar in the cleft construction – that of a focus marker. Joseph (2000) also attributes the exhaustivity and other properties of the cleft focus to *a:Nə*, which is in fact in agreement with Babu's (2006) proposals. What the *a:Nə* really is and what its function is needs further investigation.

(d) There are also empirical issues like the echo interpretation given to *wh*-s in the preverbal position, that raise the question of whether the preverbal position is actually an available focus position at all.

These are some of the questions I intend to look into in this dissertation.

CHAPTER THREE

SOME EMPIRICAL ISSUES AND ANALYSIS

We have seen in Chapter 2 that there is clearly a syntactic distinction in Malayalam between the two types of focus. Identificational focus, instantiated in Malayalam by clefting, is evidently quantificational in nature, and involves movement to a higher position. It also involves an existential presupposition, in addition to exhaustive identification from among a set of potential elements for which the predicate can hold. Information focus on the other hand is *in situ*, even though Mathew (2011) claims it to occupy Jayaseelan's (1989 et seq) preverbal focus position.

For Jayaseelan (1989 et seq), all focus in Malayalam is the same, and they are syntactically derived via movement to a preverbal focus position, the Spec of a Focus Phrase (FP) at the edge of vP/VP. Clefting is simply a biclausal variant of this very same focus movement, with the copula as the matrix verb and the clefted phrase moving to the preverbal Spec FP position of the copula. Mathew (2011) goes a step further and distinguishes between Identificational and Information Focus. She accomplishes this by showing that the cleft and non-cleft focus constructions are different, mainly in terms of exhaustive identification. She accounts for this difference in interpretation by postulating two different FPs, one at the edge of vP and another at "the left periphery of the CP". The preverbal FP at the edge of vP serves as the position for Information focus, and the FP in the CP domain serves as the position for Identificational focus. Joseph (2000) also distinguishes between the two types of focus, but does not see the projection of a focus head as necessary to explain them. The focus in the cleft is attained due to the presence of the copula. Everything else happens by PF reordering.

Now consider the following data. In a cleft like (1a) below (eg. (12) reproduced from chapter 2), we observe a Weak Crossover Effect, attributed to the quantificational nature of Identificational focus (cf. É. Kiss (1998)):

(1) (a) ?? a:r-e_i a:Nə avan-te_i amma kaND-atə

who-ACC be.PRES 3SM-GEN mother saw-NMLR

'??Who did his mother see (lit. who was it that his mother saw)?'

However, the sentence seems to improve on assigning heavy stress to *amma* 'mother':

(b) a:r-e_i a:Nə avan-te_i **amma** kaND-atə

who-ACC be.PRES 3SM -GEN mother saw-NMLR

'??Who did his mother see? (lit. who was it that his mother saw)'

Assuming that it is the phrase *avante amma* 'his mother' that has been focused by the prosodic stress assigned to *amma* 'mother', we follow Chomsky's (1977) proposal of analyzing focus/topic structures (among various other structures) as *wh*- movement (the term was understood in a general, analogous sense; later it came to be known as Operator movement). Accordingly, topicalized and focused phrases can act as quantifiers; they move at LF to a higher position (COMP, in his original conception) and can bind a variable in the original position of the moved phrase:

> "... *wh*- movement leaves a non-terminal trace, just as all movement rules do. That is, the position from which the *wh*- phrase moved remains in the derived constituent structure with its index, identical to the index of the *wh*- phrase, now in COMP. It seems clear that words such as *who, what,* etc., should be regarded (at least in questions) as quantifiers of some sort. Thus at the level LF, the sentence (30) will be represented essentially as (31):
>
> (30) *who did John see?*

(31) *for which x, x a person, John saw x*

There is good reason to suppose that the rules extending a derivation to LF form such expressions as (31), and that variables are introduced in other ways as well, in particular, by the expansion of NP quantifiers such as *every* and by a rule of FOCUS." Chomsky (1977), 82-83.

Since the phrase *avante amma* 'his mother' in (1b) has been focused, it is 'focus-raised' at LF to adjoin to the Identificational focus position (2). This must be how the weak crossover effect is overcome.

(2)	[avan-te$_i$	amma]$_j$	a:r-e$_i$	a:Nə	t_j	kaND-atə
	3SM-GEN	mother	who-ACC	be.PRES		saw-NMLR

On the other hand, this pattern is not seen in the non-cleft version, where *avante amma* 'his mother' is already in a higher position than *a:rə* 'who':

(3) (a) avan-te$_i$ amma a:r-e$_{*i}$ kaND-u
3SM-GEN mother who-ACC saw-PST
'Who did his mother see?'

(b) avan-te$_i$ **amma** aar-e$_{*i}$ kaND-u
3SM-GEN mother who-ACC saw-PST
'Who did his mother see?'

This is probably because the *wh-* here does not get wide scope[1]. Indeed, the reading I (like many other speakers) prefer for *wh-* words in the preverbal position is of an echo question

[1] É. Kiss (1998) also gives examples for weak crossover effects induced by focus in Hungarian. Her explanation for the unacceptability of the information focused sentence, which some find to be an apparent weak crossover effect, is that the unacceptability is due to discourse-pragmatic reasons: the referent of the focused

rather than a content/information question (cf. footnote 3 of Chapter 2). Compare (4a, b and c) below:

(4) (a) ni: a:r-e kaND-u ennə ni: paraɲɲ-u

2S who-ACC see-PST that 2S say-PST

‘(Lit.) You said (that) you saw WHO?’

(b) a:r-e a:Nə ni: kaND-u ennə (ni:) paraɲɲ-atə

who-ACC be.PRES 2S see-PST that 2S said-NMLR

‘Who did you say you saw?’ (Lit. ‘Who is it that you said you saw?’)

(c) ni: a:r-e kaND-u ennə a:Nə (ni:) paraɲɲ-atə

2S who-ACC see-PST that be.PRES 2S said-NMLR

‘Who did you say you saw?’ (Lit. ‘??You saw who is it that you said?’)

In (4a) the *wh*-phrase *a:re* ‘who(m)’ in the embedded clause cannot be interpreted as having matrix scope. However, it is not ungrammatical since the *wh*-phrase gets an echo interpretation. In order to get matrix scope, either the *wh*-phrase will have to be clefted, as in (4b), or the entire clause will have to be clefted, as in (4c).

Echo questions in Russian have been argued to be D-linked (Chernova (2011)). According to her, the value of an echo question is explicitly given in previous discourse. That is, the set of potential referents of the *wh*- phrase is restricted to only one element, which the speaker and hearer are aware of, but the latter is ignorant of its value. Thus Chernova argues that it is some kind of a D-linked constituent in the sense of Pesetsky (1987). Pesetsky argues that

---

constituent is first introduced by the pronoun as an individual already present in the domain of discourse, and then introduced as new information by the information focus.

> "...the crucial difference between a *which*-phrase and the *normal* occurrence of *who* or *what* is found in discourse. [...] Roughly, *which*-phrases are *discourse-linked* (*D-linked*), whereas *who* and *what* are normally not D-linked. When a speaker asks a question like *Which book did you read?*, the range of felicitous answers is limited by a set of books both speaker and hearer have in mind. If the hearer is ignorant of the context assumed by the speaker, a *which*-question sounds odd. [...] No such requirement is imposed on *wh*-phrases like *who, what,* or *how many books*. These phrases may be *non-D-linked."* (Pesetsky (1987), 107-108)

Pesetsky explains further that non-D-linked *wh*-phrases are quantifiers and adjoin to S', whereas D-linked *wh*-s are not. Thus they are able to receive a 'Baker-style' interpretation (unselective binding by a Q particle at LF) without movement.

Chernova (2011) proposes, following Sobin (2010), that in echo questions in Russian and Spanish there is a superior complementizer CP_{EQ} that selects the frozen CP structure of an echoed utterance U as its complement, and binds an EQ-induced *wh*-expression (cf. (5)).

(5) $[_{CP_{EQ}}\ C_{EQ}\ [_{CP}\ C\ \ldots\ldots\ wh_{EQ}]]$

I assume that it is some such Q_{EQ} in a specifier position of the vP that unselectively binds the echo *wh*- phrase in its preverbal position in Malayalam too. Indeed, the echo *wh*- can be in any position mirroring the structure of the echoed vP:

(6) a:rə ninn-e tall-i (ennə)

who 2S-ACC beat-PST (that)

'(that) WHO beat you?'

Notice that pronouncing the complementizer, which is in fact a quotative, clarifies this interpretation considerably. This shows that Jayaseelan's (1989 et seq) preverbal position is

not necessarily a (syntactic, at least) focus position. At least some speakers (including me) find sentences like the following to be perfectly acceptable:

(7) a:rə entinə si:ta-kkə pustakam va:ŋŋi-koDutt-u

who why Sita-DAT book buy-give-PST

'Who bought Sita books (and) why?'

In fact, there have been arguments like those of Kidwai (1999) that the occurrence of focus positions in languages is in fact superficial, and is due to some kind of PF-scrambling, driven by the checking requirements of [PF[±Interpretable]] features. She argues that these features are interpreted at a level distinct from LF termed DOMAIN D(ISCOURSE), located at the edge of the PF-component.

From the previous subsection it has become clear that *wh*-s that stay in the preverbal position get an echo interpretation rather than an information question interpretation. Thus they also cannot be focused. That the *wh*-phrase need not necessarily have to be in the preverbal position is also clear. Thus what appears to be a 'preverbal focus position' is in fact only a sort of illusion created by scrambling (possibly in the PF), in the spirit of Kidwai (1999). True questions are those in which the *wh*-phrase moves up into what is called the cleft focus.

This means that in the cleft construction something prevents the vP from spelling out before the *wh*- moves out. I argue that it is the presence of the focus marker *a:Nə* at the vP edge that

prevents the vP phase from spelling out.[2] It could well be the case that *a:Nə* somehow triggers 'phase sliding' (*a la* Gallego (2006)), and extends the vP phase.

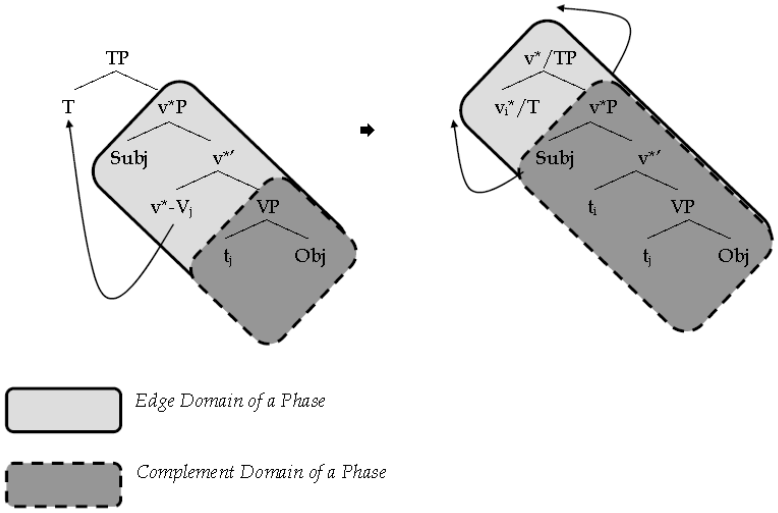
Gallego (2006) proposes that in Null Subject Languages, the T(ense) head appears to be endowed with *edge features* and become a strong phase head. This is in fact a "side-effect" of v*-to-T movement: when v* internally merges with T it re-labels the whole structure, forcing a species of *reprojection* that Gallego calls '*phase-sliding*'.

He adopts and builds on Pesetsky and Torrego's (2001) account of T-to-C movement. According to Pesetsky and Torrego (2001) the T(ense) head is the locus of Case, and phase heads select for T heads. The T head selected by v* may be called Aspect or Voice etc. Even prepositions have been shown to be a species of T heads. Gallego (2006) also adopts Lasnik's (1999) claim for head movement: if T is an affix in a given language, [$_{v*}$[V-v*]]-to-T movement must occur at the PHON component, but if T is not an affix, [$_{v*}$[V-v*]] can move to T in syntax. He investigates Null Subject Languages and argues that this movement is not an instance of canonical head movement. The moved head here targets the root of the TP, without tucking-in. The most important consequence of this is that somehow the v*P phase is 'moved upwards' to TP. Like in Chomsky (1993, 1995), it can be argued that this verbal movement results in the extension of a "checking domain". Already Chomsky (1986) had argued that v*-to-T movement could result in an amalgamated form V1, capable of removing the barrierhood of VP.

[2] It remains to be investigated exactly where in the vP *a:Nə* is actually merged. Another possibility is that the Q_{EQ} operator that binds the *wh-* phrase in a preverbal echo question occupies the same position as the one where *a:Nə* is merged. cf. (4a,b).

In short, this movement amounts to the 'pushing up' of the v*P phase. The strong phase head v* can still head the whole structure and trigger any syntactic operation from this position. A diagrammatic representation of this *Phase Sliding* is given below in (8):

(8)



Thus after the merging of the focus marker *a:Nə* in some XP above the vP, it plausibly raises to T[3] and inherits the properties of v, pushing up the phase.

[3] It is assumed that the focus marker *a:Nə* has evolved through some process like grammaticalization from the equative copula *a:Nə*. Babu (2006) says that it is another function or species of the equative copula. Sure enough, clefts are used in Malayalam with past tense form etc of equative copula, though most times it does not change the meaning, i.e., give a tensed interpretation of the clefted sentence, unlike in languages like English. For example, (i) and (ii) use the present and past forms of *a:Nə* respectively, though there is not much difference in meaning:

(i) annə ra:man a:-(y)irunnu ste:San-il vann-atə

that-day Raman be-pst station-loc came-nmlr

'It was Raman who came to the station that day.'

This would imply that the Identificational Focus position is in fact in the vP phase, and not in the CP phase as argued by Mathew (2011). Kidwai (1999) assumes that focus pertains to verbal projections:

> "We are going to keep with the traditional assumption that the feature [+FOCUS] is checked by a verbal projection, though not with the view that [+FOCUS] is a feature analogous to Case. It appears that focusing pertains more to the realm of the verbal paradigm than to the nominal one, and is in fact more predicational, and rather like mood, or choice of tense and aspect [...] [T]hen it is possible for us to maintain that the PF-checking of the feature [+FOCUS] is done by V^0." Kidwai (1999), 229.

Thus if *a:Nə* (at least a silent version of it), which is a focus marker here, is merged the *wh*-can raise to (possibly) some higher Spec of the vP and get wide scope.

But the story does not end there. Malayalam exhibits an argument-adjunct asymmetry in its *wh*- questions. Whereas argument *wh*-s cannot get wide scope in the preverbal position and get an echo interpretation unless raised to a higher position in the vP, adjunct *wh*-s seem to be fine in both the preverbal position as well as in the cleft focus. (cf. (9) and (10) below.)

(9) ni: entinə vann-u

2S why come-PST

'Why did you come?'

(ii) annə ra:man a:Nə ste:San-il vann-atə

that-day Raman be-pres station-loc came-nmlr

'It is Raman who came to the station that day.'

(10) ni: entinə a:Nə vann-atə

2S why be.PRES came-NMLR

'Why did you come?' (Lit. 'Why is it that you have come?')

In order to account for this distinction, I propose late merger (Stepanov (2001)) of the (preverbal) adjunct *wh*-s. Stepanov argues for a formal algorithm of phrase structure building within the bare phrase structure theory, which prevents the cyclic merger (merger by substitution) of adjuncts, and allows them to merge only post-cyclically. This algorithm is based on the idea of Least Tampering of Chomsky (2009), according to which, "Given a choice of operations applying to a syntactic object labelled α, select one that does not change @(α); @(X) being a set of c-command relations in a syntactic object labelled X." This can be reformulated in other terms as "Merge at the *root* when possible." (Root here means a category c-commanded by no other category.) Merge by substitution "inside" an already merged phrase marker changes the c-command relations in the phrase marker and is thus less economical. Therefore substitution is always cyclic merge. Adjunction creates segments of an existing category, and any subsequent merge other than an adjunction will alter the c-command relations in the phrase marker. Thus it would be economical to merge (by substitution) at a point before the adjunction takes place. So adjunction is forced to be post-cyclic by this algorithm.

Stepanov also argues that it is the featural make-up of the merged element that decides whether to merge by adjunction or by substitution. He suggests that there is a systematic correlation between the interpretability-uninterpretability dichotomy and the substitution-adjunction dichotomy. The definitions of adjunction and substitution in terms of the interpretability of features is given in (11a) and (b):

(11)(a) A non-projecting syntactic object α is Merged with a syntactic object β by *adjunction* iff the label of α contains no active ('unchecked') uninterpretable feature(s).

(b) A non-projecting syntactic object α is Merged with a syntactic object β by *substitution* iff the label of α contains active ('unchecked') uninterpretable feature(s).

Uninterpretable features can be 'inside' the object to be adjoined, as long as it is not in the label. Similarly the definition also does not imply that as soon as Merge by substitution takes place the uninterpretable features are checked. The only implication from these definitions is that the projection of a separate category is a property of uninterpretable features, regardless of when they are checked.

Using late merger then, we can also account for the asymmetry between argument and adjunct questions in Malayalam. I propose that the uninterpretable feature that adjunct *wh*-s (optionally) lack, in comparison with argument *wh*-s, is [FOCUS]. Whereas argument *wh*-s come with an [uFOCUS] feature that needs to be valued, adjunct *wh*-s do not. Thus it is not merged until the cycle is complete, which to us means until the v̲P phase is complete, when the lexical sub-array pertaining to the v̲P phase is otherwise exhausted. On the other hand, if they come from the numeration with a [uFOCUS] feature, then it behaves just like other *wh*-s and get clefted.

CHAPTER 4

CONCLUSION

I have primarily adopted the basic structure proposed by Joseph (2000) and extended it to give a somewhat unified account of focus in Malayalam.

However, the points of departure from her account are in the merging of *wh*-s. When the vP is merged, normal argument *wh*-phrases come with a [uFocus] feature. The adjunct *wh*-s come with no such uninterpretable features and thus are merged postcyclically. In case of the argument *wh*-s, if it is D-linked, a Q_{EQ} particle is merged in a specifier of the vP to bind its trace *in situ*. If not, the focus marker *a:Nə* is merged higher up in the vP, which attracts and raises up the *wh*-. This *a:Nə* presumably moves up to T and gives effect to a phase sliding, pulling up the vP phase. After that the derivation continues to be built, and when the next strong phase head (which here is not the C but the D), the vP is spelt out. The derivation continues to the matrix CP when everything has been spelt out. The illusion of preverbal focus position is created in the PF during/ just before the linearization.

It would be appropriate now to discuss how Rooth's (1985 et seq) Alternative Semantics fares with respect to the story of focus syntax that we have put forth. The reason we take just this semantic account of focus for discussion here is because it is the most popular account of focus semantics so far.

Rooth (1985) views focus, marked by prosodic prominence, pitch accent or other phonological or phonetic correlates, as the indication of the presence of alternatives. He also assumes that it is marked on phrases as a syntactic feature in syntactic descriptions, to be assigned both semantic/pragmatic as well as phonetic/phonological interpretation. The model

of architecture of language is the standard EST organization of the Y-model of Chomsky and Lasnik (1977), with the levels of representation D-structure, S-structure, P(honological) R(epresentation) and LF. Following the organization of grammar in this manner, focus must be marked at S-Structure because it has correlated phonetic/phonological and semantic/pragmatic aspects.

Focus affects the truth value of a sentence. Alternative semantics takes semantic account of focus by adding an additional semantic value (called focus semantic value) to the ordinary semantic value of the sentence in order to evaluate the sentence. The focus semantic value of a phrase is the set of propositions obtainable from its ordinary semantic value by substituting in the position corresponding to the focus phrase. Thus, in order to derive the semantic value of a sentence containing focused constituents, first the ordinary and focus semantic values of the proposition are derived compositionally using function application.

Alternative Semantics does not differentiate between Identificational and Informational focus. Crucially, he does not move the focused constituent at all. For him, focus is marked by prosodic prominence, which renders it contrastive. The set of alternatives induced by the syntactic focus feature on the constituent is the set of potential referents (so to say) of the focused constituent, which is evaluated by the interpretive/ semantic component of language.

However, the Rooth account seems not to hold, at least for Malayalam. Look at the example number (28) in Chapter 2. The implied contrastivity can be negated. Thus it does not invoke a set of alternatives in the sense of Rooth.

To conclude, I have attempted to give an alternative to the present approaches to focus. For one, I have tried to give a unified mechanism that can derive both the ‘preverbal focus position’ and the cleft in Malayalam. However, there still remain some loose ends to be tied

up. But due to time and space constraints it will have to be left for further research. I have tried to answer most of the questions I raised in Chapter 2. Some of the questions that would need future research are:

(a) The actual status of the copula/focus marker. Many accounts of clefts in various languages show that crosslinguistically it is the specificational/equative copula of the language (if it has any) that are used as 'focus markers' in the cleft construction. Babu (2006) says that the Malayalam equative copula *a:Nə* plays the dual role of equative copula and focus marker in Malayalam.

(b) I still have not been able to account for some interesting things that happen in the Malayalam multiple *wh*- consruction. For instance, in multiple *wh*- questions in Malayalam, according to Jayaseelan (1989 et seq), the *wh*- phrases 'stack' together at the preverbal focus position (see below).

- ninn-e aarə eppo:L entinə tall-i
  2S-ACC who when why beat-PST
  'Who beat you when (and) why?'

Jayaseelan (2001) also shows that adjunct *wh*-s in a multiple *wh*- question always tend to occupy positions closer to the verb than argument *wh*-s:

- avan a:r-e entinə konn-u
  3SM who-ACC why kill-PST
  'Why did he kill whom?'
  *avan entinə a:r-e konn-u
  3SM why who-ACC kill-PST

(c) I have not been able to propose any particular explanation for this fact. More importantly, I also could not explain the reason for adjunct whs in preverbal position allowing wide scope interpretation of argument *wh*-s too.

Bibliography

Amritavalli, R., & Jayaseelan, K. A. (2005). Scrambling in cleft construction in Dravidian. In J. Sabel & M. Saito (Eds.), *The free word order phenomenon : its syntactic sources and diversity*, Studies in generative grammar (pp. 137–162). Berlin: Mouton de Gruyter.

Babu, H. (2006). Two Copulas in Malayalam revisited. *Proceedings of the First National Symposium on Modeling and Shallow Parsing of Indian Language (MSPIL)*. Presented at the First National Symposium on Modeling and Shallow Parsing of Indian Language (MSPIL), IIT Bombay.

Chernova, E. (2011). On Russian Wh- Questions. Presented at the 9th European Conference on Formal Description of Slavic Languages, abstract.

Chomsky, N. (1977). On wh- movement. In P. W. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal syntax*. New York: Academic Press.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press.

Chomsky, N. (1998). *Minimalist inquiries: the framework*. MIT, Dept. of Linguistics.

Chomsky, N. (2001). Derivation by phase. *Ken Hale: A life in language, ed. M. Kenstowicz* (pp. 1–52). Cambridge, MA: MIT Press.

Gallego, Á. J. (2006). Phase sliding. *Ms., UAB/UMD*. Retrieved from http://webs2002.uab.es/clt/membres/professors/gallego/pdf/Phase_Sliding.pdf

Grohmann, K. K. (2009). Focus on Clefts: A Perspective from Cypriot Greek. *Selected papers from the 18th ISTAL*. Presented at the ISTAL.

Gryllia, S. (2008). *On the nature of preverbal Focus in Greek: A theoretical and experimental approach* (PhD dissertation). Universitiet Leiden.

Hiraiwa, K., & Ishihara, S. (2002). Missing links: cleft, sluicing, and "no da" construction in Japanese. *MIT working papers in linguistics, 43*, 35–54. Retrieved from http://www.sfb632.uni-potsdam.de/~shin/papers/hiraiwa-ishihara2002-cleft.pdf

Horvath, J. (2010). “Discourse features”, syntactic displacement and the status of contrast. *Lingua*, *120*(6), 1346–1369. Retrieved from http://www.sciencedirect.com/science/article/pii/S0024384109001399

Jayaseelan, K. A. (1996). Question word movement to focus and scrambling in Malayalam. *Linguistic Analysis*, *26*, 63–83.

Jayaseelan, K. A. (2001). IP-internal topic and focus phrases. *Studia Linguistica*, *55*(1), 39–75. doi:10.1111/1467-9582.00074

Jayaseelan, K. A. (2002). Question movement and how some SOV languages handle it. *GLOW in Asia Colloquium*. Presented at the GLOW in Asia Colloquium, Taipei. Retrieved from http://glow.ling.nthu.edu.tw/4-3-1Jayaseelan.pdf

Jayaseelan, K. A. (2004). Question movement in some SOV languages and the theory of feature checking. *Language and linguistics*, *5*(5). Retrieved from http://140.109.150.41/eip/FILES/journal/2007.3.9.3945559.22377378.pdf

Jayaseelan, K. A. (2008a). Bare phrase structure and specifier-less syntax. *Biolinguistics*, *2*(1), 087–106. Retrieved from http://www.biolinguistics.eu/index.php/biolinguistics/article/viewArticle/18

Jayaseelan, K. A. (2008b). Topic, focus and adverb positions in clause structure. *Nanzan Linguistics*, *4*, 43–68. Retrieved from http://www.ic.nanzan-u.ac.jp/LINGUISTICS/publication/pdf/NL4-3-jayaseelan.pdf

Jayaseelan, K. A. (2010). Stacking, Stranding, and Pied-Piping: A Proposal about Word Order. *Syntax*, *13*(4), 298–330. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9612.2010.00141.x/full

Joseph, L. (2000). *Clefts in Malayalam* (unpublished MPhil dissertation). Jawaharlal Nehru University, New Delhi.

Kidwai, A. (1999). Word order and focus positions in Universal Grammar. In L. Tuller & G. Rebuschi (Eds.), *The grammar of focus* (pp. 213–244). Amsterdam: John Benjamins Publishing Company. Retrieved from http://books.google.co.in/books?hl=en&lr=&id=o159LyeiK-gC&oi=fnd&pg=PA213&dq=kidwai+a+scrambling&ots=c_Yj74VebT&sig=AJHadNlnUNC9IHda485Jf_N6mFs

Kidwai, A. (2000). *XP-adjunction in universal grammar: scrambling and binding in Hindi-Urdu*. Comparative Syntax. Oxford University Press, USA. Retrieved from http://books.google.co.in/books?hl=en&lr=&id=vqHWkWJrFjQC&oi=fnd&pg=PA3&dq=kidwai+a+scrambling&ots=aEVbW_zKu8&sig=DZAYqYAm2glKDSz4HmTXi5kU7LY

Kiss, K. É. (1995). *Discourse configurational languages*. Oxford University Press, USA. Retrieved from http://books.google.co.in/books?hl=en&lr=&id=o4bLibKAYTsC&oi=fnd&pg=PA3&dq=katalin+e+kiss+1995&ots=mL7sXolU-k&sig=4cnZvQ5TUdrsTEwfitO7MTk_y50

Kiss, K. É. (1998). Identificational focus versus information focus. *Language*, 245–273. Retrieved from http://www.jstor.org/stable/10.2307/417867

Kiss, K. É. (1999). The English cleft construction as a focus phrase. In L. Mereu (Ed.), *Boundaries of Morphology and Syntax*, AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE (pp. 217–230). John Benjamins Publishing Company.

Lee, C., Gordon, M., & Büring, D. (2008). *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation*. Springer.

Mathew, R. (2011). Recursion of FocP in Malayalam. *ms.*

Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. *The representation of (in) definiteness*, 98–129.

Pesetsky, D., & Torrego, E. (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (Ed.), *Ken Hale: A life in language*, CURRENT STUDIES IN LINGUISTICS SERIES (Vol. 36, pp. 355–426). MIT Press. Retrieved from http://202.194.48.102/englishonline/jxyj/xslw/readings%20in%20syntax/amsa-1-0800.pdf

Rizzi, L. (1997). The fine structure of the left periphery. Retrieved from http://dspacetest.cilea.it/handle/10278/1354

Rooth, M. (1985). *Association with focus* (PhD dissertation). GLSA, Dept. of Linguistics, University of Massachusetts, Amherst. Retrieved from http://ecommons.library.cornell.edu/handle/1813/3292.

Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics*, *1*(1), 75–116. Retrieved from http://www.springerlink.com/index/k57211207j40p176.pdf

Sobin, N. (2010). Echo Questions in the Minimalist Program. *Linguistic Inquiry*, *41*(1), 131–148. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/ling.2010.41.1.131

Stepanov, A. (2001). Late adjunction and minimalist phrase structure. *Syntax*, *4*(2), 94–125. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/1467-9612.00038/abstract

Szabolcsi, A. (1981). The semantics of topic-focus articulation. *Formal methods in the study of language*, *2*, 513–541.

Uriagereka, J. (1999). Multiple Spell-out. Retrieved from http://gagl.eldoc.ub.rug.nl/FILES/root/1997-40/06/GAGL-40-1997-06.pdf