FORMAL AND PHILOLOGICAL INQUIRIES INTO THE NATURE OF INTERROGATIVES, INDEFINITES, DISJUNCTION, AND FOCUS IN SINHALA AND OTHER LANGUAGES

BY

BENJAMIN MARTIN SLADE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Linguistics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor James Hye Suk Yoon, Chair
Professor Emeritus Hans Henrich Hock, Director of Research
Professor Peter Lasersohn
Assistant Professor Karlos Arregi, University of Chicago

Abstract

In this thesis I examine a variety of linguistic elements which involve "alternative" semantic values—a class arguably including focus, interrogatives, indefinites, and disjunctions—and the connections between these elements. This study focusses on the analysis of such elements in Sinhala, with comparison to Malayalam, Tlingit, and Japanese.

The central part of the study concerns the proper syntactic and semantic analysis of Q[uestion]-particles (including Sinhala *də*, Malayalam *-oo*, Japanese *ka*), which, in many languages, appear not only in interrogatives, but also in the formation of indefinites, disjunctions, and relative clauses. This set of contexts is syntactically-heterogeneous, and so syntax does not offer an explanation for the appearance of Q-particles in this particular set of environments.

I propose that these contexts can be united in terms of semantics, as all involving some element which denotes a set of "alternatives". Both *wh*-words and disjunctions can be analysed as creating Hamblin-type sets of "alternatives". Q-particles can be treated as uniformly denoting variables over choice functions which apply to the aforementioned Hamblin-type sets, thus "restoring" the derivation to normal Montagovian semantics. The treatment of Q-particles as uniformly denoting variables over choice functions provides an explanation for why these particles appear in just this set of contexts: they all include an element with Hamblin-type semantics.

However, we also find variation in the use of Q-particles; including, in some languages, the appearance of multiple morphologically-distinct Q-particles in different syntactic contexts. Such variation can be handled largely by positing that Q-particles may vary in their formal syntactic feature specifications, determining which syntactic contexts they are licensed in.

The unified analysis of Q-particles as denoting variables over choice functions also raises various questions about the proper analysis of interrogatives, indefinites, and disjunctions, including issues concerning the nature of the semantics of *wh*-words and the syntactic structure of disjunction.

As well, I observe that indefinites involving Q-particles have a crosslinguistic tendency to be epistemic indefinites, i.e. indefinites which explicitly signal ignorance of details regarding who or what satisfies the existential claim. I provide an account of such indefinites which draws on the analysis of Q-particles as variables over choice functions. These pragmatic "signals of ignorance" (which I argue to be presuppositions) also have a further role to play in determining the distribution of Q-particles in disjunctions.

The final section of this study investigates the historical development of focus constructions and Q-particles in Sinhala. This diachronic study allows us not only to observe the origin and development of such elements, but also serves to delimit the range of possible synchronic analyses, thus providing us with further insights into the formal syntactic and semantic properties of Q-particles.

This study highlights both the importance of considering various components of the grammar (e.g. syntax, semantics, pragmatics, morphology) and the use of philology in developing plausible formal analyses of complex linguistic phenomena such as the crosslinguistic distribution of Q-particles.

Thesis Supervisor: Hans Henrich Hock

Title: Professor Emeritus of Linguistics; Sanskrit; Classics; Germanic Languages & Literatures; Russian, East European, & Eurasian

मत्प्रियतमार्थे

ती सुन्दर सुन्दर सितारा जडिएका
सुन्दर रेशमी घुम्टोलाई घृणा गर्छु,
घृणा गर्छु फूलहरूको सुगन्धी समीरलाई पनि,
म जूनको पातलो पर्दालाई पनि घृणा गर्छु,
घृणा गर्छु, प्रिय, तिमीले हाँस्तै हाँस्तै
गाएको प्रेमको मधुर गीतलाई पनि,
किनभने
किनभने ती बीचका छेका बनिदिन्चन्
मेरो ओठ र तिम्रा ओठका ॥

"घृणा गर्छु" (वि॰ सं॰ २०२४)
–बालकृष्ण सम
(बालकृष्ण शमशेर जंग बहादुर राणा)

ඇදී සසලප සෙ සුරිදුහු
සද්ඹඩලෙ පවතු වාදහසක් එක්
දවසක් සෙ මෙනෙහි මා

Sigiri graffito no. 135 (late ninth/early tenth century A.D.)

Acknowledgements

That I ended up writing a dissertation about Sinhala is the result of a serendipitous intersection of interests in Indo-Aryan languages, (morpho)syntactic change, and the semantics of questions. That a dissertation on Sinhala should be the result of serendipity is appropriate given the word's etymological origin:— Serendip being a Persian word for "Sri Lanka".[1] The English form *serendipity* itself was coined by Horace Walpole to refer to "happy and unsought after discoveries" in reference to the heroes of the tale *The Three Princes of Serendip*, who continuously made discoveries in this manner.[2]

And, indeed, serendipity played a role not only in the initial genesis of the dissertation, but also in certain aspects of the development of the analyses herein. The distribution of the particles *də* and *hari* in modern colloquial Sinhala disjunctions remained mysterious to me for a long time; it was only in the examination of the properties of indefinites formed with *də* and *hari*—which presented their own set of mysteries—that I stumbled across what I believe is the solution to the earlier mystery of the disjunctives.

However serendipity alone does not a dissertation make, and there are numerous people whose contributions to this study, direct and indirect, I would like to acknowledge.

[1] The Persian word *sarandip* or *serendip*, referring to Sri Lanka, represents a borrowing from an Indo-Aryan word ultimately related to Sanskrit *suvarṇa-dvīpa* "golden island".

[2]

> . . . I must tell you a critical discovery of mine apropos: in an old book of Venetian arms, there are two coats of Capello, who from their name bear a hat; on one of them is added a fleur-de-lis on a blue ball, which I am persuaded was given to the family by the Great Duke, in consideration of this alliance; the Medicis, you know, bore such a badge at the top of their own arms. This discovery I made by a talisman, which Mr. Chute calls the Sortes Walpolianae, by which I find every thing I want, *à pointe nommée*, whenever I dip for it. This discovery, indeed, is almost of that kind which I call Serendipity, a very expressive word, which, as I have nothing better to tell you, I shall endeavour to explain to you: you will understand it better by the derivation than by the definition. I once read a silly fairy tale, called *The Three Princes of Serendip*; as their Highnesses travelled, they were always making discoveries, by accidents and sagacity, of things which they were not in quest of: for instance, one of them discovered that a mule blind of the right eye had travelled the same road lately, because the grass was eaten only on the left side, where it was worse than on the right—now do you understand Serendipity? One of the most remarkable instances of this accidental Sagacity, (for you must observe that no discovery of a thing you are looking for comes under this description,) was of my Lord Shaftsbury, who, happening to dine at Lord Chancellor Clarendon's, found out the marriage of the Duke of York and Mrs. Hyde, by the respect with which her mother treated her at table. . . " [Horace Walpole, 4th Earl of Orford, Letter To Sir Horace Mann (Bt, KB), Arlington Street, Jan. 28, 1754 (*The Letters of Horace Walpole*, vol. 2, Letter 90).]

The Three Princes of Serendip appears in Europe first in the form of Michele Tramezzino's Italian *Peregrinaggio di tre giovani figliuoli del re di Serendippo* (Venice, 1557), adapting an earlier Persian story from the *Hasht Bihisht* "Eight Paradises" (1302) of Amir Khusrau, an Indian musician and scholar, who composed in Hindavi (an early precursor of Hindi/Urdu) as well as Persian.

The "mule" referred to in Walpole's letter is a misremembering of the camel of the original story. The princes of Serendip, upon meeting a camel driver who has lost a camel, immediately ask him if the camel happened to have been lame, blind in one eye, missing some teeth, and further had carried honey and butter and a pregnant woman. The camel driver, astonished, assumes that they must have stolen the camel, and has them thrown in prison. The princes are later released, and reveal that they inferred the attributes of the lost camel through a number of observations: that the camel was blind in one eye because the grass was only eaten on one side of the path, that the camel was missing a tooth from the half-chewed nature of the grass, the lameness from the tracks, and so on and so forth. A similar scenario appears in Umberto Eco's *The name of the rose* (*Il nome della rosa*, Bompiani 1980; English translation by William Weaver, Harcourt, 1983) in the case of the abbot's missing horse, about which the main character of the novel, William of Baskerville, infers numerous characteristics, despite never having laid eyes upon it.

First and foremost, of course, the members of my thesis committee played a major role in guiding my dissertation research, and each of them have helped me to formulate, test, discard, and reformulate hypotheses, and to integrate the many different elements of this research project.

Hans Henrich Hock has been my adviser for the last six years, and has been my primary guide not only for the present study, but also for numerous earlier studies, ranging from the synchronic and diachronic analysis of the morphosyntax of Indo-Aryan compound verbs to the reconstruction of Proto-Indo-European dragon-slaying formulae. Hans is the reason I chose to pursue graduate studies at UIUC, and the opportunity of working with him has made me very glad of this choice. His breadth of knowledge and keen insight into language have been invaluable to me not only for the diachronic aspects of this study, but for the synchronic aspects as well: Hans has provided useful comments and discussion on all facets of this study, from semantics to syntax to morphology to phonology, and his advice has often proved key to the unravelling of knotty problems. Hans has read through and provided comments on innumerable drafts and pieces of this study, in various stages of development, and this dissertation is much the better for his patient guidance, both in terms of content and coherence.

James Yoon served as chair of my thesis committee, and has long been my guide through the rather dense forest of Chomskyan syntactic theory. His advice on syntactic and morphological issues, as well as his vast knowledge of the previous literature on the analysis of SOV languages, has been of great assistance in the non-semantic side of the analysis. Discussions with James always resulted in my gaining a fresh perspective on various troublesome issues.

Peter Lasersohn kept me honest in my semantic formalism, and guided me through the analysis of a wide range of semantic issues. Given the centrality of semantics in the thesis, his advice and corrections have been invaluable in the development of some of what I believe are the most exciting parts of this study. My proposals in some cases involve rather complicated formalisms, which Peter has always managed to wade through in order to spot inconsistencies and errors. Discussions with Peter have always left me much clearer on various points of semantic theory, and his suggestions on how to approach semantic issues have helped me immeasurably in the construction of novel semantic analyses of a number of perplexing linguistic data.

Karlos Arregi provided invaluable discussion of many of the central issues of this thesis. I always came away from these discussions with new ideas and, frequently, with more questions to ponder. Karlos has a great ability to cut through the tangle of complexities to get at the heart of the issue, which not only frequently saved me heaps of time which would otherwise have been spent in pursuing red herrings, but also served to point me in the right direction for the construction of meaningful analyses.

I would be greatly remiss if I did not offer my many thanks to my Sinhala consultants: Maheshi Dassanayake (thanks for attending my defence!), Viveka Kudaligama, and Indu Rupassara (as well as Sanjeewa Rupasinghe who provided assistance in the earlier phases of the study) for all of their patient help in providing grammaticality judgements, translations, and explanations of aspects of the Sinhala language which have been crucial to this study.

Other people whose comments, suggestions, and/or kind sharing of data have greatly benefited this thesis are Jim Gair, Paul Hagstrom, Seth Cable, Malawenna Gamage Lalith Ananda, and Rohana Seneviratne. Paul Hagstrom I have known since his time as a postdoc at JHU many years ago (when I was a graduate student there), and he has offered insightful discussion not only on the issues connected with the present study, but also on preceding projects. Also, I would like to note that the comments and discussion of Kai von Fintel and two anonymous reviewers for *Semantics & Pragmatics* were extremely helpful in the case of the analysis of epistemic indefinites which is incorporated here.

I would also like to thank other members of the UIUC Department of Linguistics who have helped to shape my thinking about language and linguistics, including Richard Sproat, Rakesh Bhatt, Marina Terkourafi (also a co-author), Abbas Benmamoun, Ryan Shosted (another LaTeX aficionado), Jennifer Cole, and Mithilesh Mishra.

In addition my fellow graduate students have also helped to make life in Urbana/Champaign both intellectually- and socially-stimulating. In particular, I would like to thank—for numerous interesting discussions both linguistic and non-linguistic, both in the college and over dinners, drinks, and camp-fires, and for their general camaraderie—Matt Garley (co-author, co-editor, co-conference organiser, and sharer of interesting books); Lisa Pierce (thrower of dinner parties and sharer of interesting cats); Soondo Baek and Eunah Kim; Adriana Molina Muñoz (co-instructor for LING210, co-Indophile, and another Hock advisee); Liam Moran; Archna Bhatia; Vandana Puri; Jill Hallett; Andrew Fister (who introduced me to Ubuntu); Karen Lichtman; Sarah Grafton Simeziane; Jennifer Cramer; Erica Britt; Tim Mahrt; Eun-Kyung Lee; Amanda Huensch; Eunice Chung; Charles LaWarre; Jenna Chi (co-author); Erin Rusaw (hoster of Minecraft servers); Daniel Carr; Shawn Chang; Gary Linebaugh; Hahn Koo; Aimee Alnet; Ju Hyeon Hwang; Marco Shappeck (another former Hock advisee); and Indranil Dutta (another former Hock advisee).

My linguistic studies, however, began in Baltimore, and I would like to thank various people at the Johns Hopkins University who are to blame for my interest in language and linguistics: my undergraduate advisor, the poet/scholar Allen Grossman, who kindled my interest in formal structures of language; Andrew Kelly, in whose Sanskrit classes I got my first real taste of linguistics via *sandhi* rules and Pāṇinian grammar; Gonzalo Rubio, from whom I took my first historical linguistics class; and Luigi Burzio (my MA advisor), Géraldine Legendre, Paul Smolensky, Bob Frank, and David Lightfoot (then at the University of Maryland at College Park)—who all served to further my interest in linguistics and guide me as a neophyte linguist; as well as my then fellow graduate students in the Cognitive Science Department: Colin Wilson, Matt Goldrick, John Hale, Lisa Davidson, Uyen Le, Fero Kuminiak, and Marina Todorova (and Paul Hagstrom, then a postdoc at JHU).

I am also grateful to the School of Literatures, Cultures and Linguistics at the University of Illinois for a Dissertation Completion Fellowship (which has made the last ten months much more productive than they might have been otherwise), the Center for Global Studies at the University of Illinois for Foreign Language and Area Studies (FLAS) Fellowship (for advanced training/research in Hindi), as well as the UIUC Department of Linguistics and the Beckman Institute for Advanced Science and Technology for additional financial support. As well, I like to thank the various members of the administrative staff of the Department of Linguistics and the School of Literatures, Cultures and Linguistics who have always been extremely kind and helpful, including Pat Gallagher (now retired), Cathy Penny, Marita Romine, Lynn Stanke, Geraldine Moore, and especially Mary Ellen Fryer.[3]

I would also like to offer thanks to a number of people who have influenced my life and academic career in various (positive) ways. This includes my friends from my early academic days in Baltimore: Bob Mills, Claire Whitner, Omar Alquaddoomi, Austin Lin; my friends from Shillong: Awadesh Mishra (younger brother of Mithilesh Mishra; it's a small world), Michael Nongrum and Imnukshilla Shylla; as well as my long-time friend from Pennsylvania, Jason Whetstone.

I would also like to thank the late pandit Chavilal Sharma (my dharmic guru) and my extended Nepalese family, Hem Bahadur Chettri and Rukmani Chettri (parents-in-law), Bikram Chettri and Vicky Chettri (brothers-in-law), all of who made me feel very welcome and at home during my year-and-a-half sojourn in India, spent mainly in the lovely hill-station of Shillong, Meghalaya (Skt. "Abode of Clouds").

Special thanks to my family, including my brothers, Shaun Slade and Peter Slade, and especially my parents, John Slade and Doris Slade, for their love and support during my long, and often turbulent, academic training.

And finally, to my wife, thanks for everything.

[3]I would also like to thank TEX/LATEX 2_ε, GNU Emacs (and other GNU free "as in free speech, not free beer" software), and Linux, and their initial creators (Donald E. Knuth, Richard M. Stallman, and Linus Torvalds, respectively) and innumerable contributors (including the authors of the AUCTEX package for Emacs), for making the technical aspects of the production of this study immensely easier and more enjoyable than they would have been otherwise. Thanks also to the Department of Physics at the University of Illinois for the creation of a LATEX template/stylefile for UIUC dissertations, and to the Libertine Open Fonts Projekt for the creation of the Linux Libertine font, the primary font face used in this document.

Table of Contents

List of Abbreviations

In glosses:

1 = 1st person
2 = 2nd person
3 = 3rd person
A = *-a*/*-ā* verbal suffix ("neutral")
ABL = ablative (case)
ACC = accusative (case)
ADJ = adjective
ADV = adverb
ANIM = animate
ART = article
COMP = complementiser
COND = conditional
CONV = converb
DAT = dative (case)
DEF = definite
E = *-e*/*-ē* verbal suffix ("focussing")
EMPH = emphatic (particle)
F = focus/focussed
FEM = feminine
FUT = future
GEN = genitive (case)
GER = gerund
IMPV = imperative
INANIM = inanimate
INDEF = indefinite
INF = infinitive
INST = instrumental (case)
LOC = locative (case)
MASC = masculine
NEG = negation/negative
NEU = neuter
NMLZ = nominaliser
NOM = nominative (case)
PART = particle
PAST = past
PERF = perfect
PERM = permissive
PL = plural
POL = polite (form)
PRES = present
PTCP = participle
Q = Q(uestion)-particle
REL-PRON = relative pronoun
QUOT = quotative
SG = singular
SUBJ = subjunctive
VN = verbal noun
VOC = vocative (case)

Languages:

CS = Classical Sinhala
LS = Modern Literary Sinhala
MCS = Modern Colloquial Sinhala
OS = Old Sinhala
OT = Old Tamil
Skt. = Sanskrit
SLT = Sri Lankan Tamil
TNT = Tamil Nadu Tamil

References to texts:

Ama. = *Amāvatura*: Kodagoda (1967)
BhG = *Bhagavadgītā*: van Buitenen (1981)
BhP = *Bhāgavata-Purāṇa*
BJT = Sri Lanka Buddha Jayanti Tripitaka Series [http://www.metta.lk/sltp/index.html]
DhA = Buddhaghosa's commentary on the *Dhammapada*
Panc. = *Paṅcatantra*
PST = Pali Text Society
S.G. = *Sigiri Graffiti*: Paranavitana (1956)
RV = *Ṛg Veda*: Bandhu (1963–1966)

Chapter 1

Introduction

That day I could not refrain from questioning him further about the matter of the horse.
'All the same,' I said, 'when you read the prints in the snow and the evidence of the branches, you did not yet know Brunellus. In a certain sense those prints spoke of all horses, or at least all the horses of that breed. Mustn't we say, then, that the book of nature speaks of us only of essences, as many distinguished theologians teach?'

—Adso of Melk (*chela*) to William of Baskerville (*guru*),
regarding William's identification of the abbot's
missing horse, from Umberto Eco's *The name of the rose*

1.1 Overview

This study focusses on the formal synchronic syntactic and semantic analysis and the historical development of "Question-particle" [or Q-particles] in Sinhala, an Indo-Aryan language spoken in the island nation of Sri Lanka, with comparison of the Q-particles of Sinhala with Q-particles in other languages, including Malayalam, Japanese, and Tlingit.

In many languages, including Sinhala, these so-called Q-particles appear not only in interrogative contexts, but also in disjunctions and the formation of indefinite pronouns.[1] The particles which appear in these latter two contexts often have a surface form identical to that of the particles which appear in interrogative contexts, and therefore ideally we should like to be able to provide a unified analysis which accounts for the crosslinguistic tendency for Q-particles to appear in this syntactically heterogeneous set of contexts.

Unsurprisingly, there are a variety of language-specific differences in the distribution of Q-particles which also need to be accounted for. Thus while in Malayalam there is single Q-particle which appears in yes/no- and alternative-questions, disjunctions, and indefinites (but which is absent, in the modern language, from *wh*-questions); in Sinhala, Japanese, and Tlingit we find multiple, phonologically-distinct Q-particles. In these latter three languages, the distribution of phonologically-distinct Q-particles is, however, determined by rather different factors: e.g., in Japanese only the Q-particle *ka* may appear in all of the Q-particle environments (interrogatives, disjunctions, indefinites), the other four Q-particles in Japanese are confined to interrogatives, and in all cases politeness plays a role in which Q-particle is employed; in Tlingit, there is no overlap in the environments in which the three different Q-particles appears: only *gé* appears in yes/no and alternative questions, only *sá* in *wh*-interrogatives and

[1] In some languages, including certain stages of Sinhala, Q-particles also appear in the formation of relative clauses. The analysis of relative clauses involves a number of complexities and largely falls outside of the scope of this study, though I do offer some brief remarks on the topic in Chapter 8.

wh-based indefinites, only *khach'u* in non-interrogative disjunctions. In Sinhala, the Q-particle *hari* is employed in non-interrogative disjunction, and the formation of a *wh*-based indefinite pronoun; *də* appears in all interrogatives (*wh*, yes/no, alternative) as well as in the formation of another *wh*-based indefinite pronoun; thus there is some overlap in the distribution of these two Q-particles.

Representative examples of Q-particle constructions are provided below.

In Japanese, there is one Q-particle, *ka*, which may appear in all of the possible Q-particle environments: interrogatives (1), indefinites (2), and disjunctions (3).

(1) a. gakkoo-ni ik-imas-u ka?
school-to go-POL-PRES ***ka***
'(Are you) going to school?' (Yoshida & Yoshida 1996)

b. John-ga nani-o kaimasita **ka**?
John-NOM what-ACC bought.POL ***ka***
'What did John buy?' (Hagstrom 1998: 15)

(2) dare-**ka**-ga hon-o katta.
who-**ka**-NOM book-ACC bought.
'Someone bought books.' (Kuroda 1965: 97)

(3) John-**ka** Bill-**ka**-ga hon-o katta.
John-***ka*** Bill-***ka***-NOM book-ACC bought.
'John or Bill bought books.' (Kuroda 1965: 85)

As noted above, the other Q-particles of Japanese (*no*, *kai*, *ndai*, *kadooka*) are confined to interrogative environments, and in these environments may serve as alternatives to *ka*—such alternations seem to be largely based on considerations of politeness (*ka* is the most polite form, and *kai* and *ndai* the most informal, with the latter being appropriate only in male speech).

The Malayalam Q-particle *-oo*, like Japanese *ka*, also appears in (yes/no and alternative) interrogatives (4), indefinites (5), and disjunctions (6), but—in contrast to Japanese—not in *wh*-questions (at least in the modern language), as shown in (7).

(4) a. John wannu-**(w)oo**?
John came-***oo***
'Did John come?' (Jayaseelan 2001: 67)

b. John wannu-**(w)oo**, illa-**(y)oo**?
John came-***oo***, not-***oo***
'Did John come, or not?' (Jayaseelan 2001: 67)

(5) ñāṉ iruṭṭ-il ār-e-**(y)oo** toṭṭu
I darkness-in who-ACC-***oo*** touched
'I touched somebody in the dark.' (Jayaseelan 2001: 66)

(6) Mary John-ine-**(y)oo** Bill-ine-**(y)oo** cumbiccu
Mary John-ACC-***oo*** Bill-ACC-***oo*** kissed
'Mary kissed John or Bill.' (Jayaseelan 2008: 3)

(7) ārə(*-*oo*) wannu?
who(*-*oo*) came
'Who came?' (Jayaseelan 2001: 67)

In (modern colloquial) Sinhala, the Q-particle *də* appears in all types of questions (8), and may also occur in the formation of indefinites (9a), but a distinct Q-particle, *hari*, may also be used in the formation of indefinites as shown in (9b), and also appears in non-interrogative disjunctions[2] as in (10a), while *də* is not possible in this latter environment as shown by example (10b).

(8) a. Sunil ee potə kieuwa **də**?
Sunil that book read.PAST.A ***də***
"Did Sunil read that book?"

b. Sunil monəwa **də** kieuwe?
Sunil what ***də*** read.PAST.E
"What did Sunil read?"

c. Sunil **də** Ranjit **də** ee potə kieuwe?
Sunil ***də*** Ranjit ***də*** that book read.PAST.E
"Was it Sunil or Ranjit who read that book?"

(9) a. Kau **də** ee potə kieuwa
who ***də*** that book read.PAST.A
"Someone read that book."

b. Kauru **hari** ee potə kieuwa
who ***hari*** that book read.PAST.A
"Someone read that book."

(10) a. Sunil **hari** Ranjit **hari** ee potə kieuwa.
Sunil ***hari*** Ranjit ***hari*** that book read.PAST.A
"Sunil or Ranjit read that book."

b. *Sunil **də** Ranjit **də** ee potə kieuwa.
Sunil ***də*** Ranjit ***də*** that book read.PAST.A
"Sunil or Ranjit read that book."

Tlingit, like Sinhala, displays the use of a variety of Q-particles, but with a rather different distribution, discussed in detail later in this study. Table 1.1 provides an overview, including additional data discussed in Chapter 2.

	Mod. Sinh	**Old Mal**	**Mod Mal**	**Tlin**	**Jap**
y/n-ques.	də	-oo	-oo	gé	ka, no, kai, kadooka
wh-ques.	də	-oo	—	sá	ka, no, ndai
wh-indef.	də (aff.), hari (aff.), vat(neg.)	-oo	-oo	sá	ka
decl. disj.	hari (aff.), vat (neg.)	-oo	-oo	khach'u	ka
interr. disj.	də	-oo	-oo	gé... gwáa	[ka]

Table 1.1: Distribution of Q-particles in Sinhala, Malayalam, Tlingit, and Japanese

[2]I use the terms "non-interrogative" and "declarative" disjunctions to refer to disjunctions which are not alternative questions. "Non-interrogative"/"declarative" disjunctions can in fact appear in interrogative contexts, e.g. in yes/no-questions containing disjunctions, such as "Do you want tea or coffee?" (where the appropriate answers are "yes" or "no").

Thus, a comprehensive account of Q-particles requires an analysis which captures both the crosslinguistic similarities and differences.

1.1.1 Importance of both semantic and syntactic analysis

Beyond uncovering new empirical data and providing a formal analysis which accounts for them, this study also highlights the necessity of analysis which considers multiple components of the grammar. That is, many linguistic phenomena are given competing accounts relying on one or other component of the grammar:— for example, *wh*-fronting in English is argued alternatively as being semantically-based (quantifier raising) or syntactically-based (movement triggered by formal syntactic features). However, the current study demonstrates the importance of an analysis involving the interaction of various components of the grammar, including semantics, syntax, pragmatics, and morphology.

Specifically, I argue that the crosslinguistic tendency for Q-particles to appear in interrogatives, indefinites, and disjunction has a semantic basis, while language-specific difference in possible Q-particle environments, as well as the presence in some languages of multiple Q-particles, requires an explanation in terms of formal syntax (and, in a minority of cases, also formal pragmatics).

Furthermore, Q-particles interact with a wide range of grammatical phenomena, and thus their analysis has important implications for the proper treatment of *wh*-words (and interrogatives more generally), indefinites, and disjunction. Thus, in addition to adding to empirical coverage of previously unnoticed phenomena (e.g. the pragmatic differences between *də*-type and *hari*-type indefinites, discussed below), this study also makes an important contribution to our understanding of the formal properties of interrogatives, indefinites, and disjunction. For example, I present a novel analysis of syntax and semantics of disjunction—drawing on certain aspects of den Dikken (2006) and Alonso-Ovalle (2006)—but which, moreover, is a natural consequence of the more general semantic and syntactic analysis I propose for Q-particles.

Another contribution this study makes involves the connection between "epistemic indefinites"—indefinites which explicitly signal the lack of certain information concerning who or what satisfies an existential claim—and Q-particles. While the presence of epistemic indefinites has been noted in several languages (see Haspelmath 1997, Alonso-Ovalle & Menéndez-Benito 2003, 2010, Moore 2003), relatively little attention has been paid to the fact that in languages which employ epistemic indefinites alongside "plain indefinites" (e.g. Sinhala, Japanese, Malayalam), it is the epistemic indefinites which involve a Q-particle. Furthermore, consultation with my Sinhala informants has revealed the fact that some languages may employ a variety of different epistemic indefinites, which vary in the degree of ignorance they convey. For example, the indefinite formed with *də* in (9a), repeated below as (11a), signals a greater degree of ignorance regarding the identity of the person in question than does the indefinite formed with *hari* in (9b), repeated below as (11b).[3]

(11) a. Kau **də** ee potə kieuwa
who ***də*** that book read.PAST.A
"Someone read that book."

b. Kauru **hari** ee potə kieuwa
who ***hari*** that book read.PAST.A
"Someone read that book."

[3]The alternation between *kau* and *kauru* is simply morphophonological in nature, with *kau* being the form the pronoun takes if and only if it is immediately followed by *də*. That is, the pragmatic differences between (11a) and (11b) are due to the choice of Q-particle and not the surface form of the pronoun.

1.1.2 The Importance of Historical Data

In addition to developing a formal syntactic and semantic account of Q-particles, this study also considers the historical processes involved in the rise of Q-particles and changes in their distribution. In this section of the dissertation, I concentrate primarily on Sinhala: for here we are fortunate in that Sinhala has a more or less continuous textual tradition dating back to the 8th century A.D. (prior to the 8th century, we also find short inscriptions), thus affording an excellent opportunity for observing historical development of Q-particles, and of the 'focus construction' with which Q-particles are frequently associated. These are the two aspects of Sinhala grammar which I examine from a diachronic perspective; relevant prior scholarship on these topics does exist: from generative grammar-informed perspectives, Gair (1986[1998]b) and Paolillo (1994) examine the development of Sinhala focus (including discussion, to some extent, of interrogatives); more purely descriptive accounts which include some discussion of focus and interrogative structures in early Sinhala are provided by Geiger (1938), Paranavitana (1956), Reynolds (1964), and Wijemanne (1984). While all of these prior studies have crucially informed my understanding of the historical development of Sinhala, the goals of this study required extensive examination of primary texts as well.

I analyse what I consider to be four distinct stages of Sinhala: (1) Old Sinhala, represented by the graffiti texts on the Mirror Wall at Sihigiri (ca. 8th–10th c. A.D.); (2) Classical Sinhala, represented largely by translations and commentaries on Pāli Buddhist texts (ca. 12th–15th c. A.D.); and two varieties of modern Sinhala: (3) Modern Literary Sinhala, which differs from Classical Sinhala, but retains a number of archaisms such as overt subject-verb agreement morphology; and (4) Modern Colloquial Sinhala. The latter two varieties co-exist in a diglossic relationship, with the literary variety being employed in written and formal situations, but the general archaic nature of the literary variety justifies its treatment as representing an earlier variety than does the colloquial.[4]

The historical analysis of this thesis reveals the following overall picture. The Q particle *də* appears to originate in Sinhala in alternative questions.[5] In the oldest substantial Sinhala texts (8th–10th c.), *də* appears most frequently in yes/no questions, more rarely in *wh-questions*, and not at all in the formation of indefinites. The particle *hari* (see above (9b) and (10a)) is of more obscure origins, but appears to be somehow connected with another particle, *hō*. *Hō* in early Sinhala appears only in the formation of non-interrogative disjunctions, first appearing in the formation of indefinites only in the modern literary Sinhala period.

Thus the examination of the development of Sinhala offers an excellent opportunity to observe how Q-particles come to have the distributions that they do. Here the diachronic data can illuminate the synchronic analysis: understanding changes in the distribution of Q-particles helps us to identify which properties of Q-particles are language-particular and which are universal. Now, arguably, (synchronic) crosslinguistic data is just as informative in this respect. However, diachronic data can be more useful in the determination of the precise nature of the language-particular properties. Simple crosslinguistic data provide us only with unrelated snapshots of grammars, but if we examine different stages of a single language, we may observe how one grammar changes into another. Being able to observe changes in a grammar, of course, provides valuable information about the nature of that grammar, and, moreover, about the nature of universal grammar.

The importance of historical data for synchronic analysis can take on a most tangible form in the event of there being competing formal descriptions of a language. Here diachronic evidence can act as a metric for the evaluation of competing synchronic analyses. Suppose that L_n is a synchronic stage of a language—or rather some subset of properties of a synchronic stage of a language—and that there exist two competing formal theories, A and B, both of which are consistent with the observed characteristics of L_n. Suppose that we also have evidence of a prior

[4] On Sinhala diglossia, see Gair (1968[1998], 1986[1998]a) and Paolillo (1992).

[5] The original distribution of *də* is inferred from the distribution of its cognates in the historically prior languages (i.e. Sanskrit *utā́ho*, Pāli *udāhu*); cf. the examples in Böhtlingk & Roth (1855–1875) for Vedic, Speijer (1886) for Classical Sanskrit, and Rhys Davids & Stede (1921–1925) for Pāli. On the relationship between Vedic, Classical Sanskrit, Pāli, and other early Indo-Aryan languages, see Hock & Pandharipande (1978).

stage of L_n, which we can refer to as L_{n-1}. If theory A can account for the transition between L_{n-1} and L_n by positing a change in a single parameter/feature/property, then theory A—all else being equal—is preferable to theory B, in which L_{n-1} and L_n differ from each other in terms of several unrelated parameters/features/properties. More generally, formal synchronic analyses which allow for the diachronic evolution of a language to be described in terms of plausibly motivated changes are to be favoured over those which do not.

Thus the historical examination of various stages of Sinhala, and the changes which took place between these stages, plays an important role in delimiting the range of possible synchronic analyses of modern Sinhala.

1.1.3 Progress beyond previous studies

This study is, of course, by no means the first to discuss Q-particles and their proper formal analysis; the syntactic and semantic properties of Q-particles have been examined by a number of linguists (to cite but a few: Baker 1970, Karttunen 1977, Cheng 1991). Nor is it the first to examine Q-particles in Sinhala; important contributions having been made by Gair (1983[1998], 1986[1998]b), Gair & Sumangala (1991), Kishimoto (1992, 2005), and Kariyakarawana (1998). Moreover, the semantic analysis I adopt owes much to the previous studies of Hagstrom (1998) and Cable (2007)—who, in fact, also consider some of the same languages examined here.[6]

However, none of these previous studies provides a complete explanation of why Q-particles appear cross-linguistically in this particular set of contexts (interrogatives, indefinites, disjunctions). Cable (2007: 72–73n40) suggests—on the basis of Tlingit data—that the Sinhala Q-particle *də* and the Japanese Q-particle *ka* which appear in yes/no-questions and disjunctions are semantically and morphosyntactically distinct from the Q-particles *də* and *ka* which appear in *wh*-questions and indefinites, with their apparent identity being simply the result of homophony (though with a possible, though unspecified, historical basis). However, the fact that there exists apparent identity between the *wh*-related and yes-no/disjunction particles not only in Sinhala and Japanese, but also in Malayalam would seem to cast a degree of doubt on this hypothesis. The null hypothesis (and the more interesting hypothesis), I suggest, is that these phonologically-identical elements are also identical semantically and morphosyntactically. *Entia non sunt multiplicanda praeter necessitatem*: let us not suppose the existence of homophonous particles unless we uncover compelling evidence for such multiplicity. In this spirit, I posit that Sinhala *də*, in all of the environments in which it appears, is the same element, with the same syntactic and semantic properties, and likewise for Japanese *ka*, Malayalam *-oo*.

Accounting for underlying similarities of Q-particles crosslinguistically by positing a unified semantic analysis thus requires some means of addressing crosslinguistic differences. This is accomplished, as mentioned previously, largely in terms of language-specific differences in the formal syntactic features borne by Q-particles.

The following section provides a brief overview of previous studies of relevance to this dissertation, noting both their contributions and shortcomings; the latter of which the present study seeks to address, and to the former of which it is greatly indebted.

1.2 Previous studies

1.2.1 General studies of Sinhala Interrogatives & Focus

Jim Gair and his students have conducted numerous studies of various aspects of Sinhala syntax within the generative framework.

[6] Too late in the dissertation process did I discover Ginsburg's (2009) thesis treating Q-particles in Japanese and other languages, so few insights from that research could be incorporated in the present study.

Gair (1983[1998]) (cf. Gair 1970) investigates the synchronic properties of focus and interrogatives in Sinhala, showing that finite clauses are not islands for wh-movement or focus movement/association, though complex noun phrases are. Under this analysis, which following Kariyakarawana (1998), we may refer to as the "overt Wh-movement" analysis, focussed elements move overtly into a FOC position (some sort of complementiser-related position, thus making focus movement parallel to the overt movement of *wh*-words to COMP). Gair & Sumangala (1991) and Kariyakarawana (1998) (though differing in some details) argue that two different types of focus should be recognised: C-focus (cleft-focus), where the focussed element is "clefted", involving a biclausal structure; and E-focus (emphasis-focus), where a morphological focus marker immediately follows the focussed element, involving a monoclausal structure. I will argue that modern colloquial Sinhala focus constructions, appearances to the contrary, no longer involve any actual "clefting" (in the sense of being a biclausal construction), though ultimately deriving from "cleft" structures.[7]

Kishimoto (2005) argues that the Sinhala Q-particle *də* serves as a scope-assigner, in the sense that its position determines the scope of its associated *wh*-word. The particle *də* may undergo LF movement (movement post-Spellout/Transfer), but this movement respects island conditions (such as the Complex NPs).[8]

These studies are of great importance in terms of establishing the basic patterns for Sinhala interrogative and focus constructions; however, they are primarily syntactic in nature and thus not provide a full account of the properties of Sinhala interrogatives and focus.

1.2.2 Analyses of Q-particles

There are several analyses which treat *ka* (and similar particles) in terms of choice functions. These analyses all rely on some version of Hamblin's (1973) treatment of *wh*-words as denoting SETS of individuals. Hamblin (1973) proposes an extension to the formal semantic analysis of English developed by Montague (1970a,b, 1973) which allows for a formalisation of the semantics of *wh*-interrogatives. Hamblin (1973) notes that *wh*-words like *who* and *what* behave syntactically like proper names; that is, *who* may substitute a proper name like *John* and *what* may substitute a proper name like *Syntactic Structures*, see (12) and (13).

(12) John read *Syntactic Structures*.

(13) a. Who read *Syntactic Structures*?
b. Who read what?

However, semantically, *wh*-words are very different from proper names. So while *John* and *Syntactic Structures* denote individuals, it makes little sense to try to analyse *who* and *what* in a like manner. But *who* and *what* can be treated as denoting sets of individuals. Thus, in some pragmatic context, *who* might denote the set of individuals {John, Bill, Mary, Kim}.[9]

This treatment of *wh*-words as denoting sets in fact produces a reasonable semantics for *wh*-questions. While a sentence like (12) denotes a single proposition, a question like (13a) may plausibly be analysed as denoting a set of

[7] "Cleft" here is used in the sense with which it often appears in the literature on Dravidian languages, namely in reference to constructions involving a nominalised constituent put into a copular relationship with a focussed element—such constructions are frequently translatable as English clefts or pseudo-clefts.

[8] Kishimoto (2005) notes that there are *wh*-words in Sinhala which are not associated with an overt Q-particle. Specifically, in addition to *mokə də* "why" we also find *æi* "why", which obligatorily occurs WITHOUT *də*. Kishimoto (2005) argues that *æi* involves the insertion of a null operator which is base-generated in the closest scope position.

[9] As discussed in Chapter 9, Hamblin-type sets have special properties, distinct from those of other set-denoting elements; so that while both *who* and *man* denote sets of individuals (in fact, both denote the set of human beings), the semantic rules governing how *who* composes with other elements are rather different from those governing the composition of *man*. For the sake of presentation, I postpone the formalisation of the composition rules for Hamblin-type elements for later chapters.

propositions (one proposition for each member of the set denoted by *who*). That is, (13a) might denote (in a some possible world and pragmatic context):

(14) $[\![(13a)]\!] = \left\{ \begin{array}{c} \text{John read } \textit{Syntactic Structures}, \\ \text{Mary read } \textit{Syntactic Structures}, \\ \text{Bill read } \textit{Syntactic Structures}, \\ \text{Kim read } \textit{Syntactic Structures} \end{array} \right\}$

In other words, a question like (13a) denotes a set of possible propositions, where the utterer desires to know which member of this set is a true proposition. Another way of putting this is that "knowing a question" is equivalent to knowing what would count as an answer to that question (Hamblin 1958: 162).

Assuming a Hamblin-style analysis of *wh*-words, Q-particles can be analysed as involving a choice function, a notion which has been invoked in the analysis of indefinite and/or interrogative pronouns by numerous linguists in recent years (e.g. von Stechow 1996; Reinhart 1997, 1998; Winter 1997; Kratzer 1998; Hagstrom 1998; Sternefeld 2001; Yatsushiro 2001, 2009; Kratzer & Shimoyama 2002; Cable 2007).[10] Informally we may define a choice function as in (15).

(15) A function f is a choice function iff, for every non-empty set P, f may apply to P yielding some member of P.

For example, given two sets, A = {a, b, c} and B = {x, y, z}, there will exist—amongst other choice functions—three choice functions f_{41}, f_{42}, f_{43} such that:

(16) a. $f_{41}(A) = a$; $f_{41}(B) = y$
b. $f_{42}(A) = b$; $f_{42}(B) = x$
c. $f_{43}(A) = c$; $f_{43}(B) = z$

A choice function is thus exactly the sort of element needed to convert the Hamblin-type sets created by *wh*-words back into ordinary "Montagovian" denotations.

The earliest of these choice-functional analyses of Q-particles is Hagstrom (1998), who analyses Japanese *ka*, Sinhala *də*, and Okinawan *-GA* as denoting choice functions. More precisely, he assumes that these particles undergo obligatory movement, where the trace of this movement denotes a variable over choice functions and the particle itself denotes an operator which existentially binds the choice function variable. Hagstrom's (1998) claim that particles like *ka* bear inherent quantificational force is based on examples which appear to show that *ka* cannot "pick up" quantificational force from their environment (e.g. from an adverb like *usually*); however, Yatsushiro (2009: 151n21) reports that her judgements and those of her informants contradict those of Hagstrom's informants.

Hagstrom (1998) suggests that Japanese and Sinhala Q-particles can be given as similar analysis, namely as particles which adjoin to (the lowest) *wh*-word, subsequently undergoing a "migration operation", in which Q-particles "migrate" to a "launching-site" position outside of any syntactic islands between the *wh*-word and the interrogative C-head, and finally undergo regular movement (overt in the case of Japanese, covert in the case of Sinhala) to adjoin to the C-head. The second step, the "migration" of the particle to the "launching-site" involves an unusual type of movement which leaves no traces but which is sensitive to intervention effects. The "migration" operation is theoretically unattractive in some respects, but Hagstrom (1998) presents certain examples from Japanese which are difficult to account for otherwise.

[10] The notion of choice function itself has been around for much longer, and was originally introduced by Zermelo (1904), for general mathematical set theory, in a paper which gave a proof of the well-ordering theorem for sets.

Yatsushiro (2001, 2009) also analyses Q-particles as involving choice functions, but she argues that particle *ka* itself denotes a variable over choice functions, with the binding of the variable being accomplished by existential closure (Reinhart 1997; Kratzer 1998). Yatsushiro (2001, 2009) concentrates on the analysis of indefinite pronouns formed from *wh*-words and the particle *ka*, as well as the corresponding set of universally quantified pronouns formed from *wh*-words and the particle *mo*.

Both Hagstrom (1998) and Yatsushiro (2001, 2009) assume that indeterminate pronouns denote Hamblin-type sets of individuals (Hamblin 1973). Cable (2007), who like Yatsushiro (2001, 2009) assumes that the quantificational force of particles like Japanese *ka* derives from existential closure, takes a somewhat different approach. Cable (2007), following Beck (2006), notes that Rooth's (1985; 1992; 1996) analysis of focus utilises sets of alternatives which are functionally equivalent to those generated in Hamblin's analysis of questions. Beck (2006) and Cable (2007) therefore propose that the sets of alternatives involved in *wh*-questions are in fact focus semantic values, and that particles like Japanese *ka* make obligatory reference to the focus semantic values of their complements. I show that this theoretical move results in a number of undesirable consequences, and that thus—despite their similarities—we must continue to distinguish between the sets of alternatives generated by *wh*-words and the sets of alternatives which constitute the focus semantic value of an element.

Further, Cable (2007: 72–73n40) in fact claims that the particles which appear in yes/no- and alternative-questions and other disjunctions are morphologically and semantic distinct from the particles which appear in *wh*-questions and *wh*-indefinites. He bases this claim on the fact that in Tlingit—which resembles Japanese and Sinhala in that *wh*-, yes/no-, and alternative-questions and declarative disjunctions all involve the presence of a particle—the particles which appear in yes/no- and alternative-questions (*gé*) and declarative disjunctions (*khach'u*) are morphologically-distinct from the particle which appears in *wh*-interrogatives and *wh*-indefinites (*sá*). I argue that the particle that appears in Tlingit yes/no-questions, *gé*, is a Q-particle in the same sense that Tlingit *sá* (appearing in *wh*-interrogatives and *wh*-indefinites) is a Q-particle; and that likewise Sinhala *də* denotes a variable over choice-functions whether it accompanies a *wh*-word or appears in a yes/no-question. The basis for this argument is two-fold. Firstly, there is crosslinguistic evidence which suggests that the particles which appear in *wh*-interrogatives and *wh*-indefinites are identical to the particles which appear in yes/no-questions and other disjunctive contexts. Not only does Sinhala *də* appear both in *wh*- and yes/no-questions, but so too does Japanese *ka* and Malayalam *-oo*. It thus seems unlikely that, as Cable (2007: 74–75n40) suggests, "the use of *da*/*ka* in Sinhala/Japanese polar questions reflects the existence of a separate, homophonous 'yes/no' particle", given that this purported homophony appears fairly common crosslinguistically. It is of course possible that such homophony could reflect a common diachronic development rather than a synchronic identity, but the latter would appear to be a better null hypothesis. Secondly, given that yes/no- and alternative-questions (and other disjunctive constructions) involve a choice between alternatives, it seems eminently reasonable to expect that these too involve choice functions.

Cable (2007) also differs from Hagstrom (1998) in that he distinguishes what he refers to as "Q-adjunction" languages from "QP-projection" languages; the former involve the Q-particle directly adjoining to its sister, while the later involve the Q projecting its own XP-level category and taking its sister as its complement. Cable suggests that languages like Japanese and Korean are Q-adjunction languages, while languages like Sinhala, Tlingit, English, German are QP-projection languages. One major difference between these two types of languages is what happens when "Q" is targeted for movement. In Q-adjunction languages only the Q-particle itself undergoes movement, while in QP-projection languages the entire QP moves (which Cable suggests is a good explanation for *wh*-pied piping effects in languages like English, Tlingit etc.). Positing such a distinction between languages appears to be a productive hypothesis, given that it accounts for differences between languages like Japanese and languages like Sinhala.

Though some accounts mention the fact that Sinhala *də*, Japanese *ka* etc. also appear in disjunctive structures (e.g. Hagstrom 1998, Cable 2007), almost no-one has attempted to explain why such particles should appear in disjunctions. One notable exception is Jayaseelan (2001, 2008), who remarks on the fact that in several languages (Japanese, Sinhala, Malayalam) there is at least a phonological identity between disjunction markers and question particles, arguing that what appears to be a question particle is in fact a realisation of the disjunction operator, appealling to Baker's (1970) identification of question particles as question operators and the notion that the semantics of the question operator involves disjunction. Jayaseelan's analysis shares in common with the approach advocated here the idea that interrogatives and disjunction both involve operations over sets of possibilities. However, he does not offer a full formal account of what the semantics underlying disjunction operators would be. Further, his analysis requires the assumption that disjunction markers and disjunction operators are two distinct categories which happen to be homophonous in many languages.

1.2.3 Overview of previous work

Previous studies of Sinhala interrogatives and focus and Q-particles crosslinguistically have established a number of important empirical facts and advanced important theoretical concepts, upon both of which the present study builds. However, both in terms of empirical and theoretical coverage, several gaps remain which this dissertation seeks to address.

The importance of the studies of Jim Gair and his students in setting the groundwork for linguistic research on Sinhala in general, and on interrogative and focus constructions more specifically, cannot be overestimated. However, these studies, as well as those of Kishimoto (1992, 2005) have concentrated primarily on the syntactic component; as I argue here, a complete understanding of the properties of these constructions necessitates a complementary study of semantics alongside syntax.

Hagstrom (1998) importantly introduces the idea of treating Q-particles in terms of choice functions, and the work done by Cable (2007), following in a similar vein, introduces a number of interesting theoretical proposals connected with a choice functional treatment of Q-particles, and importantly adds empirical coverage with his close study of Tlingit. However, neither of these studies provides a complete analysis of Q-particles in all of the environments in which they appear. Detailed analysis of disjunctive constructions is notably lacking from both studies, as is examination of the pragmatics of indefinites formed with Q-particles—which show a strong tendency to form epistemic indefinites.

Thus this dissertation seeks to extend the study of Q-particles, particularly in Sinhala, both empirically and theoretically, and from both diachronic/typological and synchronic perspectives.

1.3 Summary/Roadmap

In summary, the primary goal of this dissertation is to provide a unified account of the syntactic and semantic properties of Q-particles in the full set of environments in which they appear crosslinguistically, concentrating primarily on the properties of Q-particles in Sinhala from the 8th-century to the present day, but also including detailed analysis of the properties of Q-particles in other (unrelated) languages, specifically Malayalam, Japanese, and Tlingit. This account builds upon previous analyses of Sinhala interrogatives and focus specifically,[11] and on previous formal accounts of Q-particle more generally,[12] but seeks to provide a fuller and more comprehensive analysis of the formal properties of Q-particles than has hitherto been presented.

[11] e.g. Gair (1983[1998], 1986[1998]b), Gair & Sumangala (1991), Kariyakarawana (1998), Kishimoto (2005).
[12] e.g. Hagstrom (1998), Yatsushiro (2001, 2009), Jayaseelan (2001, 2008), Cable (2007).

My analysis of Q-particles incorporates certain aspects of the analyses of Hagstrom (1998), Yatsushiro (2001, 2009), and Cable (2007) discussed above. However, all of these previous accounts focus on the analysis of *wh*-interrogatives and/or *wh*-indefinites, and provide no complete formal semantic account of the role of such particles in disjunctive contexts (including alternative and yes/no questions), and thus do not present an account which covers the full range of Q-particle environments. As stated previously, a unified analysis of Q-particles is obviously highly desirable.

In brief, I provide a semantically unified analysis of Q-particles as denoting variables over choice functions. This choice-functional analysis can account for the fact that, crosslinguistically, Q-particles often appear in a set of seemingly heterogeneous environments (interrogatives, disjunctions, indefinites). Given that all of these environments can be analysed as including elements denoting Hamblin-type sets and that the role of a choice function is to apply to a set and return a single element from that set, the appearance of Q-particles in just this set of environments can be explained naturally if they are treated as denoting variables over choice functions. Crosslinguistic and diachronic differences in the distribution of Q-particles are accounted for (primarily) by positing that Q-particles may bear different (language-particular) sets of formal syntactic features.

For example: Sinhala, like Tlingit, possesses multiple Q-particles—though the distribution of these particles is rather different than that of Tlingit. Specifically, alongside of *də*, which appears in *wh*-, yes/no-, and alternative-questions, as well as in the formation of certain *wh*-indefinites, we also find the particles *hari* and *vat*. Neither *hari* nor *vat* appear in interrogative contexts (which is perhaps why they have not been much discussed in the literature on Q-particles), but they are involved in the formation of non-interrogative disjunctions (*vat* in NPI contexts, *hari* elsewhere) and, like *də*, in the formation of *wh*-indefinites (again, *vat* in NPI contexts, *hari* elsewhere).

The apparent multiplicity of Q-particles in certain languages (e.g. Tlingit *sá* vs. *gé* vs. *khach'u*, Sinhala *də* vs. *hari* vs. *vat*) and not in others (e.g. Malayalam *-oo*) suggests that there may be language-specific rules which govern the surface form taken by a Q-particle in different syntactic contexts. That is, in some languages there may exist different Q-particles with identical semantic denotations (i.e. as variables over choice functions), but bearing different syntactic features (of the sort described in Chomsky 2000, 2001, 2008 and Pesetsky & Torrego 2007).

The multiplicity of Q-particles is interesting in other ways as well. Though I maintain that all of these particles bear the same semantic denotation (though possibly differing in which formal syntactic features they bear), there is evidence which suggests that Q-particles may differ in terms of their pragmatic as well as syntactic features. This line of investigation intersects with another aspect of this study which investigates the fact that indefinites pronouns formed from a *wh*-word ("indeterminate pronoun") and a Q-particle, like Japanese *nani-ka* "something", Sinhala *mokak də* "something" etc., have a crosslinguistic tendency to pragmatically signal that the speaker lacks certain information about who or what satisfies the existential claim (see Bhat 1981 on Kannada, Jayaseelan 2001 on Malayalam, Moore 2003 on Japanese, and more generally Haspelmath 1997: 45–51). This fact is in itself interesting and deserving of a formal account, but Sinhala presents an additional intriguing situation: the combination of *wh*-words with different Q-particles results in indefinites which signal different degrees of ignorance. Specifically, a *wh+də* indefinite signals a higher degree of ignorance about identity of the indefinite than does a *wh+hari* indefinite.

I assume that the epistemic properties of such indefinites are pragmatic in nature (cp. von Fintel's 2000b treatment of *wh*-based *-ever* free relatives in English, and, with more immediate relevance, the treatments by Alonso-Ovalle & Menéndez-Benito 2003, 2010 of the Spanish epistemic indefinite determiner *algún*), and, further, that their formal analysis can be naturally cast in terms of choice functions. My analysis draws on the work of Boër & Lycan (1975), who present formalisation of what it is to "know" who or what someone or something is, developing the idea of "Important Predicates". Important Predicates are singular predicates (i.e. predicates which denote singleton sets of individuals) for a particular individual with respect to a particular purpose or project (or, we might add, from a particular perspective). Like choice functions, Important Predicates involve the selection of a unique individual and

so a correspondence relationship can be established between the two concepts, allowing us to define the notion of choice functions which approximate Important Predicates. This, combined with the idea of intensional choice functions (Romero 1999)—choice functions which operate over individual concepts rather than individuals—allows for a choice-functional explanation of the "ignorance component" associated with epistemic indefinites.

In addition, the analysis of Q-particles as consistently denoting variables over choice functions requires a novel syntactic and semantic analysis of disjunction, where elements which on the surface appear to play the role of the disjunction operator are instead analysed as variables over choice functions. Such an analysis still requires the presence of some element which can apply to the disjuncts and form a Hamblin-type set (cp. Alonso-Ovalle 2006) to which a choice function can apply; however, at least in the languages examined here, this set-forming element is phonetically null (which dovetails with the syntactic analysis of disjunction proposed, for independent reasons, by den Dikken 2006). The analysis advanced herein has implications for the treatment of disjunction and conjunction more generally, suggesting that in the formation of disjunction, and perhaps in the formation of conjunction as well, choice-functions are ubiquitous across languages.

Given the interrelation between Q-particles and focus in Sinhala, I also examine the formal properties of the Sinhala focus construction and their relation to the formal properties of interrogatives.

Finally, I trace the historical development of (1) the focus construction of modern Sinhala, and (2) the Sinhala Q-particles *də* and *hari*. The diachronic component of the dissertation—while important in its own right since it adds to our knowledge of the history of Sinhala—also crucially informs the synchronic analysis by showing which formal properties are subject to change and, perhaps more importantly, in what ways these properties undergo change.

I choose to focus this study on the following set of languages: Sinhala, Japanese, Malayalam, and Tlingit, in part because this set of languages provides a good sampling of different grammatical systems which involve (overt) Q-particles. Malayalam has only one Q-particle, *-oo*, which appears in all of the possible Q-particle contexts, with the exception that *-oo* does not appear in the formation of *wh*-questions in the modern language (though it did in earlier Malayalam). Sinhala, Japanese, and Tlingit constitute important examples of languages which employ multiple Q-particles, and also exhibit rather different patterns of distribution of these particles. In modern colloquial Sinhala, the Q-particle *də* appears in a wide range of environments: in *wh*-, yes/no-, and alternative-questions, as well as in the formation of one type of indefinite; while *hari* appears in disjunctions in affirmative declaratives, and in the formation of another type of indefinite; and *vat* acts as the NPI counterpart of *hari*. In Tlingit, on the other hand, the distribution of the various Q-particles is rather different. Here, *sá* appears in *wh*-interrogatives and in the formation of indefinites (in other words, in environments including *wh*-pronouns), while *gé* appears in yes/no and alternative questions, and *khach'u* in disjunctions in non-interrogative contexts. These differences in Q-particle distributions thus form a solid basis for theorising about the nature of the range of possible language-specific differences in the properties of Q-particles—which I suggest are largely morphosyntactic in nature. Sinhala forms the central focus of the dissertation, in part because of its rich literary history, which allows for diachronic study of Q-particles.

The remainder of the dissertation is structured as follows. Chapter 2 provides an overview of Q-particles crosslinguistically, presenting relevant data from Sinhala, Malayalam, Japanese, and Tlingit. Chapters 3–8 are concerned primarily with the analysis of Q-particle and focus constructions in Sinhala. Chapter 3 lays out the basic syntactic formalism adopted and applies this formalism to the full range of Q-particle constructions found in Sinhala. Chapter 4 examines the syntax of the modern colloquial Sinhala focus construction and Chapter 5 provides a formal semantic account of Sinhala focus constructions. Chapter 6 presents a formal semantic analysis of Q-particles and *wh*-words in the full range of interrogative constructions in which they are found in modern colloquial Sinhala (the syntax and semantics of disjunction are treated more fully in Chapter 10). Chapter 7 examines indefinites, in particular epistemic indefinites, especially in modern colloquial Sinhala, and their connection with Q-particles,

namely the crosslinguistic tendency for indefinites formed with Q-particles to be epistemic indefinites, and provides a formal pragmatic analysis of the same. Chapter 8 provides a brief excursus on the appearance of Q-particles in relative clauses in Dravidian and early Sinhala. Chapter 9 proposes that while all Q-particles have the same semantic denotation, they may differ in (a) the set of formal syntactic features they bear, (b) the lexico-semantics of *wh*-words, and (c) in the presuppositions they bear. Here I show that a complete understanding of the distributional properties of Q-particles requires semantic, syntactic, and pragmatic analysis, drawing on the analyses proposed in the previous chapters. Chapter 10 presents a more complete discussion of the role of Q-particle in disjunctions, and develops a novel analysis of the syntax and semantics of (dis)junction more generally. Finally, Chapter 11 provide examinations of the diachronic development of the Sinhala focus construction and the Sinhala Q-particles *də*, *hari*, and discusses the relevance of these data for the formal synchronic analyses. Chapter 12 provides a summary and discussion of the greater implications of the synchronic and diachronic analyses proposed here for generative theory, and suggestions about possible future areas of investigation.

Chapter 2

Q[uestion]-particles crosslinguistically

In this chapter I provide a brief overview of the environments in which Q-particles appear in (modern colloquial) Sinhala, Malayalam, Japanese and Tlingit. As indicated in Chapter 1, Q-particle environments are not limited to interrogatives, but include also disjunctions (even in non-interrogative contexts) and the formation of certain *wh*-based indefinite pronouns. In some languages, such as Malayalam, a single type of Q-particles appears in all contexts, while other languages, such as Sinhala and Tlingit, use Q-particles with different morphological forms in different syntactic environments. Here I provide an overview for each language, detailing which Q-particles appear in which syntactic environments.

2.1 (Modern Colloquial) Sinhala

Sinhala is an Indo-Aryan language spoken in the island nation of Sri Lanka (the former Ceylon). It is the southernmost Indo-Aryan language (together with Dhivehi,[1] a closely-related language spoken in the Maldives), and has been isolated from the Indo-Aryan languages of the north Indian mainland for over two millennia. It has, however, been in contact with southern Dravidian languages (forms of Tamil and the ancestor of Tamil and Malayalam) and exhibits some degree of convergence with Dravidian in terms of its phonology, syntax, morphology, and lexicon—but remains recognisably Indo-Aryan.[2] See further Gair 1982[1998] for a general description of Sinhala and the impact of Dravidian.

One of the salient features of modern colloquial Sinhala is its use of two morphologically distinct verbal endings: one, *-a*, is used in neutral contexts, the other, *-e*, is found in "focussing" contexts, which in Sinhala includes most *wh*-interrogatives and some yes/no-interrogatives. Sinhala, like other South Asian languages, displays default SOV word order.

Another relevant feature of modern colloquial Sinhala is its use of "Question particles". Interrogatives of all types (*wh*-, yes/no-, and alternative-questions) employ the particle *də*, which also appears in the formation of certain *wh*-indefinite pronouns. Additionally, we find two other particles, which I will suggest are also Q-particles: *hari*, used in the formation of non-interrogative disjunctions and, like *də*, in the formation of certain *wh*-indefinite pronouns; and *vat*, which is an NPI counterpart of *hari*.

Wh-questions in Sinhala employ the Q-particle *də*, and the verb takes the special "focussing" *-e* ending (following Kishimoto 2005, I refer to this as the *-e* ending, glossed as -E), distinguished from the neutral ending (the *-a* ending, glossed as -A). Compare the declarative in (1) with the corresponding interrogative in (2).[3]

[1] On which see Cain (2000).

[2] See Gair (1976[1998]: 200–201) who notes ". . . the survival of Sinhala as a clearly Indo-Aryan language might be looked on as a minor miracle of linguistic and cultural history"; see also Karunatillaka (1977).

[3] Examples in this section are taken from Kishimoto (2005) unless otherwise noted.

(1) Chitra potə gatta
Chitra book bought-A
'Chitra bought the book.'

(2) Chitra monəwa də gatte?
Chitra what *də* bought-E
'What did Chitra buy?'

The *-e*-ending also appears in focus constructions (as discussed in Chapters 4 and 11.1)—providing evidence that Sinhala *wh*-interrogatives involve focus.

Wh-words along with their associated Q-particle (and any intervening material) may also be dislocated to the right of the verb of the clause over which they take scope, as in example (3). This movement is characteristic of focussed elements, which also optionally undergo an identical operation (see below Chapter 4).

(3) Chitra gatte monəwa də?
Chitra bought-E what *də*
'What did Chitra buy?'

The Q-element is obligatory in *wh*-questions, regardless of the form of the verb, as shown in (4).

(4) *Chitra monəwa gatta/gatte?
Chitra what bought-A/bought-E
'What did Chitra buy?'

The *-e* marking of the verb is obligatory in *wh*-questions, see (5) and (6), respectively.[4]

(5) *Chitra monəwa də gatta?
Chitra what *də* bought-A
'What did Chitra buy?'

(6) *Chitra monəwa gatta/gatte də?
Chitra what bought-A/bought-E Q
'What did Chitra buy?'

Without the *-e* marking, a *wh*-word accompanied by a Q-element is interpreted as an indefinite, as can be seen by the contrast in (7a) and (7b).[5]

(7) a. mokak də wætuna.
what *də* fell-A
'Something (unidentified) fell.' (Gair & Sumangala 1991)

b. mokak də wætune?
what *də* fell-E
'What fell?' (Hagstrom 1998)

The verb in *-e* marks the scope of the *wh*-word. This can be seen in examples (8a), (8b) below.[6]

[4]Unless *də* appears clause-finally, which it cannot generally do when the *wh*-word is in the matrix clause, see further at (14) ff.

[5]See Ramchand (1997), who discusses a similar situation in Bengali.

[6]The form *kiyəla*, glossed as 'that', is morphologically a gerund-form (of the type Hindi *kar ke* 'having done' or Nepali *gar-era* 'having done') of a verb meaning 'to speak', and thus in origin appears to be a quotative (like Nepali *bhanī, bhanera, bhanne (kura)*). However, it perhaps is better analysed as having become a complementiser by this stage.

(8) a. Ranjit [[[kau də] aawa] kiyəla] danne?
Ranjit [[[who *də*] come.PAST.A] that] know.PRES.E
'Who does Ranjit know came?'

b. Ranjit [[[kau də] aawe] kiyəla] dannəwa.
Ranjit [[[who *də*] come.PAST.E] that] know.PRES.A
'Ranjit knows who came.'

In (8a) the matrix verb bears the *-e* marking, giving a matrix *wh*-question reading, whereas in (8b) the embedded verb bears the *-e* marking, thus the *wh*-word takes scope only over the lower clause.

As early as Gair 1983[1998] it was noted that though *wh*-words appear to be insensitive to islands, in the sense that *wh*-words can scope out of islands, there can be no island barrier (e.g. complex NP) between the Q-element and the verb in *-e*, cp. (9) with (10).

(9) *Chitra [[[[Ranjit [monəwa] də gatta] kiənə] kaṭəkataawə] æhuwe ?
Chitra [[[[Ranjit [what] *də* bought-A] that] rumour] heard-E
'What$_i$ did Chitra hear the rumour that Ranjit bought t$_i$?'

(10) Chitra [[[Ranjit monəwa gatta] kiənə] kaṭəkataawə] də æhuwe ?
Chitra [[[Ranjit what bought-A] that] rumour] *də* heard-E
'What$_i$ did Chitra hear the rumour that Ranjit bought t$_i$?'

That the complex NP in (9) is indeed an island is demonstrated by the fact that overt extraction out of Complex Noun Phrases [CNPs] is also ungrammatical, no matter whether the extracted element is a *wh*-word (11), (12) or not (13)—regardless of the placement of *də*.

(11) *Chitra [[[[Ranjit t$_i$] də gatta] kiənə] kaṭəkataawə] æhuwe monəwa$_i$?
Chitra [[[[Ranjit t$_i$] *də* bought-A] that] rumour] heard-E what$_i$?
'What$_i$ did Chitra hear the rumour that Ranjit bought t$_i$?'

(12) *Chitra [[[Ranjit t$_i$ gatta] kiənə] kaṭəkataawə] æhuwe monəwa$_i$ də ?
Chitra [[[Ranjit t$_i$ bought-A] that] rumour] heard-E what$_i$ *də* ?
'What$_i$ did Chitra hear the rumour that Ranjit bought t$_i$?'

(13) *Chitra [[[Ranjit t$_i$ gatta] kiənə] kaṭəkataawə] æhuwe ee potə$_i$.
Chitra [[[Ranjit t$_i$ bought-A] that] rumour] heard-E that book
'It was that book$_i$ which Chitra heard the rumour that Ranjit bought t$_i$?'

The Q-element associated with a *wh*-phrase may appear in clause-final position in a restricted set of contexts: (i) when the *wh*-phrase is embedded under a verb like *dannəwa* 'know' as in example (14), (ii) when the *wh*-phrase is in the matrix clause, only when the *wh*-phrase is *kīdenek* 'how many (animate)' or *kī-ak* 'how many (inanimate)' as in (15).[7]

(14) Ranjit [[[kauru aawa] də] kiyəla] dannəwa.
Ranjit [[[who came-A] *də*] that] know-A
'Ranjit knows who came.'

(15) kiidenek potə kieuwa də?
how.many book read-A *də*
'How many (people) read the book?'

[7]E-forms of the verb appear in yes/no questions only when the verb in question takes scope over *də*.

Note that when the Q-element appears clause-finally, the verb takes the neutral *-a* marking rather than the *-e* marking.

Whenever the Q-element can appear in clause-final position, it is also possible for it to occur in the non-final position (adjacent to the *wh*-word or a constituent containing the *wh*-word), as shown in (16) and (17).

(16) Ranjit [[kau də aawa] kiyəla] dannəwa.
Ranjit [[who *də* came-A] that] know-A
'Ranjit knows who came.' (Gair 1992)

(17) kiidenek də potə kieuwe?
how.many *də* book read-E
'How many (people) read the book?'

However, these two variants carry different presuppositions. The clause-final positioning of the Q-element carries no presupposition that there must be at least one value which satisfies the proposition, whereas the non-clause-final positioning of Q does (see Kishimoto 2005: 9–11). That is, when the Q-particle appears clause-finally in a *wh*-interrogative like "How many people read the book?", the speaker does not assume that there is a true answer to the question, allowing the possible response "No-one read it". When Q-particles in *wh*-interrogatives occur in their usual position (adjacent to the *wh*-word, or, in case of an island, on the edge of the island containing the *wh*-word), the speaker does suppose that there is some true answer, thus making a response like "No-one read the book" infelicitous. Consider the following question-answer pairs in (18)–(21).[8]

(18) a. Q: kiidenek potə kieuwa də?
how.many book read-A Q
'How many (people) read the book?'

b. A: kauru-wat kieuwe nææ.
anyone read-E not
'No-one read it.'

(19) a. Q: kiidenek də potə kieuwe?
how.many *də* book read-E
'How many (people) read the book?'

b. A: # kauru-wat kieuwe nææ.
anyone read-E not
'No-one read it.'

[8]In the answers [A] in (18)–(21) below, it should be noted that the E-form of the verb of the matrix clause appears because of the negation *nææ*; this negation always triggers E-forms of the verb, for reasons which are at this point unclear to me.

(20) a. Q: oyaa [[kauru aawa də] kiyəla] dannəwa də?
you [[who came-A *də*] that] know-A *də*
'Do you know who came?'

b. A: oo. kauru-wat aawe nææ.
yes anyone came-E not
'Yes. No-one came.'

(21) a. Q: oyaa [[kau də aawe] kiyəla] dannəwa də?
you [[who *də* came-E] that] know-A *də*
'Do you know who came?'

b. A: # oo. kauru-wat aawe nææ.
yes anyone came-E not
'Yes. No-one came.'

Additionally, not all verbs allow for clause-final *də* in their complement clauses. Verbs which do so include *dannəwa* 'know', *hoya bərənəwa* 'examine', *parikśā kərənəwa* 'look into, inspect', and *tērenəwa* 'understand'; verbs not allowing clausal final *də* include *ahanəwa* 'ask', *prasnə kərənəwa* 'question', and *hitenəwa* 'consider' (see Kishimoto 2005: 8 and Gair 1983[1998]). The parameters determining these classes of verbs are not clear. The classes cannot be defined on the basis of the ability of a verb to take an interrogative subordinate clause (*ahanəwa* 'ask' does not allow clause-final *də*) or by their ability to select a declarative complement in addition to an interrogative complement (*hoya bərənəwa* 'examine' allows for clause-final *də*, but cannot take a declarative complement).

In contrast to the use of *də* in *wh*-interrogatives, *də* in yes/no questions can freely occur in clause-final position even in matrix clauses. Indeed, clause-final placement of *də* is the unmarked position for the yes/no particle (22).

(22) Chitra ee potə kieuwa **də**?
Chitra that book read-A ***də***
'Did Chitra read that book?' (Kishimoto 2005: 11)

The particle *də* may also appear after a constituent smaller than IP—in which case it marks that constituent as focussed (23), and, as expected, the verb appears in the *-e* form.

(23) Chitra ee potə **də** kieuwe?
Chitra that book ***də*** read-E
'Was it that book which Chitra read?' (*Ibid.*)

Də is also used to form interrogative disjunctions, as shown in example (24), but is not possible in non-interrogative disjunctive contexts, as shown by the ungrammaticality of (25).

(24) Gunəpālə **də** Chitra **də** gamətə giyē?
Gunapala ***də*** Chitra ***də*** village.DAT go.PAST.E
'Was it Gunapala or Chitra who went to the village?'

(25) *Gunəpālə **də** Chitra **də** gamətə giyā.
Gunapala ***də*** Chitra ***də*** village.DAT go.PAST.A
'Gunapala or Chitra went to the village.'

In addition to *də* we also find two other particles which are used in the formation of both indefinites and disjunctions: *hari* and *vat*. The particle *də* is used to form interrogative disjunctions, but cannot be used in declarative disjunctions as shown in (25), where we instead find the particle *hari* in affirmative contexts, and *vat* in negative contexts.

(26) Gunəpālə **hari** Chitra **hari** gamət̪ə giyā.
Gunapala ***hari*** Chitra ***hari*** village.DAT go.PAST.A
'Gunapala or Chitra went to the village.'

(27) Gunəpālə **vat** Chitra **vat** gamət̪ə giyē nǣ
Gunapala ***vat*** Chitra ***vat*** village.DAT go.PAST.E NEG
'Neither Gunapala nor Chitra went to the village.'

In addition, both *hari* and *vat* can be used, like *də*, to form indefinite pronouns. Again, *vat* appears in negative contexts, (28) and *hari* in affirmative contexts, (29).

(28) Kauru **hari** gamət̪ə giyā
who ***hari*** village.DAT go.PAST.A
'Someone went to the village.'

(29) Kauru **vat** gamət̪ə giyē nǣ
who ***hari*** village.DAT go.PAST.E NEG
'No-one went to the village.'

Both *hari* and *də* indefinites pragmatically signal that the speaker lacks knowledge concerning the identity of the individual in question—in contrast to 'plain' indefinites formed from a noun+the indefinite suffix *-ek*, as in example (30).

(30) Ken-ek gamət̪ə giyā
human-INDEF village.DAT go.PAST.A
'Someone went to the village.'

Example (30), unlike the *hari* and *də* indefinites in examples (28) and (7a), carries no additional pragmatic signals and thus can be uttered even when the speaker knows who it is who went to the village.

The difference between *hari* and *də* indefinites is subtle, but the general distinction is that *də* indefinites signal a greater degree of ignorance than do *hari* indefinites, see further Chapter 7.

In the following sections I examine the distribution of Q-particles in Malayalam (including both Old Malayalam and Modern Malayalam), Japanese, and Tlingit.

2.2 Malayalam

Malayalam is a Dravidian language spoken in the state of Kerala in southern India. It is closely related to Tamil, spoken in nearby state of Tamil Nadu as well as in Sri Lanka—in fact the language referred to as Old Tamil (ca. 1st – 8th centuries A.D.) is the common ancestor of both Tamil and Malayalam.[9] Malayalam bears a special relation to Sinhala in that Sinhala has long been in contact with the closely related Dravidian language Tamil.

Like many other Dravidian languages, Malayalam employs particles in a wide variety of syntactic contexts, including interrogatives, indefinites, relative clauses (optionally), and disjunctions. Unlike many of the other languages examined in this study, we find only one Q-particle in Malayalam, *-oo*. Like Sinhala and other South Asian languages, default word order in Malayalam is SOV.

The particle *-oo* appears in yes/no and alternative questions, as in examples (31) and (32).

[9]What I refer to as Old Malayalam dates from roughly the 15th – 19th centuries A.D.

(31) John wannu-**(w)oo**?
John came-***oo***
'Did John come?' (Jayaseelan 2001: 67)

(32) John wannu-**(w)oo**, illa-**(y)oo**?
John came-***oo***, not-***oo***
'Did John come, or not?' (Jayaseelan 2001: 67)

This particle also appears in declarative disjunctions, as in example (33).

(33) Mary John-ine-**(y)oo** Bill-ine-**(y)oo** cumbiccu
Mary John-ACC-***oo*** Bill-ACC-***oo*** kissed
'Mary kissed John or Bill.' (Jayaseelan 2008: 3)

Here we may note an important difference between Sinhala and Malayalam: while Sinhala *də* is restricted to interrogative disjunctions (with *hari* or *vat* appearing in non-interrogative contexts), Malayalam *-oo* appears in both interrogative (32) and declarative disjunctions (33).

Like Sinhala *də*, Malayalam *-oo* can also be used to form indefinites, as in (34).

(34) ñāṉ iruṭṭ-il ār-e-**(y)oo** toṭṭu
I darkness-in who-ACC-***oo*** touched
'I touched somebody in the dark.' (Jayaseelan 2001: 66)

And, again like Sinhala indefinites formed with *də* or *hari*, the indefinite formed from *wh*-word+*-oo* carries a pragmatic signal that the speaker lacks further knowledge of the individual in question.[10]

In contrast to Sinhala, *wh*-questions in modern Malayalam do not employ *-oo*, as shown by examples (35), (36).

(35) ārə wannu?
who came
'Who came?' (Jayaseelan 2001: 67)

(36) [awaṉ ewiḍe pooyi ennə] ñāṉ coodiccu
[he where went C] I asked
'I asked where he went.' (Jayaseelan 2001: 67)

However the particle is present in archaic, (37), and old Malayalam, (38), (39), *wh*-questions.

(37) it-entu katha-**(y)oo**?
this-what story-***oo***
'What story is this?' (Raman Pilla 1918: 151, cited in Jayaseelan 2001: 68)

(38) entu-kil-**oo** rājya-ttiṉṉu want-a upadrawam?
what-be-***oo*** kingdom-DAT came-RELATIVISER trouble
'What is the trouble that has come to the kingdom?' ("Ambarrīshoopākhyānam", Narayanapilla 1971: 21)

[10] Again, a 'plain', pragmatically-unmarked indefinite, formed from a NP preceded by "one", is available, as in example (i).

(i) ñāṉ innale **oru āḷ-e** paricayappeṭṭu
I yesterday **one person**-ACC met
'I met someone yesterday' (Lit. 'I met a person yesterday.') (Jayaseelan 2001: 66)

(39) maharṣi nintiruwaḍi entu-nimittam-ākil-**oo** iwiḍam nookki ezhunnaḷḷi?
great-sage (hon. title) what-reason-be-***oo*** this-place seeing came.HON
'For what reason is it that the great sage has been pleased to come to this place?' (*ibid.*, p. 32, cited in Jayaseelan 2001: 68)

This difference between the old and modern languages seems to reflect differences in the semantics of the *wh*-word itself (discussed further in Chapter 9.2).

2.3 Japanese

Japanese, like Sinhala and Malayalam, uses SOV word order. Unlike Malayalam, Japanese employs a number of different Q-particles, though the particle *ka* has the widest distribution and is the only Q-particle which appears outside of interrogative contexts.

Interestingly, early Japanese employed a construction referred to as *kakari-musubi* (see Sansom 1928, Ogawa 1976, 1977, Whitman 1997, Hagstrom 1998: 24–28, Watanabe 2002, Yanagida 2006, amongst others), which is reminiscent of constructions employing *-e* verbal forms in Sinhala—in that it involves a clause-internal (rather than clause-final) particle which induces a special marking on the verb. Particles participating in this construction include not only the Q-particle *ka*, but other particles including *koso*, *zo*, and *namu*, which seem to be emphatic particles. An example of an Old Japanese *wh*-construction is given below in (40).

(40) sisi husu-to **tare ka** kono koto oomae-ni maos**u**
beasts lie-QUOT **who *ka*** this thing Emperor.DAT say.**M**[11]
"Who reported to the Emperor that beasts were lying?" (*Nihon Shoki* [720]:75, Ogawa 1977: 221, from Hagstrom 1998: 25)

In Old Japanese, unlike modern Japanese (see below), but similar to the situation of Sinhala *də* in *wh*-questions, the Q-particle *ka* occurs suffixed to the *wh*-word *tare* "who", and the verb takes a special form (recalling the special *-e* verbal form of Sinhala). See further, Hagstrom (1998: 24–28, 37–40). In this study, however, I will be primarily concerned with the distribution of Q-particles in the modern language.

In modern Japanese the particle *ka* may appears in *wh-*, alternative, and yes/no-questions, as well as in the formation of declarative disjunctions and indefinites. In addition, Japanese has a number of Q-particles which are used in a subset of the contexts in which *ka* appears, these are: *no*, *ndai*, *kai*, and *kadooka*. The choice of Q-particle in part is a matter of politeness: *ka* is the most polite form, while *no* is less polite than *ka* but not as informal as *kai* or *ndai* (see further Miyagawa 1987, 1998; Ginsburg 2009).

Similar to the pattern observed for Sinhala and Malayalam, in Japanese, *ka* appears clause-finally in yes/no questions, see example (41).

(41) gakkoo-ni ik-imas-u ka?
school-to go-POL-PRES ***ka***
'(Are you) going to school?' (Yoshida & Yoshida 1996)

The particles *no* or *kai*[12] may also occur here rather than *ka*, as well there is the possibility of zero-marking:[13]

[11] The gloss "M" (for *musubi*) indicates the special adnominal form that the verb takes in *kakari-musubi* constructions.
[12] *Kadooka* may also be employed, but only in embedded clauses.
[13] Yes/no and *wh*-questions formed without any Q-particle are possibilities found in informal speech (see Yoshida & Yoshida (1996)).

(42) gakkoo-ni ik-imas-u (no/kai/ø)?
school-to go-POL-PRES ***no*/*kai*/*ø***

In contrast to Sinhala and old/archaic Malayalam, in *wh*-questions *ka* appears clause-finally, rather than following the *wh*-word, as shown in (43), (44).

(43) John-ga nani-o kaimasita **ka**?
John-NOM what-ACC bought.POL ***ka***
'What did John buy?' (Hagstrom 1998: 15)

(44) John-ga [Mary-ga nani-o katta **ka**] sitteiru
John-NOM [Mary-NOM what-ACC bought ***ka***] knows
'John knows what Mary bought.' (Hagstrom 1998: 16)

In *wh*-questions like (43), *ka* may be replaced by *no*, *ndai*, or no marking:

(45) John-ga nani-o kaimasita **no/ndai/ø**?
John-NOM what-ACC bought.POL ***no*/*ndai*/*ø***

Like Sinhala *də*, *hari* and Malayalam *-oo*, Japanese *ka* can also be used to form indefinites, as in example (46). In contrast to *wh*-questions, in Japanese *wh*-based indefinites the particle *ka* follows the *wh*-word rather than appearing clause-finally. Again, like Sinhala and Malayalam *wh*-indefinites, Japanese *wh*-based indefinites like (46) signal the speaker's lack of knowledge of the identity of the referent (Moore 2003).[14]

(46) dare-**ka**-ga hon-o katta.
who-**ka**-NOM book-ACC bought.
'Someone bought books.' (Kuroda 1965: 97)

It appears that Japanese, like Sinhala (see Section 2.1), may be able to encode two levels of speaker ignorance, but using a rather different mechanism. Yatsushiro (2001: 12) provides examples of *ka* indefinites in which *ka* attaches to a postposition or to a higher noun phrase containing a *wh*-word which seem to signal a greater degree of speaker ignorance than do corresponding examples in which *ka* attaches directly to the *wh*-word: compare (47a) with (47b) and (48a) with (48b).[15]

[14] As in Sinhala and Malayalam, there are 'plain' NP-based indefinites in Japanese which carry no pragmatic signal about the speaker's knowledge.

(i) Watashi-ni mise-tai **mono-ga** aru n daroo.
I-DAT show-DES **thing-NOM** exist NMLZ probably
'I take it you've got something for me to look at.' (Moore 2003: 605)

[15] See also Nishigauchi (1990: 121–123) and Shimoyama (2006).

(47) a. Dare-**ka**-kara tegami-ga todoita.
who-***ka***-from letter-NOM arrived.
'A letter arrived from someone.' (*ibid.*)

b. Dare-kara-**ka** tegami-ga todoita.
who-from-***ka*** letter-NOM arrived.
'A letter arrived from someone (or other).' (*ibid.*)

(48) a. [Dare-**ka**-no hahaoya]-ga paatii-ni kita.
[who-***ka***-GEN mother]-NOM party-LOC came
'Someone's mother came to the party.' (*ibid.*)

b. [Dare-no-**ka** hahaoya]-ga paatii-ni kita.
[who-GEN-***ka*** mother]-NOM party-LOC came
'Someone (or other)'s mother came to the party.' (*ibid.*)

Though Yatsushiro (2001) does not discuss any possible difference in the epistemic signals of these indefinites, her translations are suggestive. If these examples really do represent something parallel to the distinction between Sinhala *də* and *hari* indefinites, it is interesting that Japanese encodes the pragmatic distinction morphosyntactically rather than lexically. Note that the distinction is only possible in certain contexts (e.g. where the *wh*-word takes a postposition or is an argument of a higher noun), as *ka* appears not be able to attach directly to the right of a case-marker like *ga*.

Japanese *ka* can also form declarative disjunctions like Sinhala *hari* and Malayalam *-oo*, as in example (49)

(49) John-**ka** Bill-**ka**-ga hon-o katta.
John-***ka*** Bill-***ka***-NOM book-ACC bought.
'John or Bill bought books.' (Kuroda 1965: 85)

Japanese alternative questions are somewhat more complicated, involving *ka* appearing after the disjuncts (as in a declarative disjunction), but with an additional Q-particle occurring clause-finally (perhaps as a sort of 'scope-marker'; see Fukutomi 2006), as in (50).

(50) John-wa coffee **ka** ocha **ka** docchi-o nonda no
John-TOP coffee ***ka*** tea ***ka*** which-ACC drank *no*
'Which of these two things did John drink: coffee or tea?' (Fukutomi 2006)

Japanese, like Sinhala, thus exhibits a variety of Q-particles. Some of this variation is tied to politeness: *ka* is the most polite form, with *no* occupying a middling level of politeness above *kai* and *ndai* (the last of these is apparently appropriate only to informal male speech, see Miyagawa 1998); the particle *kadooka* apparently is neutral with respect to politeness (see Ginsburg (2009: 81)). Only *ka* and *no* can occur with both *wh-* and yes/no-questions; *kai* and *kadooka* are restricted to yes/no-questions and *ndai* to *wh*-questions. There are also differences with respect to which particles may occur in main and embedded clauses: *ka* may occur in both matrix and embedded clauses, while *no*, *ndai*, and *kai* are restricted to matrix clauses, and *kadooka* to embedded clauses only. Most relevantly for the purposes of the present study, only *ka* occurs in the formation of indefinites and declarative disjunctions.

2.4 Tlingit

Tlingit is classified as a member of the Na-Dene language phylum, which also contains Eyak and Athabaskan languages such as Navajo and Apache (Campbell 1997). It is spoken primarily in the southeastern part of Alaska, as

well as in areas of British Columbia and the Yukon Territory; and has only about 300–400 living native speakers (Cable 2007: 41–43). Tlingit word order is most typically (S)OV, but (S)VO ordering is also common, and in general word order is fairly free (Cable 2007: 56).

Tlingit, like Sinhala, displays morphologically-distinct Q-particles whose distribution is regulated largely by differences in syntactic environment. There are at least three distinct Q-particles: *gé*, appearing in yes/no and alternative questions; *sá*, appearing in *wh*-interrogative and indefinites; and *khach'u*, appearing in declarative disjunctions.

I have only two examples of *gé* in Tlingit ((51), (58)), but, based on these examples, *gé* appears to be a second-position particle.

(51) Lingít **gé** x̲'eeya.áx̲ch?
Tlingit ***gé*** you.understand.it
'Do you speak Tlingit?' (Cable 2007: 74)

Tlingit *wh*-questions, as in English, involve fronting of the *wh*-word; as in Sinhala and Malayalam, the Q-particle *sá* occurs following the *wh*-word, as shown in example (52).

(52) Daa **sá** aawax̲áa i éesh?
what ***sá*** he.ate.it your father
'What did your father eat?' (Cable 2007: 75)

The Tlingit particle *sá* also forms indefinites in a limited set of circumstances. It appears to freely form NPI-indefinites, as in (53), and free-choice indefinites, as in (54).

(53) Tlél goodéi **sá** x̲wagoot.
not where.to ***sá*** I.went
'I didn't go anywhere.' (Cable 2007: 73)

(54) Kéet a x̲á daa **sá**.
killer.whale he.eats.it what ***sá***
'A killer-whale will eat anything.' (Cable 2007: 66)

Sá also appears to be able to form plain existential indefinites when followed by the focus particle *wé*, (55), or preceded by *ch'a* "just", (56). Whether these indefinites signal any degree of speaker ignorance, I do not know.

(55) Daa **sá**.wé yóo dikéenax̲.á
what ***sá***.FOC-PART yonder far.out.across.one
'There was something up there.' (Nyman & Leer 1993: 14, cited in Cable 2007: 107)

(56) Ch'a daa **sá** aagáa k̲ukk̲watées'. . .
just what ***sá*** it.for I.will.search . . .
'I'll look for something there.' (Nyman & Leer 1993: 180, cited in Cable 2007: 107)

Declarative disjunctions involve yet a third particle, *khach'u*, as shown in example (57).[16]

(57) Tlél aadóoch sá kóox awuxhá **khach'u** cháayu awdaná.
not who.ERG *sá* rice ate ***khach'u*** tea drank
"Nobody ate rice or drank tea." (Seth Cable, p.c.)

[16]The *sá* of (57) is part of the NPI indefinite *aadóoch sá* "nobody" (see above discussion) and plays no role in the formation of the disjunction.

Finally, alternative questions have the appearance of yes/no questions, in that *gé* appears in the second position, with the second and subsequent disjuncts being followed by an element *gwáa*, as shown in (58).

(58) Káxwei **gé** i tuwáa sigóo, cháau **gwáa**, héen **gwáa**?
coffee ***gé*** you.want, tea ***gwáa***, water ***gwáa***
"Do you want coffee, or tea, or water?" (Seth Cable, p.c.)

Note that though Tlingit is like Sinhala in having a set of Q-particles whose distribution depends on syntactic environment, the relevant syntactic environments are very different.

2.5 Summary

Thus, crosslinguistically, what has come to be known as a "Question" or "Q-particle" (cf. Baker 1970; Hagstrom 1998; Cable 2007)—nomenclature notwithstanding—surfaces in a wide variety of syntactic environments, occurring not only in interrogative contexts, but also in disjunctions and in the formation of certain types of indefinites. See Table 2.1.[17]

	Mod. Sinh	**Old Mal**	**Mod Mal**	**Tlin**	**Jap**
y/n-ques.	də	-oo	-oo	gé	ka, no, kai, kadooka
wh-ques.	də	-oo	—	sá	ka, no, ndai
wh-indef.	də (aff.), hari (aff.), vat(neg.)	-oo	-oo	sá	ka
decl. disj.	hari (aff.), vat (neg.)	-oo	-oo	khach'u	ka
interr. disj.	də	-oo	-oo	gé. . . gwáa	[ka]

Table 2.1: Distribution of Q particles in Sinhala, Malayalam, Tlingit, and Japanese (repeated)

As indicated previously, I will argue that the crosslinguistic uniformity of the set of possible Q-particle contexts can be explained by positing a single, unified denotation for Q-particles, namely as variables over choice functions (see Chapters 6, 7). The crosslinguistic differences in the distribution of Q-particles within particular languages can then be accounted for by supposing that different Q-particles may bear different formal syntactic feature specifications—and it is these features which determine that Q-particle's licit environments.

The next chapter lays out the basic syntactic formalism adopted and presents an in-depth analysis of the syntax of Q-particle in modern colloquial Sinhala.

[17]Square brackets indicate some additional complication.

Chapter 3

Syntax of Sinhala Q-particle constructions

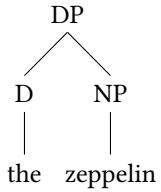
> 'Not entirely, dear Adso,' my master replied. 'True, that kind of print expressed to me, if you like, the idea of "horse", the *verbum mentis*, and would have expressed the same to me wherever I might have found it. But the print in that place and at that hour of the day told me that at least one of all possible horses had passed that way. . . '
>
> —William of Baskerville to his pupil Adso of Melk,
> in Umberto Eco's *The name of the rose*

In this chapter I provide a formal syntactic account of the syntax of Sinhala Q-particle constructions. I begin with a brief discussion of the basic syntax assumed in Section 3.1 and an overview of the basic syntactic properties of modern colloquial Sinhala in Section 3.2. I then discuss a more general theory of possible base-positions for Q-particles, building on Cable (2007), and propose a preliminary account of the possible base-positions of Q-particles in Sinhala in Section 3.3. Sections 3.4–3.7 provide specifics of the syntax of various Q-particle constructions in Sinhala.

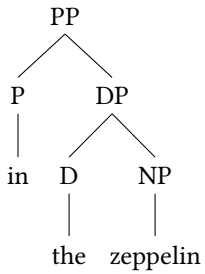
3.1 Preliminaries: Minimalism, basic clausal structure, and formal syntactic features

In this study the syntactic formalism adopted is essentially that of the Minimalist Program (Chomsky 1995, 2001, 2004, 2007). Using this formalism, a syntactic derivation proceeds in several steps. Firstly, lexical elements are selected from the lexicon, where the lexicon consists of "atomic elements, lexical items LI, each a structured array of properties" (Chomsky 2007: 6). Then these elements are Merged to create a binary-branching structure. So if the numeration (the set of elements chosen from the lexicon) contains {the, zeppelin, in}, then we may Merge *the* and *zeppelin* to create (1) and then Merge *in* with the output of the first Merge, as in (2).

(1)

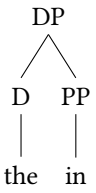


(2)



The set of possible Mergers can be constrained by positing that elements enter the derivation with certain selectional requirements, e.g. determiners select for nouns.[1] Such a restriction rules out Merging, for instance, *the* and *in*:

(3) *



At some point the derivation is Spelled-out; that is, it is handed off to both the phonological module and the semantic module of the grammar, the PF (phonological form) and LF (logical form) interfaces, respectively (re-named the sensorimotor and conceptual-intentional interfaces, respectively, in Chomsky 2008).[2] Recent work (e.g. Chomsky 2001, 2008) suggests that the derivation is not Spelled-out all at once and only when the numeration is empty, but rather that the Spelling-out (or Transfer) takes place in "phases". The distinction between these two approaches will not concern us here for the most part.

Note that—in additional to local dependencies like selectional requirements—elements may bear other features which are dependent on the presence of certain other features on some other element in the derivation. That is, Chomsky (2001) suggests that in addition to (semantically) interpretable features, elements may also bear uninterpretable features. Uninterpretable features cannot be interpreted by the interfaces (LF & PF) and thus the presence of any uninterpretable features in the derivation when it is Spelled-out causes the derivation to crash. Uninterpretable features are features with no value—in contrast to interpretable features, which are always valued. Uninterpretable features may pick up a value in the course of the derivation, at which point they are deleted and thence present no interpretation problem for the interfaces.

Only unvalued features are active, and only active features can Probe. Note that, unlike Chomsky, I do not assume that only elements bearing active (that is, unvalued) features can be the TARGET of a Probe. An element may Probe within its c-command domain,[3] therefore an element bearing an unvalued uα[ ] feature can only acquire a value by Probing when the element bearing a matching feature occurs lower in the tree, as in example (4), but not when a matching instance of the feature occurs higher in the tree, as in (5).[4] As shown by (4), when an element bearing an unvalued feature Probes and Agrees with an element bearing a valued instance of that feature, the unvalued feature acquires a matching value and can then be deleted, as in (4b).

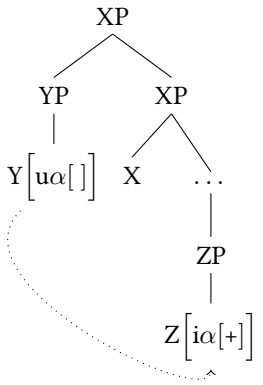
[1] In some recent analyses, e.g. Chomsky (2005), selectional properties are recast in terms of "edge features".

[2] In the remainder of this study I usually refer to these interfaces as PF and LF, for familiarity's sake.

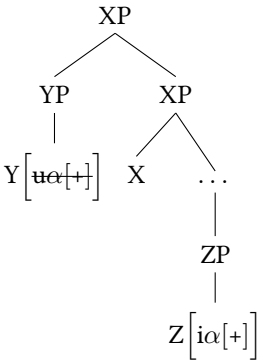
[3] There are further restrictions on the domain within which an element may Probe imposed by the Phase-Impenetrability Condition, though these will not concern us here for the most part: see Chomsky 2000, Chomsky 2004.

[4] I indicate uninterpretable features with a *u-* prefix, e.g. uα, and interpretable features with an *i-* prefix, e.g. iβ.

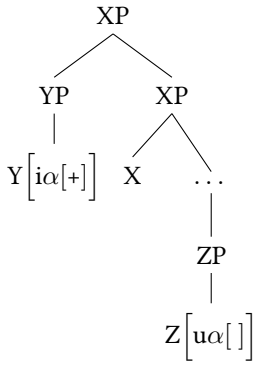
(4) a.



b.

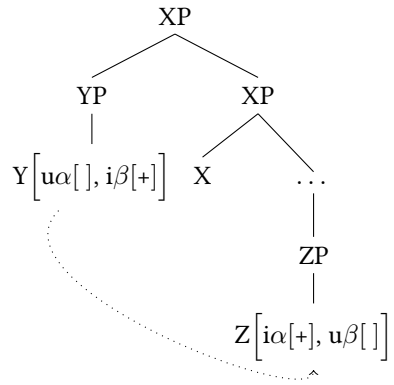


(5) *

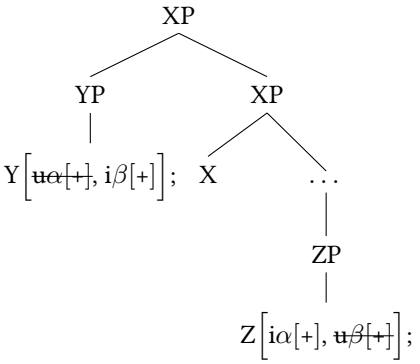


Additionally, I assume that when an Agree relation is established between two elements, Y, Z, any unvalued feature of either Y or Z, matching any valued feature on the other element, acquires a matching value. This is illustrated by example (6). Here Y bears an unvalued feature uα and a valued feature iβ, and being thus Active, Probes and locates Z, which bears a valued instance of iα and an unvalued instance of uβ. In this configuration, not only does the uα feature of Y acquire a value matching that of the iα feature of Z, but the uβ feature of Z also acquires a value matching that of the iβ feature of Y, as shown by (6b).

(6) a.

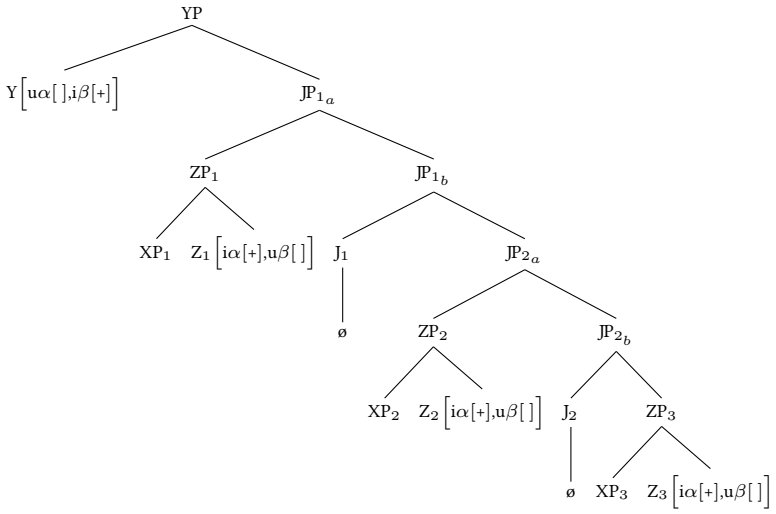


b.

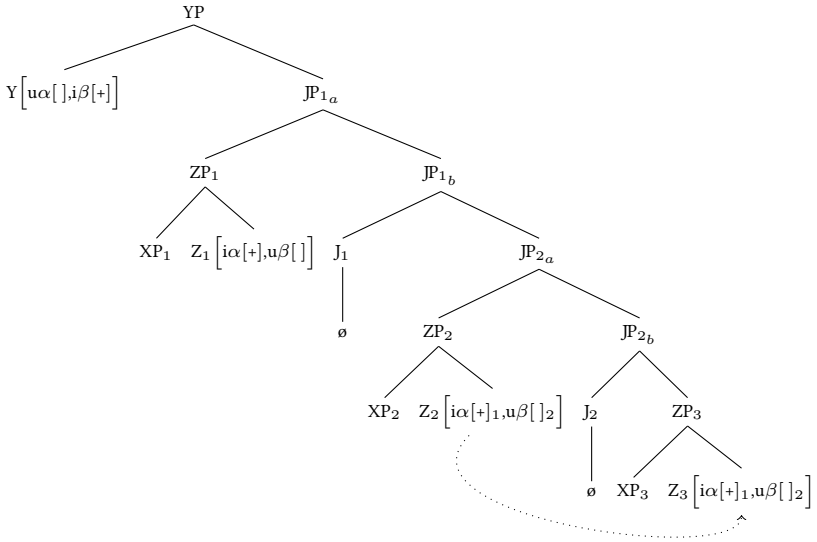


I do adopt one aspect of the feature-valuation system of Pesetsky & Torrego (2007), namely the notion of 'feature-sharing', so that the result of Agree is actually the sharing of a value by both instances of the matching feature. One important consequence of this move is that Agree between two unvalued features is not vacuous: if two unvalued features have established a feature-sharing via Agree, then any subsequent Agree relation which is established between either of the 'shared' unvalued features and a valued instance of that feature will result in *both* of the unvalued features acquiring a value. Consider the tree in (7) below, where subscripts on features indicate a shared value. In (7b) an Agree relation is established between Z_2 and Z_3, and consequently both features become 'shared' (indicated by the subscripting); the same occurs at the next step, as shown in (7c), between Z_1 and Z_2. Thus, when an Agree relation is established between Y and Z_1 in (7e), the result is that not only is Y's uα feature valued, but also the uβ feature receives a value on Z_1 and—due to the feature-sharing—also on Z_2 and Z_3.

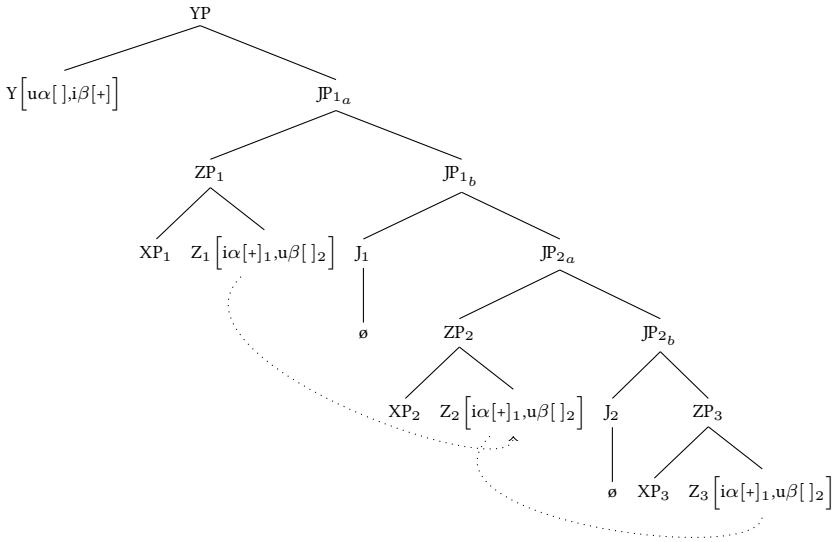
(7) a.


YP
Y [uα[ ],iβ[+]]
JP1a
ZP1
JP1b
XP1
Z1 [iα[+],uβ[ ]]
J1
JP2a
ø
ZP2
JP2b
XP2
Z2 [iα[+],uβ[ ]]
J2
ZP3
ø
XP3
Z3 [iα[+],uβ[ ]]

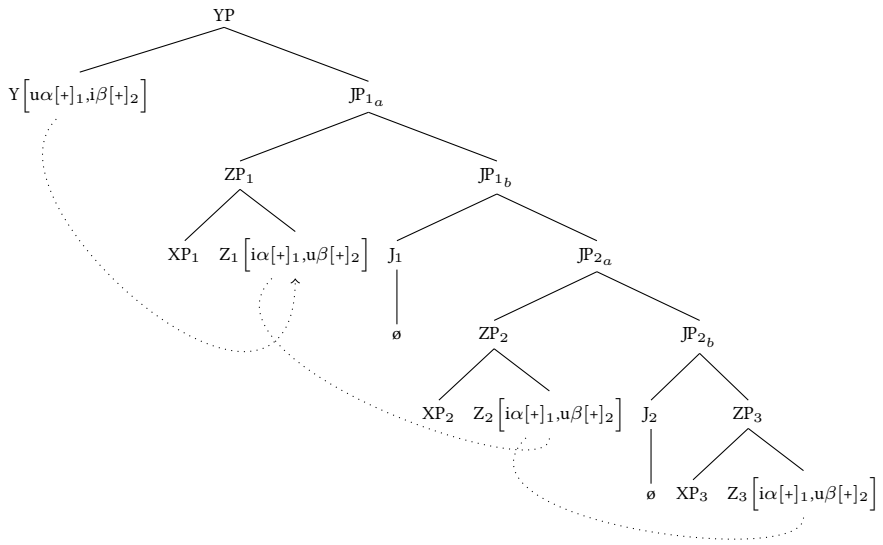
b.


YP
Y [uα[ ],iβ[+]]
JP1a
ZP1
JP1b
XP1
Z1 [iα[+],uβ[ ]]
J1
JP2a
ø
ZP2
JP2b
XP2
Z2 [iα[+]1,uβ[ ]2]
J2
ZP3
ø
XP3
Z3 [iα[+]1,uβ[ ]2]

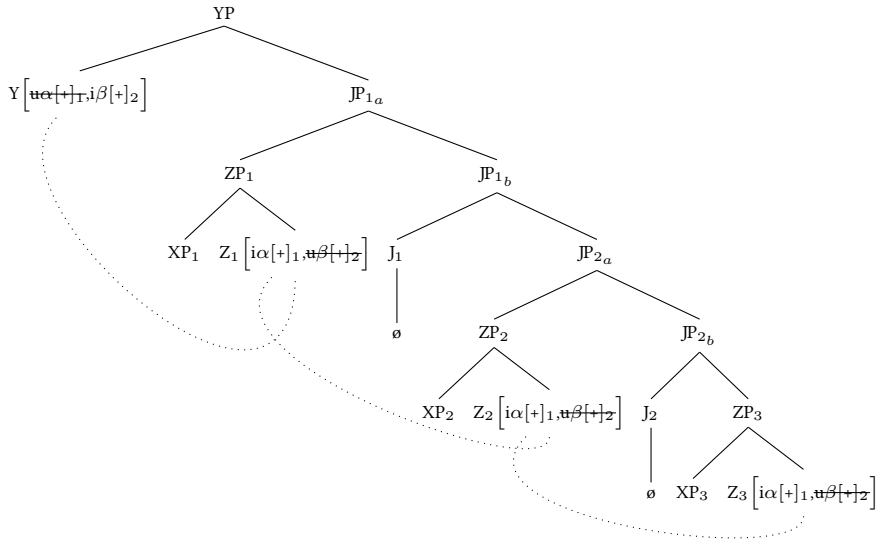
c.


YP
Y [uα[ ],iβ[+]]
JP1a
ZP1
JP1b
XP1
Z1 [iα[+]1,uβ[ ]2]
J1
JP2a
ø
ZP2
JP2b
XP2
Z2 [iα[+]1,uβ[ ]2]
J2
ZP3
ø
XP3
Z3 [iα[+]1,uβ[ ]2]

d.



e.



The merger operations shown above in (1) and (2) are instances of "external Merge" (Chomsky 2001), which is an operation which introduces new material into the derivation (taken from the numeration). In addition there is also an operation of "internal Merge" (Chomsky 2001)—called Move in the older generative framework—which does not introduce any new material into the derivation. An element with an unvalued/uninterpretable feature which Probes and Agrees with an element bearing a matching valued/interpretable instance of that same feature may serve as a trigger for "internal Merge" movement. Formally this involves the Probe bearing an EPP feature (Chomsky 2000, re-termed "edge feature" in Chomsky 2008), triggering the Remerger of the element which the Probe Agrees with. For example, entertaining the idea that in English interrogative complementisers bear unvalued uWh[ ] features and also EPP/edge-features can be seen as explaining the overt "fronting" of *wh*-words in English. That is, assuming the derivation has reached the stage:

(8) [IP John ate what],

at which point an interrogative COMP is merged:

(9) $[_{CP}$ COMP^{int} $[_{IP}$ John ate what]]

Assuming COMP^{int} bears uWh[ ] and an EPP/edge-feature and that *what* bears a valued iWh[+] feature, when COMP^{int} Probes in order to value its uWh[ ] feature; its EPP/edge-feature triggers Remerger of *what* at this point of the derivation, resulting in *what* being Remerged into SpecCP:

(10) $[_{CP}$ what$_i$ COMP^{int} $[_{IP}$ John ate t$_i$]]

Assuming the syntactic formalism developed in this section, I turn in the next section to the examination of the basic syntax of modern colloquial Sinhala.

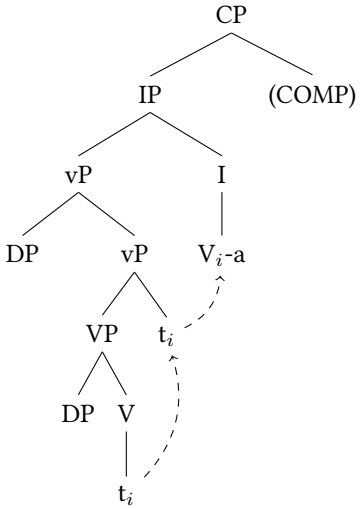
3.2 The basic syntactic structure of Sinhala

Sinhala displays a basic SOV word order, with focussed constituents (including *wh*-words, their associated Q-particle, and any material intervening between the two) being optionally displaced to the right of the surface position of the verb—a position which is generally the rightmost position of the clause (see further Chapter 4).

In order to account for the basic word order of Sinhala, I assume that VP, vP, IP, and CP are left-branching.[5] As discussed in Chapter 4, I propose that the *-e* focus-associated verbal ending of modern colloquial Sinhala resides in the head of FocusP. I assume FocusP to be left-headed, with focussed elements optionally undergoing overt movement to SpecFocusP, which thus linearly follows the *-e* morpheme.

Let us then assume the following abstract structures for "non-focussing" (11) and "focussing" (12) constructions in Sinhala (see Chapters 4 and 5).

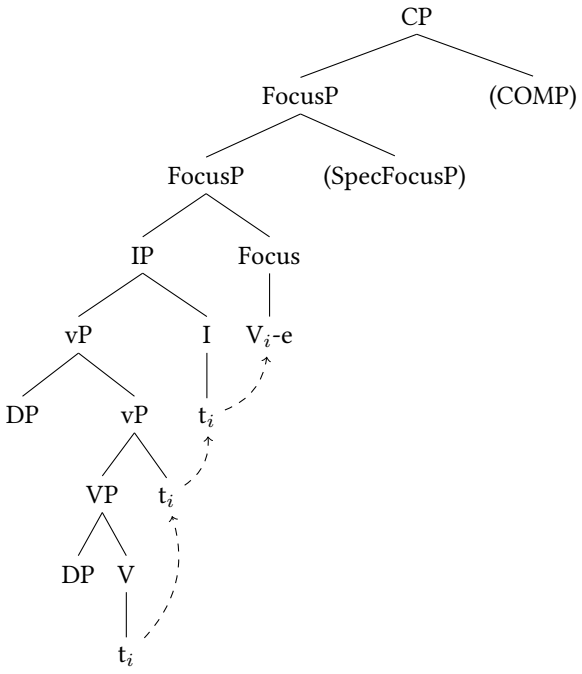
(11)



In a neutral construction like (11), the head of VP raises successively to v and then to I—accounting for the verb-final basic structure of Sinhala.

[5]Certain other phrases, such as NP, DP, and JP ("Junction Phrase", see below Section 3.5 below and Chapter 10) are right-branching.

(12)



In a focus construction, like (12), we find an additional projection, the Focus projection, in which resides the focus-associated *-e* morpheme. Here the verb raises from V to v to I to the head of FocusP, where it picks up the *-e* morpheme.[6]

3.3 The base position of Q in Sinhala and other languages

Previous accounts, including those of Hagstrom (1998) and Kishimoto (2005), have analysed the Q-particles of both Japanese and Sinhala as elements which adjoin to *wh*-words or to constituents containing *wh*-words. I follow Cable (2007) in assuming that while this is the correct analysis for languages like Japanese, in other languages, including Sinhala, English, Tlingit, and German, Q-particles should be analysed as heading their own projections.

Thus in QP-Projection languages, where Q is targeted for movement, the entire QP is moved, rather than just the Q-particle itself. This, Cable argues, underlies *wh*-pied piping like that in (13a) (contrast with the ungrammatical (13b)).

(13) a. $[_{QP}$ What friend of his mother's] did John see t_i?
b. *What$_i$ did John see $[_{QP}$ t_i friend of his mother's]?

In QP-Projection languages with overt Q-particles, like Sinhala, we see that when Q-related movement occurs, it is the entire QP which moves, rather than the Q-particle alone. Observe the sentences in (14): in (14a) the *wh*-word and its associated Q-particle remain *in situ*; in (14b) the *wh*-word and its associated Q-particle both move to the focus position right of the verb; but the Q-particle alone may not move to the focus position, as shown by the ungrammatical (14c).

[6] Alternatively, rather than raising from I to Focus, we might assume that the verb remains in I pre-Spellout and that some sort of (morpho)phonological restructuring takes place post-Spellout which concatenates the V+v+I element and the Focus *-e* element. This could be analysed in terms of Distributed Morphology (see Halle & Marantz 1993, Embick & Noyer 2001, Embick & Noyer 2007, amongst other).

(14) a. Chitra [$_{QP}$ **monəwa də**] gatte?
Chitra [$_{QP}$ **what** ***də***] buy.PAST.E

b. Chitra t_i gatte [$_{QP}$ **monəwa də**]$_i$?
Chitra t_i buy.PAST.E [$_{QP}$ **what** ***də***]$_i$

c. *Chitra [$_{QP}$ **monəwa** t_i] gatte **də**$_i$?
Chitra [$_{QP}$ **what** t_i] buy.PAST.E ***də***$_i$
"What did Chitra buy"

In contrast, in Japanese *wh*-questions it is in fact the Q-particle alone which moves to the clause-final position, as shown by (15).

(15) John-ga **nani**-o kaimasita **ka**?
John-NOM **what**-ACC bought.POL ***ka***
"What did John buy?"

Cable (2007) presents further differences between Q-Projection and Q-Adjunction languages which concern possible positions of Q-particles. For example, in Tlingit, which Cable deems a Q-Projection language, a Q-particle may not intervene between a post-position and its complement:

(16) (from Cable 2007: 100)

a. Tléil **aadóo** teen **sá** x̲wagoot.
not **who** with ***sá*** I.went
"I didn't go with anyone."

b. *Tléil **aadóo sá** teen x̲wagoot.
not **who** ***sá*** with I.went

But such configurations are possible in Q-Adjunction languages like Japanese, as shown by:

(17) a. Taroo-wa doko-**ka**-e itta.
Taro-TOP where-***ka***-to went.
"Taro went somewhere."

b. Taroo-ga [dono tosi]-**ka**-e ryoko sita-rasii.
Taro-NOM [which city]-***ka***-to travel did-seems.
"Taro seems to have travelled to some city." (Cable 2007: 171)

To account for these differences, Cable (2007: 122) proposes the following constraint on Q-Projection languages:

(18) **The QP-Intervention Condition**
A QP cannot intervene between a functional head and a phrase selected by that functional head. (Such an intervening QP blocks the selectional relation between the functional head and the lower phrase.)

Assuming, as Cable (2007) does, that the particles associated with yes/no and alternative questions (like Tlingit *gé*) are not "true" Q-particles, the constraint in (18) correctly predicts the distribution of Tlingit particle *sá*.[7]

Though (18) appears promising and potentially well-motivated,[8] it presents certain empirical difficulties. One

[7]Correctly disallowing *sá* from positions between a post-position and its complement, a possessor and the possessed NP, and on the right edge of a matrix clause.

[8]Cable (2007: 122n71) remarks that if we accept a distinction between s-selection and c-selection (Grimshaw 1981; Pesetsky 1982), and assume that only functional heads c-select (select for a particular syntactic category) for their arguments while lexical heads only s-select (select for a semantic type) for their arguments, a QP will be unable to satisfy all of the selectional requirements of a functional head.

major difficulty is that (18) is tenable only if we accept Cable's assertion that particles associated with disjunctions, yes/no- and alternative-questions are not "true" Q-particles, even where—as in the case of Sinhala, Japanese, and Malayalam—the particle which appears in these contexts is form-identical with the particle appearing in *wh*-questions. In this study, I argue against this assertion and for what I would consider to be the null-hypothesis, namely that these phonologically-identical elements are in fact one and the same particle. Given that, for instance, Sinhala *də* routinely appears in clause-final position in yes/no-questions as in (19), Cable's (18) is obviously problematic if we accept that the particle *də* which appears in yes/no-questions is the same *də* that appears in *wh*-questions.

(19) Chitra ee potə kieuwa **də**?
Chitra that book read-A ***də***
'Did Chitra read that book?' (Kishimoto 2005: 11)

However, even if we were to accept Cable's claim that only the *wh*-associated particles are "true" Q-particles, Sinhala is still problematic for the condition in (18), given that even the *də* associated with *wh*-questions can occur in positions which (18) does not allow. Recall from Section 2.1 above that under special circumstances, *wh*-questions may be formed in which the verb takes the "neutral" A-form rather than the "focussing" E-form otherwise obligatory in *wh*-questions, and that in such constructions the Q-particle *də* appears not adjacent to the *wh*-word, but rather clause-finally.[9] The relevant examples are repeated below as (20) and (21), respectively.[10]

(20) oyaa [[kauru aawa **də**] kiyəla] dannəwa də?
you [[who came-A ***də***] that] know-A *də*
"Do you know who came?"

(21) kiidenek potə kieuwa **də**?
how.many book read-A ***də***
"How many (people) read the book?"

The configuration in (20) might be argued to conform to the condition in (18), if we analyse *kiyəla* "that" as a lexical rather than a functional element—although that would seem to require special pleading.

In (21), however, the Q-particle *də* unarguably occurs in the clause-final position of the matrix clause—a position which (18) rules out for Q-Projection languages like Sinhala. Note, again, that the configurations of (20), (21) are only possible where the verb appears in the A-form:

(22) *oyaa [[kauru aawe **də**] kiyəla] dannəwa də?
you [[who came-E ***də***] that] know-A *də*
("Do you know who came?")

(23) *kiidenek potə kieuwe **də**?
how.many book read-E ***də***
("How many (people) read the book?")

We may correctly predict the possible base positions of Q-particles in Sinhala by assuming a syntactic restriction to the effect that QP can select any of the following: DP, AdvP, IP,[11] in addition to the semantic restriction that QP must

[9] Recall also that such constructions differ pragmatically from the normal constructions in that they do not presuppose that there is at least one value which satisfies the proposition.

[10] Again, we are here concerned with the position of *də* in the subordinate clause, not the *də* of the matrix yes/no question.

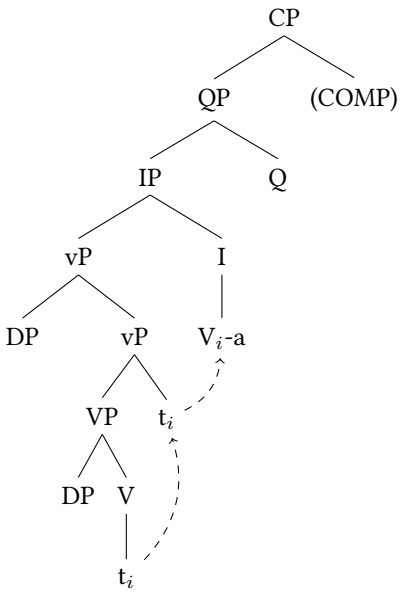
[11] Whether this can be reconciled in some way with Cable's (2007) proposal of a QP-Intervention Condition (18) I leave for future research. The fact of the matter is that Cable's QP-Intervention Condition as it stands, regardless of one's position on the status of particles in non-*wh*-environments, cannot account for the complete range of Sinhala data.

select an element of the right semantic type (i.e. a set to which a choice function can apply)—see Chapter 9.1.

When QP selects IP, then *dɔ* appears clause-finally on the surface. We have seen this position of *dɔ* in "neutral" yes/no-questions and in "special" non-presupposition *wh*-questions. However, whenever *dɔ* appears clause-finally, the verb must appear in the A-form (and not the "focussing" E-form). It remains to account for this restriction.

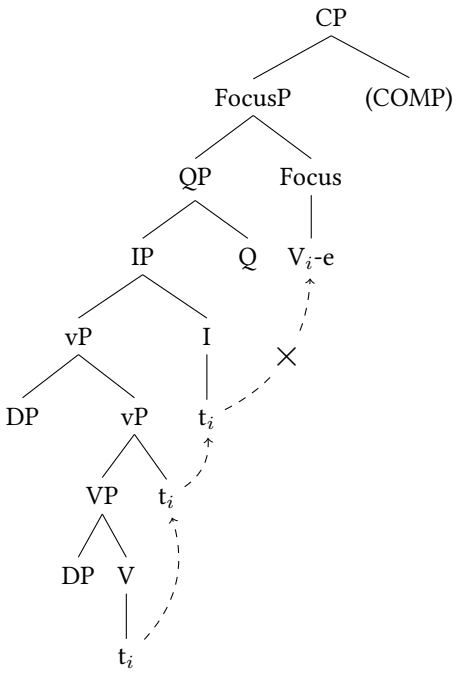
In "neutral" sentences where the verb appears with the A-form, QP may select for IP, as shown below in (24).

(24)



However, in focussing constructions where the verb appears in the E-form, if the QP selects for IP, then the raising of I to Focus is blocked, as shown below in (25).

(25)



I assume that the structure in (25) involves a violation of the Head Movement Constraints (see Travis 1984: 181, Chomsky 1986)—since the head of IP attempts to move over the head of QP to reach FocusP—accounting for its

ungrammaticality.[12]

3.4 The syntax of modern colloquial Sinhala *wh*-questions

Given the constraints on the base-position of Q-particles discussed in the above Section 3.3, the remainder of the syntax of *wh*-questions is relatively simple. Q may select any category from DP, AdvP, or IP which contains a *wh*-word (on the latter constraint see Chapter 6).

One point which requires further explanation is why *wh*-words but not their associated Q-particles may be internal to islands (relative to the CP over which the *wh*-word takes scope), as in the examples below repeated from Chapter 2.1.

(26) *Chitra [[[[Ranjit [monəwa] də gatta] kiənə] kaṭəkataawə] æhuwe ?
Chitra [[[[Ranjit [what] *də* bought.A] that] rumour] heard-E
'What$_i$ did Chitra hear the rumour that Ranjit bought t$_i$?'

(27) Chitra [[[Ranjit monəwa gatta] kiənə] kaṭəkataawə] də æhuwe ?
Chitra [[[Ranjit what bought.A] that] rumour] *də* heard-E
'What$_i$ did Chitra hear the rumour that Ranjit bought t$_i$?'

These data indicate the existence of a dependency relation between Q and C which cannot be satisfied if Q is internal to an island. I assume that this dependency relation involves feature-valuation; specifically positing that interrogative complementisers bear valued iInt[+] features and unvalued uQ[ ] features, the latter of which must be valued via Agree with *də*, which bears a valued iQ[+] feature. The details of this feature-based analysis are provided in Chapter 9.3.[13]

There is, however, an additional issue involved in *wh*-interrogatives. Recall also that, like focussed elements, *wh*-words along with their associated Q-particle may optionally dislocate to the right of the verb of the clause over which they take scope, as in example (28b), contrast with example (28a), both repeated from Chapter 2.1.

(28) a. Chitra monəwa də gatte?
Chitra what *də* bought-E
'What did Chitra buy?'

b. Chitra gatte monəwa də?
Chitra bought-E what *də*
'What did Chitra buy?'

Taken together, these data suggest that there is an additional dependency relation between QPs (more properly, between the heads of QPs) in *wh*-questions and some element within the CP-layer. Since non-*wh* elements which are focussed also optionally undergo an identical movement operation, it is reasonable to assume that the dependency relation involves Focus. Specifically, I assume that *də* in *wh*-questions obligatorily bears an interpretable iFocus[+]

[12] Alternatively, if we assume that a post-Spellout PF operation—like Local Dislocation (see Embick & Noyer 2001, Embick & Noyer 2007)—is responsible for the concatenation of I (the verb) and Focus (the *-e* suffix), then the intervening material, i.e. *də* in the head of QP, would block this morphophonological operation.

[13] This can be formalised in terms of a phase-based derivation (Chomsky 2001, 2005, 2008). We can take relevant phases to consist of CP, vP, and DP (on DPs as phases, see Chomsky 2005). A phase is Spelled-out/Transferred only once the next relevant phase is reached (Chomsky 2001: 13). Assuming that a phase HP which complete (but not yet Spelled-out), in the next phase ZP, the domain of H is not accessible (to operations like Agree), only the head H itself and its Edge are accessible (Chomsky 2001: 14). In the case of (27), this implies that the QP *Ranjit. . . kaṭəkataawə də* must have first been raised (covertly) to the Edge of vP, from where it would be accessible to the head of CP in the next phase. In the case of (26), the QP *monəwa də* is trapped inside of the complex NP island, and thus is unable to raise to a position from which is accessible to operations in subsequent phases.

feature and an unvalued uExist[ ] feature (see further below Chapter 4.2), and that the head of FocusP bears an uninterpretable uFocus[ ] feature and an interpretable iExist[+] feature. This allows us to capture the entailment that *wh*-questions trigger the appearance of the *-e* morpheme (which is the head of the FocusP).[14] Optionally, the head of FocusP may bear an EPP feature, which triggers movement of the QP to SpecFocusP, accounting the possibility of (28b).

Kishimoto (1997, 2005) proposes a rather different analysis (also adopted by Hagstrom 1998) which attempts to unite the presence of *-e* verbal marking and the appearance of *də* in interrogatives. Specifically he suggests that finite verbs bearing *-e* endings entails that the verb bears an uninterpretable [+Q] feature, which can be checked by the Q-particle *də* (Kishimoto 2005: 22). Where *də* moves overtly to SpecCP, this [+Q] is checked and deleted, resulting in the verb bearing an *-a* ending; where no such overt movement takes place, the [+Q] feature surfaces in PF as the *-e* ending. Kishimoto proposes this analysis to account for the co-existence of examples like (29a), the standard case where the verb appears with the *-e* ending and *də* occupies a position within its scope, with examples like (29b) (which, again, are permissible only under special circumstances, e.g. when the verb in question is the complement of *dannəwa* "know"), where the verb appears with the "neutral" *-a* ending and is immediately followed by *də*.

(29) a. Ranjit [[[kau də] aawe] kiyəla] dannəwa.
Ranjit [[[who *də*] come.PAST.E] that] know.PRES.A
'Ranjit knows who came.'

b. Ranjit [[[kauru aawa] də] kiyəla] dannəwa.
Ranjit [[[who came-A] *də*] that] know-A
'Ranjit knows who came.'

This account is problematic in a number of respects. Firstly, as shown in Chapter 4 below, verbs with the *-e* suffix also appear in non-interrogative contexts: they appear wherever a focussed element is present. This makes the assumption that *-e* is some sort of PF-reflex of an unchecked [+Q] feature difficult to maintain, since *-e* can appear even when there is no interrogative element present in the derivation. Secondly, allowing PF to "interpret"/Spellout unchecked features is theoretically unappealling. In the current analysis, unchecked features sent to the interfaces cause the derivation to crash, which would make this proposal especially problematic. Kishimoto (2005) adopts an earlier Minimalist analysis (of the Chomsky 1995-era) in which features may be either strong or weak, with weak features allowing checking/movement to be Procrastinated until LF. Even assuming this version of Minimalism, it is not clear how PF would be able to see and interpret unchecked features. If the features are "weak" and thus do not trigger a crash within PF (though they must obtain a value within LF), then presumably they should also be invisible with respect to PF, and not subject to interpretation.

I thus propose that it is preferable to treat *-e*, not as the reflex of an unchecked [+Q] feature, but rather as the head of FocusP. This allows us to naturally express the fact the *-e* appears in focussing contexts, and not only in interrogatives, and does not require any dubious PF-interpretation of unchecked features. The possibility of (29b) alongside (29a) can be accounted for instead by positing an analysis which directly reflects the difference of the two with respect to the presence of an existence presupposition—which I take to be carried by the *-e* morpheme itself. Before this alternative analysis can be developed, however, more needs to be said about the syntax and semantics of focus (see Chapters 4 and 5, respectively). The restrictions on the position of the Q-particle with respect to the presence/absence of the *-e* morpheme have been dealt with in Section 3.3 above; the semantic consequences are

[14] These feature specifications are in fact not entirely obligatory for Q-particles in *wh*-questions. Under special circumstances, described above in Chapter 2.1, *wh*-questions do not carry existence presuppositions, and under these circumstances Q-particles may enter without Focus or Exist features. See further below. In general, however, as in English and other languages, *wh*-questions do tend to involve an existence presupposition (i.e. they tend to presuppose that there is some true answer to the question).

deferred to the appropriate place for their discussion, namely Chapter 6.

At this point I turn to the consideration of alternative questions—which also involve the presence of *də*—and disjunction more generally.

3.5 The syntax of modern colloquial Sinhala alternative questions & disjunctions

The Q-particle *də* also appears in the formation of alternative questions in Sinhala, as in example (30) below.

(30) Gunəpālə **də** Chitra **də** gamət̪ə giyē?
Gunapala ***də*** Chitra ***də*** village.DAT go.PAST.E
'Was it Gunapala or Chitra who went to the village?'

Note here that *-e* verbal marking is obligatory when the disjunct is in the c-command domain of the verb; the *-a* marking is inadmissible in this context, as shown by (31).

(31) *Gunəpālə **də** Chitra **də** gamət̪ə giyā.
Gunapala ***də*** Chitra ***də*** village.DAT go.PAST.A
'Gunapala or Chitra went to the village.'

I account for this fact by assuming that *də*, as it does in (normal) *wh*-questions, also bears iFocus[+] and uExist[ ] features in alternative questions. This is a reasonable assumption given that alternative questions of this sort in Sinhala also carry existence presuppositions (see Chapter 6 below).

The syntax of alternative questions, especially the use of multiple Q-particles (one for each disjunct), requires a novel analysis for disjunction more generally—if the semantically-unified analysis of Q-particles is to be maintained. This analysis is laid out in detail in Chapter 10; here I provide a much briefer sketch.

At first blush, *də* in alternative questions like (30) above, as well as *hari* in non-interrogative contexts like (32), appear to be similar to English *or*, which is traditionally analysed as a "disjunction operator".

(32) Gunəpālə **hari** Chitra **hari** gamət̪ə giyā
Gunapala ***hari*** Chitra ***hari*** village.DAT go.PAST.A
'Gunapala or Chitra went to the village.'

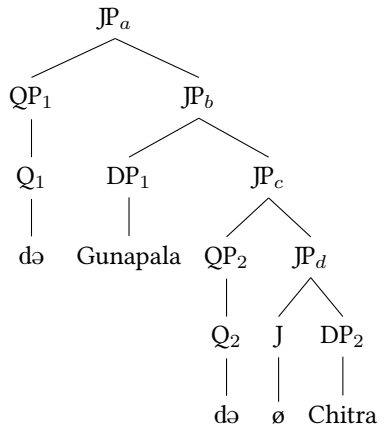
However, given that this thesis assumes a unified semantic denotation for Q-particles across different syntactic contexts as variables over choice functions, we cannot treat *də* or *hari* as "disjunction operators" (in Chapter 10 I argue that English *or*, in fact, should not be analysed as a "disjunction operator" either). This entails that the actual disjunction must be accomplished by some other element. Adopting an analysis similar to that proposed in den Dikken (2006) (see Chapter 10 for further discussion of this analysis and its motivations), I assume that universal grammar makes available a "Junction" element J—which is the actual "(dis)junction operator" in contexts like (30).[15] Such an analysis makes it possible to maintain a semantically-unified treatment of *də* and other Q-particles across different syntactic contexts. See further Chapters 6 and 10.

Under this analysis, the structure of disjunction in (30) would be as shown in (33), with JP containing the disjuncts and QPs adjoining to each disjunct.[16]

[15] The head of JP, at least in English and Sinhala, is phonologically-null.

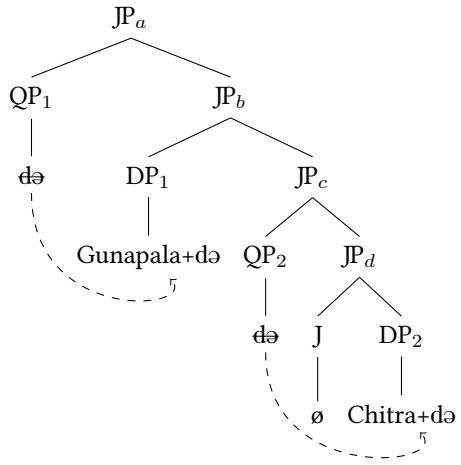
[16] More precisely, each QP adjoins to the minimal JP containing both the head of J and the disjunct with which it is associated.

(33)



The surface ordering of *də* with the respect to the disjunct I treat as a reflex of post-syntactic clitic alignment (on which, see Chapter 10):

(34) "PF-Structure" of (33):



In the next section I examine the use of Q-particles in yes/no-questions.

3.6 The syntax of modern colloquial Sinhala yes/no-questions

I treat yes/no-questions as a special sub-type of alternative question which involves ellipsis.[17] Note that the sentences in (35) are all semantically equivalent, suggesting that (35a) and (35b) can be treated as instances of (35c) where certain constituents have been elided, as represented in (36).

(35) a. Did John drink tea?
b. Did John drink tea or not?
c. Did John drink tea or did John not drink tea?

(36) a. Did John drink tea ~~or did John not drink tea~~?
b. Did John drink tea or ~~did John~~ not ~~drink tea~~?
c. Did John drink tea or did John not drink tea?

[17] See also Han & Romero (2002), Romero & Han (2004).

Thus a yes-no question like (37) can be analysed as involving an elided *or not. . .* component, as shown in (38).

(37) Chitra ee potə kieuwa **də**?
Chitra that book read-A ***də***
'Did Chitra read that book?'

(38) [$_{JP}$ [[$_{IP}$ Chitra ee potə kieuwa] **də**] J [[$_{IP}$ *Chitra ee potə kieuwe næ̀æ̀*] ***də***?]]
[$_{JP}$ [[$_{IP}$ Chitra that book read-A] ***də***] J [[$_{IP}$ *Chitra that book read-E*[18] *nɛ́ɛ́*] ***də***]]
'Did Chitra read that book *(or) did Chitra not read that book*?'

So (22) is really an alternative question, where the alternatives are "Chitra read that book" and "Chitra didn't read that book". In other words, in (neutral) yes-no questions the disjunction takes place at the IP level. Since the disjunction takes place at the IP level, the focussing form of the verb in -E is not available, because if the *-e* morpheme is generated in the head of FocusP, a Q-particle which takes the IP as a complement blocks the required raising of I to Focus (see above Chapter 3.3). Additionally, in "neutral" yes/no questions there does not appear to be any element which bears focus in any case, and thus no trigger for the appearance of E-form of the verb in the first place. Therefore, in "neutral" yes/no-questions the verb must appear in the A-form.

On the other hand, in "focussed" yes/no questions like (39) the disjunction takes place at a lower level, in the case of (39) at the level of the DP. Since "focussed" yes/no questions do involve a focussed element, we find the verb appearing in the E-form. As in the case of (37) the elided element is an *or not. . .* component, where the negation is constituent negation, which applies to the DP (see Chapter 5.4 for discussion of the semantics of constituent negation).

(39) Chitra ee potə **də** kieuwe?
Chitra that book ***də*** read-E
'Was it that book which Chitra read?' (*Ibid.*)

(40) Chitra [$_{JP}$ [[$_{DP}$ ee potə] **də**] J [[$_{DP}$ *ee potə nɛ́ɛ́*] ***də***]] kieuwe?
Chitra [$_{JP}$ [[$_{DP}$ that book] ***də***] J [[$_{DP}$ *that book nɛ́ɛ́*] ***də***]] read-E
'Was it that book *(or) not that book* which Chitra read?'

Again, focussed yes/no-questions like (39) also involve an elided *or not. . .* component, as shown in (40).

3.7 The syntax of modern colloquial Sinhala Q-particle indefinites

Sinhala Q-particle indefinites are form-identical with *wh*-interrogatives, as is the case as well in other languages like Japanese (for which the term "indeterminate particle" was coined, indicating a form used both as an interrogative pronoun and an indefinite pronoun, see Kuroda 1965). In modern colloquial Sinhala, *wh*-interrogative constructions differ morphosyntactically from *wh*-indefinite constructions in that the former trigger the use of the "focussing" *-e* verbal suffix, while in the case of the latter the "neutral" *-a* verbal suffix is employed; compare the declarative in (41a) with the interrogative in (41b).

[18]The appearance of an apparent *-e* verbal form in the negative disjunct is a quirk of the morphosyntax of negation in Sinhala, see below Chapter 4.3.2 for discussion.

(41) a. mokak də wætuna.
what *də* fell-A
'Something (unidentified) fell.' (Gair & Sumangala 1991)

b. mokak də wætune?
what *də* fell-E
'What fell?'

In the case of *wh*-based indefinites, unlike in the case of interrogatives, there seem to be no Q-related dependencies. Semantically and pragmatically, however, these indefinites present a number of interesting features, as discussed in Chapter 7.

3.8 Summary, and why syntax cannot account for the distribution of Q-particles

The syntactic analysis of Sinhala Q-particle constructions developed in this study is couched within a Minimalist theory of syntax, assuming that syntax consists of binary-branching structures built through successive application of the function Merge (including both external Merge, which introduces new elements into the derivation, and internal Merge, which operates over elements already present in the derivation). There may exist both local (head-complement) and more distant dependencies between elements of syntax; these are formally encoded using the notion of syntactic features. Syntactic features may be interpretable/valued or uninterpretable/unvalued; uninterpretable features present when the derivation is shipped off to the interfaces result in an interpretation crash (≈ ungrammaticality). Syntactic elements which enter the derivation with uninterpretable features may value and delete such features if the uninterpretable features can be valued, via the establishment of an Agree relationship with another syntactic element bearing a valued instance of the same feature. Agree relationships are established by an active element (an element which bears an unvalued uninterpretable feature α) Probing within its c-command domain in order to locate an element with an instance of feature α. Such Probe domains may also be limited in terms of syntactic phases (CP, vP, DP etc.).

Given these syntactic assumptions, we see in the case of Sinhala *wh*-interrogatives that there are no dependency relationships between either a *wh*-word and the CP of the clause over which it takes scope or between a *wh*-word and its associated Q-particle. However, there is a syntactic dependency between a Q-particle and the CP of the clause over which the Q-particle's associated *wh*-word takes scope. For disjunctive contexts, including alternative questions, I have posited a novel syntactic treatment of disjunction (similar in some respects to that of den Dikken 2006) in which (dis)junction involves a JP (junction phrase) with a phonologically-null head J (this is discussed in much more detail in Chapter 10). Yes/no-questions I treat as a special subtype of alternative question with ellipsis of certain elements.

Syntactically, the set of contexts in which Q-particles appear in Sinhala (and in Malayalam, Japanese, and Tlingit—see above Chapter 2) is heterogeneous. *Wh*-interrogatives and *wh*-indefinites both involve *wh*-words, but Q-particles also appear in non-*wh*-contexts, such as yes/no and alternative questions—thus the presence of *wh*-words alone cannot account for the presence of Q-particles. Nor can we say that Q-particles occur only in interrogative contexts, since we find Q-particles not only in *wh*-indefinites but also in non-interrogative disjunctions. Syntactically, there is nothing which unites disjunctive, interrogative, and *wh*-contexts. However, these contexts can be united semantically, as I show in Chapters 6, 7, 9.1, and 10, as all being contexts involving an element with Hamblin-type semantics.

We can then turn to the semantic interpretation of the syntactic structures described in this chapter in Chapter 6, which treats the semantics of interrogatives in Sinhala, and Chapter 7, which treats the semantics (and pragmatics)

of indefinites in Sinhala. However—since focus is involved in many of these Q-particles constructions in Sinhala, and since the presence of focus has semantic consequences—before examining the semantics of Sinhala Q-particle constructions, I first provide analyses of the syntactic and semantic structure of focus constructions in Sinhala, in Chapters 4 and 5, respectively.

Chapter 4

Syntax of Sinhala focus constructions

Given that many types of interrogative constructions in Sinhala require the use of the special "focussing" form of the verb appearing with the *-e* suffix, in order to understand the syntactic structure of such interrogatives we must examine the structure of focus constructions in more detail. This chapter therefore provides an account of the syntax of "focussing" constructions (where we find the verb appearing with the *-e* suffix) in modern Sinhala.

I consider the morphological and syntactic evidence which pertain to the structure of modern colloquial and modern literary Sinhala "focus" constructions, arguing that—in contrast to neighbouring Dravidian "cleft" constructions, which are superficially similar—modern colloquial Sinhala "focus" constructions are best analysed as monoclausal, not biclausal.

4.1 Preliminaries

Though the modern colloquial Sinhala verb does not show any overt subject agreement morphology—see example (1) below—it does have a special 'focussing' form that appears when there is a focussed element in the verb's c-command domain; *wh*-constituents are obligatorily focussed as well.[1] The type of construction examined in this chapter, especially when it involves dislocation of the focussed element, is often referred to in the literature on South Asian languages (particularly Dravidian languages and Sinhala) as a "cleft" construction. Such South Asian constructions are not, in fact, structurally equivalent to English clefts or pseudo-clefts, but they bear semantic interpretations comparable to those of an English cleft. The structure assumed for South Asian "clefts" usually involves a nominalised constituent put into a copular relationship with a focussed element; however, I show in this section that such an analysis is not appropriate for modern colloquial Sinhala focussing/"cleft" constructions.[2]

(1) a. mamə gaməṭə yanna
I.NOM village-DAT go.PRES.A
'I go to the village.' [MCS]

b. eyaa gaməṭə yanna
he.NOM village-DAT go.PRES.A
'He goes to the village.' [MCS]

When one of the constituents of the clause bears focus (notated here as a superscripted F), the verb takes the E-form, as shown by (2) and (3) below. Again, there is no morphological realisation of subject-predicate agreement.[3]

[1] In most cases. The exception to this rule is that when the Q-particle *də* has moved to the clause-final position following the main verb, in which case the verb appears in the 'neutral' form, see Chapter 2.1.

[2] Though the ancestor of the modern focus construction was in fact a true (South Asian type) "cleft" in certain earlier stages of its development; see Chapter 11.1.

[3] Following Kishimoto 2005 I gloss the 'focussing' form of the verb, which appears as the suffix *-e/-ē*, as *-E*, and the neutral default form, appearing as *-a/-ā*, as *-A*. The alternation in the endings of the verbs in Colloquial Sinhala between *-a* and *-ā* and *-e* and *-ē* is of no morphological significance, but simply reflects a phonological rule.

(2) a. mamə gamə̣ə^{F} (y/tamay) yanne
I.NOM village^{F}-DAT (EMPH) go.PRES.E
'It is to the village I go.' [MCS]

b. eyaa gamə̣ə^{F} (y/tamay) yanne
he.NOM village^{F}-DAT (EMPH) go.PRES.E
'It is to the village he goes.' [MCS]

(3) a. mamə gamə̣ə^{F} (y/tamay) giyee
I.NOM village^{F}-DAT (EMPH) go.PAST.E
'It is to the village I went.' [MCS]

b. eyaa gamə̣ə^{F} (y/tamay) giyee
he.NOM village^{F}-DAT (EMPH) go.PAST.E
'It is to the village he went.' [MCS]

The focussed element is marked by prominence in intonation. It may be followed by an emphatic particle, such as *y(i), tamaa, tamay*, and the focussed element itself often occurs to the right of the verb, but intonational prominence alone is sufficient to mark focus—so neither dislocation of the focussed element nor the presence of a particle is obligatory. These various possibilities are illustrated in (4) (assuming prosodic focus on *gamə̣ə*).

(4) a. mamə gamə̣ə^{F} yanne
I.NOM village^{F}.DAT go.PRES.E
'It is to the village I go.' [MCS]

b. mamə gamə̣ə^{F}-y yanne
I.NOM village^{F}.DAT-EMPH go.PRES.E
'It is to the village I go.' [MCS]

c. mamə yanne gamə̣ə^{F}
I.NOM go.PRES.E village^{F}.DAT
'It is to the village I go.' [MCS]

d. mamə yanne gamə̣ə^{F}-y
I.NOM go.PRES.E village^{F}.DAT-EMPH
'It is to the village I go.' [MCS]

Let us contrast this pattern with that found in modern literary Sinhala [LS], where the neutral forms of the verb DO show morphological agreement with the subject, in both present, (5), and past, (6) tense.

(5) a. mama gamaṭa yami
I.NOM village.DAT go.PRES.1SG
'I go to the village.' [LS]

b. hē gamaṭa yayi
he.NOM village.DAT go.PRES.3SG
'He goes to the village.' [LS]

(6) a. mama gamaṭa giyemi
I.NOM village.DAT go.PAST.1SG
'I went to the village.' [LS]

b. hē gamaṭa giyāya
he.NOM village.DAT go.PAST.3SG
'He went to the village.' [LS]

As shown below in (7) and (8), Literary Sinhala too has E-forms of the verb which appear in focus constructions. The verbal forms in (5) and (6) above correspond in usage to Colloquial A-forms.

In contrast to sentences with neutral forms, focus constructions in modern literary Sinhala are characterised by two phenomena: (i) the verb shows no agreement morphology but rather appears in an invariant form (in *-e*), and (ii) the logical subject takes accusative case (in modern colloquial Sinhala, as shown above, the subject retains the normal case assigned by the verb, usually nominative). See the modern literary Sinhala present tense focus constructions in (7) and past tense focus constructions in (8).

(7) a. mā yanne gamaṭa^{F} ya
I.ACC go.PRES.E village^{F}.DAT *ya*
'It is to the village I go.' [LS]

b. ehu yanne gamaṭa^{F} ya
he.ACC go.PRES.E village^{F}.DAT *ya*
'It is to the village he goes.' [LS]

(8) a. mā giyē gamaṭa^{F} ya
I.ACC go.PAST.E village^{F}.DAT *ya*
'It is to the village I went.' [LS]

b. ehu giyē gamaṭa^{F} ya
he.ACC go.PAST.E village^{F}.DAT *ya*
'It is to the village he went.' [LS]

Unlike in Colloquial Sinhala, the element *ya* (equivalent to Colloquial *y(i)*) must mark the focussed element, cp. the grammatical MCS (4c), repeated below as (9a), with ungrammatical LS (9b). Further, the focussed element in Literary Sinhala obligatorily appears to the right of the verb and cannot remain in its base position as in Colloquial Sinhala, cp. grammatical MCS (4b), repeated below as (10a), with ungrammatical LS (10b). The grammatical LS construction appears in (11).

(9) a. mamə yanne gamə̣ə̣^{F}
I.NOM go.PRES.E village^{F}.DAT
'It is to the village I go.' [MCS]

b. *mā yanne gamaṭa^{F}
I.ACC go.PRES.E village^{F}.DAT
'It is to the village I go.' [LS]

(10) a. mamə gamə̣ə̣^{F}-y yanne
I village^{F}.DAT-EMPH go.PRES.E
'It is to the village I go.' [MCS]

b. *mā gamaṭa^{F} ya yanne
I.ACC village^{F}.DAT *ya* go.PRES.E
'It is to the village I go.' [LS]

(11) mā yanne gamaṭa^{F} ya
I.ACC go.PRES.E village^{F}.DAT *ya*
'It is to the village I go.' [LS]

As noted by Gair (1986[1998]b: 155), the focus constructions of modern Sinhala (both colloquial and literary) "bear an unmistakable resemblance, in both form and meaning, to constructions in several Dravidian languages" (i.e. Tamil,

Malayalam, Telugu, and Kannada). Observe the structure of the focus construction in (12) below, from Sri Lankan Tamil [SLT].[4]

(12) naan poonatu yaaLppaaNattukku^{F}
I.NOM went.PAST.VN.NOM Jaffna.DAT^{F}
'It was to Jaffna that I went.' (cited from Gair 1986[1998]b: 156) [SLT]

The Tamil structure exemplified by (12) has a long history in Dravidian, as evidenced by the appearance of the same type of structure in Old Tamil [OT] (300 B.C. to A.D. 700), as shown below in (13).

(13) yāṉ oṟu-ppatu numar-ai^{F}
I.NOM punish-NONPAST-VN.NOM relation-ACC^{F}
'It is the relations whom I punish.' (*kali* 58.20; cited from Lehmann 1998: 97) [OT]

However, there are important differences between Dravidian and both modern literary and colloquial Sinhala, with respect to the status of the verb in the "clefting"/"focussing" structures described above. The Tamil "focussing" verbal forms in (12)–(13) are actually nominalised verbs, derived from an attributive participle with the addition of an affix which is morphologically identical to the default inanimate third-person singular ending (glossed as NOM in the examples shown here).[5] The same form appears also in the formation of action nominals in nominalised clauses, as in (14) below.

(14) mani pooRatu
Mani.NOM go.PRES.VN.NOM
'Mani's going' (cited from Gair 1986[1998]b: 156) [SLT]

The Tamil 'focus' construction is commonly referred to as a 'cleft sentence', see Annamalai & Steever (1998), who represent the alternation between the modern mainland Tamil Nadu Tamil [TNT] sentences in (15) in terms of a transformation.

(15) a. nāṉ maturai.y-il piṟa-nt-ēṉ
I.NOM Madurai-LOC be_born-PAST-1SG
'I was born in Madurai.'

b. [$_{S}$ [$_{NP_1}$ nāṉ piṟa-ntatu] [$_{NP_2}$ maturai(-y.il)]^{F}]
[$_{S}$ [$_{NP_1}$ I.NOM be_born-PAST-VN.NOM] [$_{NP_2}$ Madurai-NOM(-LOC)]^{F}]
'Madurai is where I was born.' (cited from Annamalai & Steever 1998: 123) [TNT]

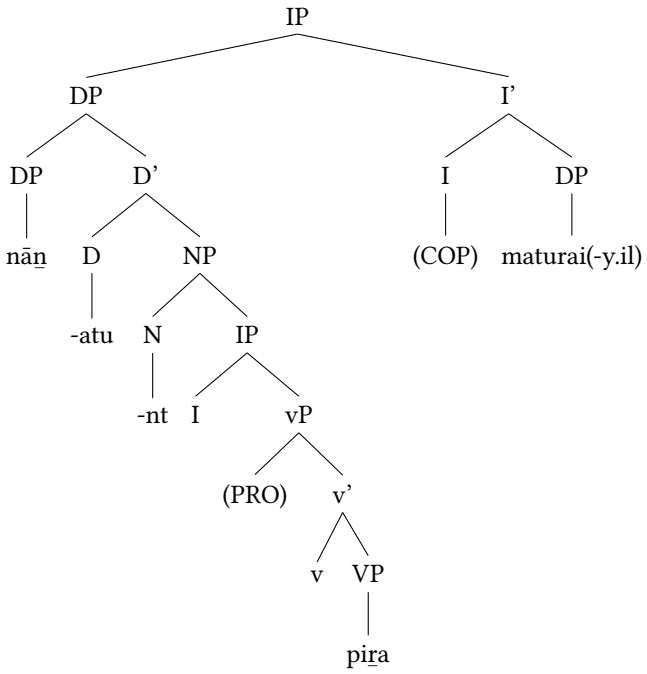
Annamalai & Steever (1998) describe the relation between (15a) and (15b) as a transformation of (15a) by the postposing of the focused element *maturai(y.il)* "(in) Madurai" to the right of the verb, with the simultaneous nominalisation of the verb, noting further that the locative case marker *-il* is optional in the clefted sentence in (15b), and is frequently substituted by the nominative.

Using more modern terminology, we might treat (15b) as a nominalisation of the verb, as in (16) (cp. Baker & Vinokurova 2009).

[4]Sri Lankan Tamil is very similar to the language of mainland Tamil Nadu; the differences between the two varieties do not appear to concern the structure of focus constructions, see Gair & Suseendirarajah 1981[1998].

[5]In Tamil, endings of this sort attach not only to verbal forms but also to nouns.

(16)



Whatever precise analysis is adopted for the (morpho)syntax of Dravidian "cleft" constructions, it is clear that they involve a type of phrasal nominalisation (see Yoon (1996)).

Note that, in contrast to Tamil (see example (14) above) where the same morphological form appears both for general nominalisation and in "cleft" constructions, in Sinhala the E-form of the verb which appears in "focus" constructions is not employed in nominalisations in either colloquial or literary varieties. Rather, in both varieties of modern Sinhala, verbal nominalisation involves either the 'gerund' form of the verb in *-iimə* or *-illə*, (17), or, more commonly, the use of the 'adjectival' form of the verb (generally followed by *ekə* if there is no overt noun; *ekə* is etymologically the inanimate numeral "one"), (18).

(17) [miniha-ge pot liviimə/livillə] hoňdə nææ
[man-GEN books write-GER] good not
'The man's writing books is not good.' (cited from Gair 1976[1998]: 207) [MCS]

(18) [[miniha pot liyənə] ekə] hoňdə nææ
[[man-NOM books write.PRES.ADJ] *ekə*] good not
'That the man writes books is not good.' (cited from *Ibid.*) [MCS]

Note that the 'gerund'-type nominalisation exemplified by (17) requires the logical subject of the nominalised clause to appear in the genitive case, unlike the 'adjectival'-type nominalisation of (18) where the logical subject appears with the expected nominative case.

With respect to the second construction, literary Sinhala differs in that the logical subject appears with accusative case, cp. the literary construction in (19a) with its colloquial equivalent in (19b).

(19) a. siri [[mā kiyevuva] potə] nokiyevuvēya
siri.NOM [[I.ACC read.PAST.ADJ] book] not-read.PAST.3SG
'Siri did not read the book that I read.' (cited from Gair 1995[1998]: 243) [LS]

b. siri [[mamə kiyewwə] potə] kiyewwe nææ
Siri.NOM [[I.NOM read.PAST.ADJ] book] read.PAST.E neg
[MCS]

Example (19a) shows that here, as in 'clefted' clause of focus constructions, the logical subject of the literary nominalised clause appears with accusative case (unlike colloquial, where the logical subject in both constructions appears with the expected case, usually nominative).

Thus, there is no evidence in either variety of modern Sinhala that the "focussing" *-e* forms of the verb represent actual nominalisations of the verb.

In Tamil, as in colloquial Sinhala, the appearance of an emphatic marker does not trigger obligatory dislocation of the focussed element, but this is the case in literary Sinhala. Compare Tamil (20), literary Sinhala (21), and colloquial Sinhala (22).[6]

(20) avar-taan naaLaykku koLumpukku pooRaar
he-EMPH tomorrow.DAT Colombo.DAT go.PRES.3SG.MASC
'He is going to Colombo tomorrow.' (cited from Gair 1986[1998]b: 161) [SLT]

(21) a. mā yanne gamaṭa^{F} ya
I.ACC go.PRES.E village^{F}.DAT *ya*
'It is to the village that I go.' [LS]

b. *mā gamaṭa^{F} ya yanne
I.ACC village^{F}.DAT *ya* go.PRES.E
'It is to the village that I go.' [LS]

(22) a. mamə yanne gamə̀ṭə^{F} (yi/tamaa/tamay)
I go.PRES.E village^{F}.DAT (EMPH)
'It is to the village I go.' [MCS]

b. mamə gamə̀ṭə^{F} (yi/tamaa/tamay) yanne
I village^{F}.DAT (EMPH) go.PRES.E
'It is to the village I go.' [MCS]

Table 4.1 summarises the properties of focus constructions in modern Tamil, modern literary Sinhala, and modern colloquial Sinhala.

	Tamil	**Literary Sinhala**	**Colloquial Sinhala**
(1)	optional 'clefting' in focus constructions	obligatory 'clefting' in focus constructions	optional 'clefting' in focus constructions
(2)	non-obligatory element *-taan*	obligatory element *ya*	non-obligatory element *-y(i), tamaa, tamay*
(3)	subj. of 'cleft' clause in normal case	subj. of 'cleft' clause in acc. case	subj. of 'cleft' clause in normal case
(4)	same morphology for verb of 'cleft' and verbal nouns	distinct morphology for verb of 'cleft' and verbal nouns	distinct morphology for verb of 'cleft' and verbal nouns

Table 4.1: Properties of focus constructions in Tamil and Sinhala

[6]Here Malayalam more closely resembles literary Sinhala than does Tamil, as Malayalam requires both that focussed elements follow the verb and that they are followed by the copula *aaNū* (unless the emphatic marker *tanne* already occurs on the focus), cf. Gair (1986[1998]b: 161–162).

Tamil and colloquial Sinhala agree on points (1), (2), and (3), against literary Sinhala; and literary and colloquial Sinhala agree on point (4), against Tamil. This might be taken as evidence for the progressive convergence of the syntax of Sinhala focus constructions with that of Tamil. However, it is important to note that both literary and colloquial varities of Sinhala differ from Tamil with respect to point (4). This point is important, since it raises the possibility that *-e* verbal forms may no longer synchronically be treated as nominalisations—though their antecedents in Classical and Old Sinhala clearly are—in sharp contrast to the situation in Tamil (see further Chapter 11.1).

I shall argue in Chapter 11.1 that in fact, though on the surface the colloquial Sinhala focus construction appears to have progressively undergone more and more convergence with the focus construction of Tamil, the colloquial Sinhala 'focus' E-verbal form has been reanalysed as a special form of the finite verb, as opposed to a nominalisation of the verb, and that therefore the colloquial Sinhala focus construction has diverged signficantly from the structure of the Dravidian construction.

One pair of related questions which should be addressed at this point is: (a) what is the status of the Tamil 'emphatic' particle *-taan*, and (b) what is the status of 'emphatic' particles of both literary Sinhala (*ya*) and colloquial Sinhala (*(y)i)*, as well as *tamaa, tamay*)?

Tamil *-taan* marks emphasis, and apparently derives from the reflexive use of *taan*; literary Sinhala *ya* and colloquial *yi* are historically identical, deriving from a clitic form of the copula. Literary Sinhala *ya* retains this identity as a third-person singular clitic form of the copula (or perhaps just a realisation of AGR, see Gair 1995[1998], who calls this an 'agreement-clitic'); while colloquial Sinhala *yi* has become an (optional) marker of emphasis/focus. That literary Sinhala *ya* is a copula or overt agreement-clitic can be seen by the fact it appears with this function in simple equational sentences, as in (23a) below; and by the fact that it is part of a larger paradigm of agreement-clitics, as shown by the appearance of the clitic *mi* (first-person singular agreement) in (23b).[7]

(23) a. hetema goviyek ya / hetema goviyek-i
he.NOM farmer.INDEF 3SG / he.NOM famer.INDEF-3SG
'He is a farmer.' [LS]

b. mama goviyek mi
I.NOM farmer.INDEF 1SG
'I am a farmer.' [LS]

Note that while in literary Sinhala all sentences require overt agreement of some sort, colloquial Sinhala employs no overt copula in this context, as shown below in (24).[8]

(24) a. mamə goviyek
I.NOM farmer.INDEF
'I am a farmer.' [MCS]

b. eyaa goviyek
he.NOM farmer.INDEF
'He is a farmer.' [MCS]

Further, in literary Sinhala focus constructions, *ya* (25a) may be substituted by a lexical copula (25b), but one or other of these forms must occur, as shown by the ungrammatical (25c).[9]

[7]The *-i* in (23) is a reduced form of the "agreement marker" *ya*.
[8]Modern colloquial Sinhala, in fact, has no copula.
[9]Cf. Gair (1995[1998]: 255).

(25) a. mā kiyavannē [ema potə]^{F} ya
I.ACC read.PRES.E [that book]^{F} 3SG
'It was that book that I read.' [LS]

b. mā kiyavannē [ema potə]^{F} veyi
I.ACC read.PRES.E [that book]^{F} be.3SG
'It was that book that I read.' [LS]

c. *mā kiyavannē [ema potə]^{F}
I.ACC read.PRES.E [that book]^{F}
'It was that book that I read.' [LS]

On the other hand, in colloquial Sinhala, we find that the particle *y(i)* is optional, and further may be substituted—not by a copula—but rather by other emphatic particles like *tamaa, tamay*, as in above example (22).

4.2 The syntax of focus constructions in modern colloquial Sinhala

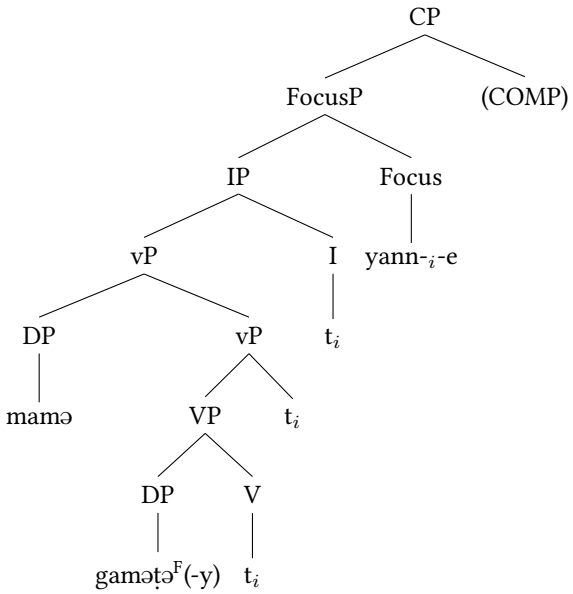
I argue that the syntax of focus constructions in Modern Colloquial Sinhala therefore can be given a monoclausal analysis.

I propose that the *-e* of focussing verbs is an affix which is generated in the head of FocusP.[10] The main verb (which has raised from V to v to I) then raises from I to the head of FocusP and picks up the *-e* affix.[11] Thus the sentences (26a) and (26b) would have the structures shown in (27) and (28), respectively.[12]

(26) a. mamə gaməṭə^{F}(-y) yanne
I village^{F}.DAT(-EMPH) go.PRES.E
'It is to the village I go.'

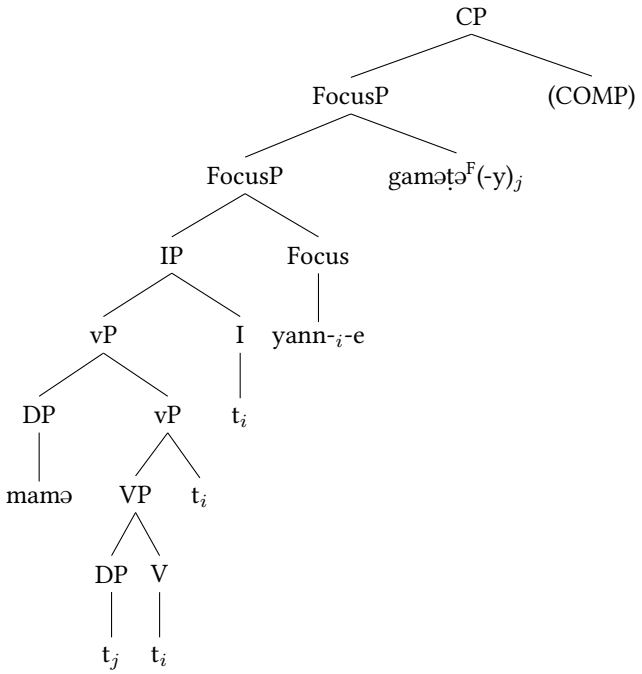
b. mamə yanne gaməṭə^{F}(-y)
I go.PRES.E village^{F}.DAT(-EMPH)

(27)



[10] Since the *-e* marking obviously is associated with focus, it is natural to assume that it is generated in the head of FocusP.
[11] I posit that FocusP is left-headed, given that focussed elements appear to the right of the *e*-marked verb.
[12] Presumably, the focussed element in (26a) (with the structure shown in (27)) would raise to SpecFocusP at LF.

(28)



Given this structure, I suggest that the following rules successfully predict the behaviour of focus constructions in Sinhala:

(29) Syntactic licensing of focus

a. A focus-marked element must be in the c-command domain of *-e* (which occupies the head of FocusP).
b. The focus-associated element *-e* must have a focus-marked element in its c-command domain.

We can formalise the descriptive rules in (29) as follows. Let us assume, using the system of syntactic features discussed above in 3.1, that focussed elements enter the derivation with a valued iFocus[+] feature and an unvalued feature uExist[ ]. The morpheme *-e*, residing in the head of FocusP, bears a unvalued uFocus[ ] feature, a valued iExist[+] feature, and an optional Edge feature (which triggers overt movement of the Agreeing element to SpecFocusP).[13] If the head of FocusP bears an Edge feature, the focus element is moved overtly to SpecFocus (appearing, on the surface, to the right of the *-e* marked verb), and is further frozen in place (unable to undergo further movement).[14] This formal analysis correctly captures the characterisation in (29).

4.3 Further arguments for a monoclausal analysis of focus constructions in modern colloquial Sinhala

Gair & Sumangala (1991) and Kariyakarawana (1998)—Gair's 1991 co-author, having in the meantime changed his name—argue that Sinhala "cleft" constructions involve overt movement of the focussed constituent to the right of the verb represent biclausal constructions. I argue, rather, that in modern colloquial Sinhala focus constructions no longer involve clefting (or any sort of nominalisation+biclausal structure) and thus should be analysed as monoclausal—though as discussed in Chapter 11.1, there is good evidence that focus did involve clefting (i.e. biclausal structures) in earlier stages of Sinhala.

[13] The naming of the feature Exist reflects the association of the *-e* morpheme with an existence presupposition.
[14] On freezing, see Rizzi (2006, 2007).

I begin by examining the evidence for a cleft analysis of focus in Sinhala. Consider again the data in (4), repeated below as (30).

(30) a. mamə gamə̣ə^{F} yanne
I village^{F}.DAT go.PRES.E
'It is to the village I go.' [MCS]

b. mamə gamə̣ə^{F}-y yanne
I village^{F}.DAT-EMPH go.PRES.E
'It is to the village I go.' [MCS]

c. mamə yanne gamə̣ə^{F}
I go.PRES.E village^{F}.DAT
'It is to the village I go.' [MCS]

d. mamə yanne gamə̣ə^{F}-y
I go.PRES.E village^{F}.DAT-EMPH

Kariyakarawana (1998) distinguishes between two different type of focus, what he calls C-focus (that is, "cleft"-focus) and E-focus (that is, "emphasis"-focus). C-focus involves the focussed constituent appearing in a position to the right of the verb, as in (30c), while in sentences with E-focus, the focussed constituent does not occupy a surface position to the right of the verb. Kariyakarawana (1998) argues for a biclausal analysis of C-focus, and suggests that while E-focus involves a monoclausal structure, there is still movement of an operator (associated with the focus element) to high structural position.

Here I examine Kariyakarawana's evidence for distinguishing two different types of focus constructions in Sinhala, and argue that all of his data can be accounted for by the analysis proposed above.

4.3.1 Position of Sentential Complementisers

Kariyakarawana (1998) notes that the possible positions of "clefted" elements with respect to sentential complementisers differ from the positions allowed in neutral sentences. The examples in (31) show that elements may be scrambled to the right of sentential complementisers.[15]

(31) a. gunē koləmbə yanəva kiyəla
Gunee Colombo go.PRES.A that
'. . . that Gunee goes to Colombo. . . '

b. koləmbə yanəva kiyəla gunē
Colombo go.PRES.A that gunee
'. . . that Gunee goes to Colombo. . . '

However, where the sentence contains a focussed element, that element may not appear to the right of sentential complementisers, as shown by (32b) (contrast with (32a)).

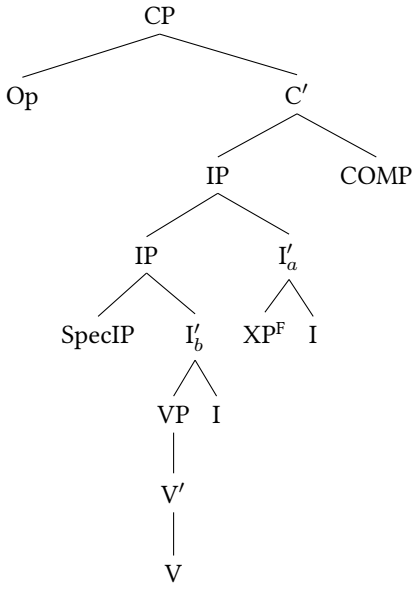
(32) a. koləmbə yanne gunē^{F} kiyəla
Colombo go.PRES.E Gunee^{F} that
'. . . that GUNEE that goes to Colombo. . . '

b. *koləmbə yanne kiyəla gunē^{F}
Colombo go.PRES.E that gunee^{F}
'. . . that GUNEE that goes to Colombo. . . '

[15] Any of *nan* "if", *lu* "they say" (reportative particle), or *vennə æti* "may be" may be substituted for *kiyəla* in (31) and (32) without any difference in the grammaticality/ungrammaticality.

Kariyakarawana (1998) argues that what he calls the "presupposition clause" (the clause the clefted element, XP^F, occupies, assuming a biclausal structure), I'_a in tree (33), is a bare IP and thus has no position for COMP elements, as shown in (33).

(33)



However, we can more straightforwardly account for the ungrammaticality of (32b) simply by applying the rules in (29): if *gunē* moves to the right of the complementisers then it is no longer within the scope of the -E marker. In more formal terms, the focussed element may not move beyond SpecFocusP due to freezing (which takes place once the focussed element's unvalued feature acquires a value), see above.

Kariyakarawana (1998) also examines distributional differences in the possible position of adverbs in neutral and focussing sentences. He notes that both VP adverbs and sentential adverbs may appear following the verb in neutral sentences:

(34) a. īye amma kæum hæduva
yesterday mother oil-cake make.PAST.A
'Yesterday, mother made oil-cake.'

b. amma kæum hæduva īye
mother oil-cake make.PAST.A yesterday

c. amma kæum ikmənəṭə hæduva.
mother oil-cake quickly make.PAST.A
'Mother made oil-cake quickly.'

d. amma kæum hæduva ikmənəṭə.
mother oil-cake make.PAST.E quickly

However, in focussed sentences where the focussed element appears to the right of the verb, only sentential adverbs may follow the verb, as shown by (35) below.

(35) a. īye amma hæduve kæum^{F}
yesterday mother make.PAST.E oil-cake^{F}
'It was OIL-CAKE that mother made yesterday.'

b. amma hæduve kæum^{F} īye
mother make.PAST.E oil-cake^{F} yesterday

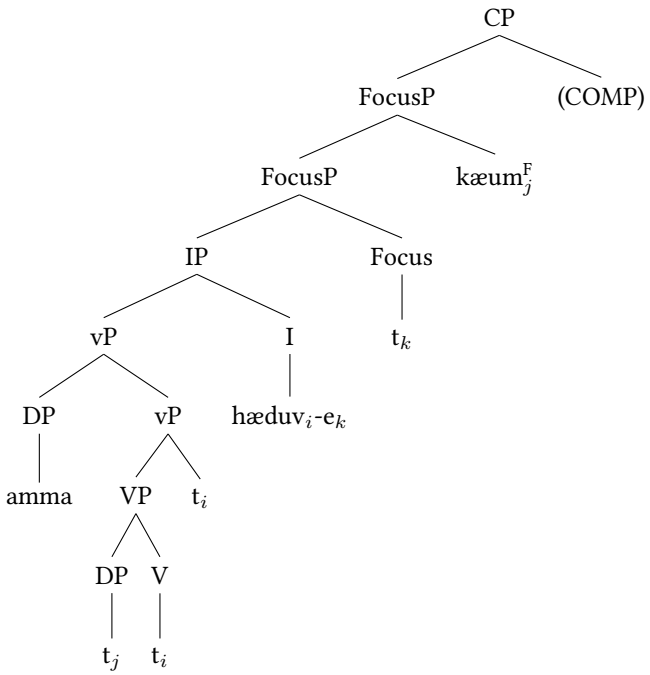
c. amma ikmənəṭə hæduve kæum^{F}.
mother quickly make.PAST.E oil-cake^{F}
'It was OIL-CAKE that mother made quickly.'

d. *amma hæduve kæum^{F} ikmənəṭə.
mother make.PAST.E oil-cake^{F} quickly

Kariyakarawana (1998) argues that the ungrammaticality of (35d) shows that such "cleft" sentences involve a biclausal structure, otherwise the VP adverb should be able to scramble to the right of verb, as in (34d).

Yet we can capture the same facts assuming the structure for Sinhala focussed sentences suggested above in Section 4.2. Given that -E occupies the head of FocusP, a VP-adverb will not be able to appear to the right of -E-type verb, as shown by (36).[16]

(36)



4.3.2 Scope of Negation

The examples in (39)–(41) below are the data Kariyakarawana (1998) uses to argue for the scope of negation as providing evidence for a biclausal analysis of "cleft" sentences in modern colloquial Sinhala. He glosses both *nevey* and *nætte* as "not", and does not offer any further remarks about differences between them. So, before discussing these examples in the context of deciding between syntactic analyses of "cleft" sentences, I provide a number of observations about the properties of these two negation elements.

Thus, I first examine various negative constructions in Sinhala, and their associated syntactic, semantic, and pragmatic properties. There are a number of negators in Sinhala, including a prefixal *n(o)-*, alongside non-affixing

[16]Note that this does require us to allow sentential adverbs to be able to adjoin to FocusP or some higher functional projection.

nevey[17] and *nææ*.

Nææ may serve as the negative counterpart of existential verbs like *tiyenəwa* "be, exist (inanimate)", *innəwa* "be, exist, stay (animate)", e.g.:

(37) a. mehē alut pot tiyenəwa.
here new books be
'There are new books here.'

b. mehē alut pot nææ.
here new books *nææ*
'There aren't new books here.'

Further, *nææ* is the basic negation found in neutral sentences like (38).

(38) gunē koləmbə giyē nææ.
Gunee Colombo go.PAST."E"$_1$ *nææ*
'Gunee did not go to Colombo.'

Note that in (38) the verb, *giyē*, appears to take the E-form associated with focussing sentences. Historically, the E-form in (38) is doubtless identical with the focus-associated E-form. However, synchronically, the E-form in (38) has none of the properties associated with the focussing E-form: it does not require or allow a focussed element (i.e. an element that can be followed by *tamay*, *-y* etc.) in its scope, and consequently it does not carry any focus-related presupposition (see Chapter 5). I thus gloss this "pseudo E-form" as "E"—with a subscript 1, as we shall encounter another, morphologically distinct, "pseudo E-form".

Nevey, in contrast to *nææ*, is a specifically focus-associated constituent negator. In (39) it associates with the entire IP; in (40) it associates specifically with the DP *Colombo*.

(39) [[gunē koləmbə giyā] nevey]
[[Gunee Colombo go.PAST.A] *nevey*]
'It is not that Gunee went go to Colombo.'

(40) [gunē giyē [[koləmbə^F] nevey]]
[Gunee go.PAST.E [[Colombo^F] *nevey*]]
'It is not to COLOMBO that Gunee went.'

Thus, in (39) the entire proposition is negated, with no presupposition that Gunee or anyone else went anywhere; while in (40) there is a presupposition that Gunee did go somewhere, and the sentence asserts that Colombo is not the place where he went.

In example (41), on the other hand, the sentence in fact asserts that there is place that Gunee did not go, and that that place is Colombo.

(41) gunē koləmbə^F(-y) giyē nætte
gunee Colombo^F(-EMPH) go.ADJ."E"$_2$ *næti*.E
'The place that Gunee didn't go is Colombo.'

[17] With dialectal variants *nevi*, *nemey*, and *nemē*.

As in (38), here too in (41) we find a "pseudo E-form" of the verb, *giyē* (alongside of *nætte*, which is true focussing E-form). *Nætte* is the focussing E-form of the verbal negation *næti*.[18] *Næti*/*nætte* combine with adjectival participles (which normally end in *-ə*), but always require that the participle appear with an *-e* ending (see Reynolds 1980: 162).[19] This is, like the *-e* ending triggered by *nææ*, a sort of "pseudo E-form" of the verb, as it has none of the properties typical of the focussing E-form. Since, unlike *giyē* in (38) (which is finite), the *giyē* we find in (41) is adjectival, I label it as "E"$_2$.

Kariyakarawana (1998: 102–6) suggests that the differences between (39)–(41) have to do with the scope of the negation; he further asserts that we should expect the negative elements in (40) and (40) to negate the entire proposition if these were monoclausal.

However, in fact, (38)–(40) all have the same truth conditions, i.e. they all assert that Gunee did not go to Colombo; where they differ is with respect to their presuppositions (which reflect how negation can interact with focus rather than differences in scope). Only (41) has different truth conditions, because it actually represents a different syntactic construction (which is obscured by the fact that *næti*/*nætte* combines with adjectival participles—which otherwise end in *-ə*—and by a, synchronically construction-specific, rule which triggers a morphophonological change in their ending to *-e*, resulting in what looks like a focussing E-form of the verb).

The important point here is that, since Sinhala possesses a constituent negator, as discussed in this section (see Chapter 5.4 for a semantic analysis of this element), the question of monoclausal vs. biclausal constructions is orthogonal to the differences in semantics and pragmatics in (39)–(41), which derive from (i) the structural position of the constituent negator *nevey* (in the case of (39), (40)) or (ii) the use of a adjective participle (in (41))—which takes a "pseudo E-form" as discussed above—rather than a "focussing" form of the verb with a ("true") *e*-morpheme.

4.3.3 Apparent uniqueness of focus

One potential problem with the monoclausal analysis of modern colloquial Sinhala focus constructions as non-'clefts' is the fact that only one instance of *y(i)/tamaa/tamay* may occur per clause, compare the grammatical (42a), (42b) to ungrammatical (42c). This, I believe, is the strongest argument of Kariyakarawana (1998) for a biclausal analysis of Sinhala "cleft" constructions.

(42) a. gunapala^{F}-y ee potə kieuwe
Gunapala^{F}-EMPH that book read.PAST.E
'It is GUNAPALA who read that book.' [MCS]

b. gunapala ee potə^{F}-y kieuwe
Gunapala that book^{F}-EMPH read.PAST.E
'It is THAT BOOK that Gunapala read.' [MCS]

c. *gunapala^{F}-y ee potə^{F}-y kieuwe
Gunapala^{F}-EMPH that book^{F}-EMPH read.PAST.E
'It is THAT BOOK that GUNAPALA read.' [MCS]

If we were to treat focus constructions as true (South Asian) 'clefts', then this restriction might be more naturally explained. In English, for instance, only one cleft is allowed per clause, compare grammatical (43a), (43b) to ungrammatical (43c), (43d).

[18] *Næti*/*nætte* are morphologically the negative counterparts of *æti*/*ætte*. However, in terms of their synchronic semantics, they seem to have little connection: *æti*/*ætte* mean "is enough" or "probably is" (Reynolds 1980: 201). Historically *æti*/*ætte* are cognate with Sanskrit *asti*, Pāli *atthi* "be (3sg)" (and likewise *næti*/*nætte* are cognate with Sanskrit *nāsti* [that is, *na* "not" + *asti*], Pāli *natthi*).

[19] An alternative method of forming negative participles is to prefix *no-* to the participle, see Reynolds (1980: 162).

(43) a. It is John^{F} that Mary saw.
b. It is Mary^{F} that saw John.
c. *It is John^{F} that it is Mary^{F} that saw.
d. *It is Mary^{F} that it is John^{F} who saw.

However, a constraint on the occurrence of multiple focus-marking particles in a single clause is not restricted to Sinhala. In Hindi, the focus particle *hī* may only occur once per clause, see the examples below in (44) (cited from Sharma 1999).

(44) a. uske-**hī** jūte mere kamre-meṁ paḍe the
his-EMPH shoes my room-in lie.PAST.PTCP.MASC.PL be.PAST.MASC.PL
'His shoes were lying in my room.'
b. uske jūte mere-**hī** kamre-meṁ paḍe the
his shoes my-EMPH room-in lie.PAST.PTCP.MASC.PL be.PAST.MASC.PL
'His shoes were lying in MY room.'
c. *uske-**hī** jūte mere-**hī** kamre-meṁ paḍe the
his-EMPH shoes my-EMPH room-in lie.PAST.PTCP.MASC.PL be.PAST.MASC.PL
'His shoes were lying in MY room.' [Hindi]

There is no evidence, either synchronic or historical, which would suggest that Hindi sentences containing *hī* are 'clefts'. Rather there appears to be a constraint which, both in modern colloquial Sinhala and in Hindi, restricts the occurrence of focus-marking particles to one per clause.

Further, observe that there is no restriction on the number of focussed elements per clause in Sinhala, only on the number of occurrences of focus-marking particles, as illustrated by the dialogue below in (45).[20]

(45) a. Person A: itin, chitra ranjiṭṭa kiyuwa ...
so, Chitra.NOM Ranjit.DAT tell.PAST.A ...
'So, Chitra told Ranjit that ...'
b. Person B [interrupting]: nehe, gunapala^{F}-y rohanaṭa^{F} kiwwe
no, Gunapala.NOM^{F}-EMPH Rohana.DAT^{F} tell.PAST.E
'No, it was GUNAPALA who told ROHANA.' [MCS]

Thus the constraint against multiple occurrences of *y(i)/tamaa/tamay* in a single clause should be understood to be a morphosyntactic restriction (perhaps on the licensing of focus-marking particles), as shown by comparison to Hindi, rather than evidence that the modern colloquial construction remains a 'cleft'.

4.4 Summary

The structural analysis and rules proposed in Section 4.2 can thus account for the distribution of E-verb forms and focussed elements in modern colloquial Sinhala. Therefore, we can maintain a monoclausal analysis of focus constructions in modern colloquial Sinhala, which can account for both what Kariyakarawana (1998) calls C-focus and E-focus, assuming that in-situ focussed elements raise to SpecFocusP at LF.

As discussed in Section 4.1, the *-e* "focussing" verbal forms of both modern varieties of Sinhala are no longer directly connected with nominalisation (nominalisation exists in both forms of Sinhala, but is accomplished through

[20] The context for this dialogue: Person A and Person B are telling Person C about a past conversation of some mutual friends.

other morphological means).[21] Though in earlier forms of Sinhala, specifically Classical Sinhala and modern literary Sinhala, focus constructions are best analysed as involving a biclausal "cleft" (see Chapter 11.1), in modern colloquial Sinhala a biclausal analysis is not well motivated.

Firstly, subjects of *-e* "focussing" verbs retain their expected case (nominative, except where the verb assigns dative case) and neither "focussing" *-e* verbs nor "neutral" *-a* verbs bear agreement morphology—thus in terms of morphology *-e* verbs do not behave differently from *-a* verbs. Secondly, modern colloquial Sinhala focus constructions contain no copular element and displacement of the focussed element is optional. Rather than positing two different structures, one, monoclausal, where the focussed element remains *in situ*, the other, biclausal, in case of overt movement of the focussed element; a single, monoclausal, analysis can account for both structures. In both structures the focussed element can be analysed as moving to SpecFocusP, either overtly, or else covertly (at LF). Restrictions on the possible positions of overtly moved focussed elements can be accounted for if we assume that once the focussed element moves to SpecFocusP, a feature-driven movement, it cannot undergo further movement (due to "freezing", as discussed above in Section 4.1).

The historical predecessors to the modern Sinhala focus construction, however, involve rather different structures, as discussed in Chapter 11.1; one of these, that of Classical and modern literary Sinhala, is a true "clefting" construction, in the sense that it involves a biclausal structure, similar to that of Dravidian languages like Tamil and Malayalam.

Having argued for a monoclausal analysis for modern colloquial Sinhala focus constructions, in which the focussed element moves (either overtly or covertly) to SpecFocusP, in the next chapter I provide a formal semantic analysis for such structures.

[21] Again, historically the *-e* verbal form derives from a nominalised form of the verb, as discussed in Chapter 11.1, but is not synchronically connected with nominalisation in the modern forms of the language.

Chapter 5

Semantics of Sinhala focus constructions

5.1 A Roothian semantics of focus

In my analysis of the semantics of modern colloquial Sinhala focus constructions, I adopt the system of alternative semantics Rooth (1985, 1992, 1996) develops for the treatment of focus. Rooth proposes that every expression ϕ has two semantic values. In addition to its ordinary semantic value, $[\![\phi]\!]^{g}$, ϕ also has a focus semantic value, $[\![\phi]\!]^{g,F}$, which is a set of semantic objects identical to $[\![\phi]\!]^{g}$ except that they represent 'alternatives' to any focussed elements in ϕ. The focus semantic value of a focussed element is generated as per (1). Examples are given in (2).

(1) $[\![X_F]\!]^{g,F} = \{x \in D_\tau \mid [\![X]\!]^{g} \in D_\tau\}$

(2) a. $[\![\text{John}]\!]^{g,F} = \{x \in D_e\}$
b. $[\![\text{come}]\!]^{g,F} = \{P \in D_{\langle e,t\rangle}\}$

Thus the focus semantic value of an expression like $[\![\text{John saw [Bill]}_F]\!]^{g,F}$ (=(4b)) would be as in (3), where the focussed element, *John*, is replaced by various 'alternatives'.

(3) a. $\{\lambda w.\text{John saw x in w} \mid x \in D_e\}$, e.g. more informally,
b. {John saw Mary, John saw Kim, John saw George, John saw Bill, . . . }

Observe that *wh*-words and focus both involve alternative semantic values; as pointed out by Rooth (1992), there is a correlation between *wh*-questions and the position of focus in answers. Thus, (4b) is an appropriate reply to (4a), whereas (4c) sounds distinctly odd as a reply to (4a).[1]

(4) a. Who did John see?
b. John saw Bill.
c. John saw Bill.

Adopting a semantics which treats questions as sets of propositions (see Chapter 6), (4a) would be represented as in (5) (for the derivation, see (3b) above).

(5) $\{\lambda w.\text{John saw x in w} \mid x \in D_e \wedge x \in \text{human}' \text{ in w}\}$

Note the similarity between (5) and (3); specifically, (5) is a subset of (3). If we assume that the ordinary semantic value of a question must be a subset of the focus semantic value of a corresponding answer (see Rooth (1992)), then it makes sense why (4b) is an appropriate answer to (4a) whereas (4c) is not.[2]

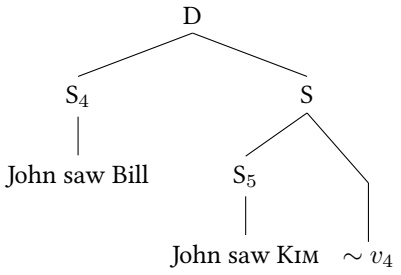
[1] I use SMALLCAPS to indicate focus.
[2] The focus semantic value (3a) is a superset of the ordinary denotation of the question (5) since the former includes x's which are not people.

Another case in which focus semantic values are necessary is in contexts like that shown in (6), a typical configuration in which focus indicates contrast.

(6) A: John saw Bill.
B: No, John saw KIM.

On the alternative semantics analysis developed in Rooth 1985, 1992, the dialogue in (6) can be given the (simplified) representation in (7).

(7)



The $\sim$ operator invokes the focus-semantic value of a constituent, as in (8), where v_4 refers back to a previous utterance in the discourse. That is, for (7) to be felicitious, the following must hold:

(8) $[S_5 \sim v_4]$ is felicitous if $[\![v_4]\!]^g \in [\![S_5]\!]^{g,F}$

For the dialogue in (6) to be felicitous, v_4 must be a member of the focus semantic value of S_4. The calculation of the relevant ordinary and focus semantic values, as shown in (9), reveals that (7) does indeed meet the felicity condition in (8).

(9) a. $[\![\text{John saw } [\text{Kim}]_F]\!]^g = \lambda w.\text{John saw Kim in w}$
b. $[\![\text{John saw Bill}]\!]^g = \lambda w.\text{John saw Bill in w}$
c. $[\![\text{John saw } [\text{Kim}]_F]\!]^{g,F} = \{\lambda w.\text{John saw x in w} \mid x \in D_e\}$

The ordinary semantic value of S_4 (=v_4) is a member of the focus semantic value of S_5, and thus the dialogue is correctly predicted to be felicitous. More generally, we may define $\sim$ as in (10), where C is a contextually-determined variable:

(10) $[\![\sim C\ \alpha]\!]^g = [\![\alpha]\!]^g$; iff $g(C) \subseteq [\![\alpha]\!]^{g,F}$, undefined otherwise

5.2 The semantics of modern Sinhala focus constructions

In this section I consider the semantics of modern colloquial Sinhala focus constructions in detail. As indicated by the translations given, the Sinhala focus construction is similar in certain respects to an English cleft construction. That is, semantically, Sinhala focus constructions involve more than the mere presence of a focussed element.

Rooth (1999) argues that the semantics of focus are distinct from the semantics of existential presupposition. In support of this position he points out that focus is felicitous in contexts where clefting is not.

Consider the dialogue in (11).

(11) Person 1 asks: "Did someone hit John?"

a. Person 2 replies: "I don't know, but MARY (certainly) didn't hit John."

b. Person 2 replies: #"I don't know, but it (certainly) wasn't MARY who hit John."

While the reply in (11a), with focus on "Mary", is possible, the reply in (11b), where "Mary" is clefted, is infelicitous. Clefts in English, it seems, involve an existential presupposition, which may project. Thus in (11b), the cleft invokes a presupposition of the form [$\exists x \in D_e$.hit(john,x)], which conflicts with the first thing that Person 2 said, namely "I don't know [if someone hit John]".

How do Sinhala focus constructions behave? Consider the question in (12), and the answers in (13).

(12) kavuru hari rohanəṭə gehuva də?
who *hari* Rohan.DAT hit.PAST.A *də*?
'Did someone hit Rohan?'

(13) a. mamə danne nææ, ēt sita rohanəṭə gehuve nææ kiyəla nam mamə dannəwā.
I know.PRES."E_1" not, but Sita Rohan.DAT hit.PAST."E_1" not that EMPH I know.PRES.A
'I don't know, but I know that Sita didn't hit Rohan.'

b. #mamə danne nææ, ēt sita^F nevey rohanəṭə gehuve kiyəla mamə dannəwā.
I know.PRES."E_1" not, but Sita^F *nevey* Rohan.DAT hit.PAST.E that I know.PRES.A
'I don't know, but I know that SITA didn't hit Rohan.'

As the data in (13) show, the Sinhala focussing construction behaves like an English cleft construction. That is, it seems to invoke an existence presupposition. Thus (13b) is infelicitous since the existence presupposition ($\exists x \in D_e$.hit(rohan,x)) conflicts with the first part of the speaker's utterance ("I don't know [if someone hit Rohan]").

Presumably, this property of the Sinhala focussing construction is a relic of its early status as a true "clefting" construction. The question remains: how to account for it in the synchronic grammar (given that I argue that modern colloquial Sinhala focus constructions are monoclausal, see above Chapter 4)?

I posit that the morpheme *-e* has the following semantics:[3]

(14) $[\![\text{-}e]\!]^g = \lambda P.\lambda q{:}q \in P \;\&\; \exists p \in P[p(w)=1].q$

Given the definition in (14) we can correctly derive the semantics of a sentence like (15), as shown in (16).

(15) mamə^F(-y) rohanəṭə gehuve.
I^F(-EMPH) Rohan.DAT hit.PAST.E
'It was I who hit Rohan.'

(16) $[\![[_{\text{FocusP}}\ [_{\text{IP}} \sim C\ [\text{mamə(-y) rohanəṭə gehuv-}]\ \text{-e C }]]\!]^g =$

a. $[\![\text{-e}]\!]^g(g(C))(\sim C\ [\text{mamə(-y) rohanəṭə gehuv-}]) =$

b. $\lambda P \lambda q{:}q \in P \;\&\; \exists p \in P[p(w)=1].q(g(C))([\lambda w.\text{I hit Rohan in w}]) =$

c. $\lambda q{:}q \in g(C) \;\&\; \exists p \in g(C)[p(w)=1].q([\lambda w.\text{I hit Rohan in w}])$

The value g assigns to the C must be a subset of the focus semantic value of [I^F hit Rohan], as shown in (17). The focus semantic value of [I^F hit Rohan] will be a set of alternatives to "I" hitting Rohan, e.g. "Mary hit Rohan, John hit Rohan" etc., as shown more formally in (18).

[3]Following the notation in Heim & Kratzer (1998), the expression to the right of ':' (and before the '.') is a definedness condition (in this case, a presupposition).

(17) $g(C) \subseteq [\![I^F \text{ hit Rohan}]\!]^{g,F}$

(18) $[\![I^F \text{ hit Rohan}]\!]^{g,F} = \{\lambda w.x \text{ hit Rohan in } w \mid x \in D_e\}$

Let us assume in this instance that g(C) is identical to $[\![I^F \text{ hit Rohan}]\!]^{g,F}$, in which case the first part of the definedness condition will be satisfied, since [λw.I hit Rohan in w] is a member of {λw.x hit Rohan in w | x ∈ D_e}.

(19) $[\![(15)]\!]^g$ = λw.I hit Rohan in w
iff ∃q[q ∈ {λw.x hit Rohan in w | x ∈ D_e}].q(w)=1

(19) entails that (15) is felicitous so long it is true that someone hit Rohan:– which is the observed felicity condition for a sentence like (15), as noted above.

5.3 Q-particles and focus

If we assume that the Q-particle *də* can bear focus, the appearance of the *-e* verbal form in *wh*-questions, alternative, and certain yes/no-questions follows naturally.

Thus, just as in a declarative sentence containing a focussed element, the focussed element may appear either *in situ* or else overtly moved to the postverbal focus position, as shown in (20), so too the *wh*-word plus its following Q-particle may appear either *in situ* or else in the postverbal focus position, see (21).

(20) a. Sanath ee potə^F(-y) kieuwe.
Sanath that book^F(-EMPH) read.PAST.E
"It was that book which Sanath read."
b. Sanath kieuwe ee potə^F(-y).
Sanath read.PAST.E that book^F(-EMPH)

(21) a. Sanath monəwa **də**^{F} kieuwe?
Sanath what ***də***^{F} read.PAST.E
"What did Sunil read?"
b. Sanath kieuwe monəwa **də**^{F}?
Sanath read.PAST.E what ***də***^{F}

The analysis of examples like (21b) is straightforward if we assume that the Q-particle *də* here bears a focus feature.

5.4 The semantic of constituent negation

Another element of Sinhala grammar which involves the calculation of focus semantic values is constituent negation. Consider a sentence like (22), where we see the use of the constituent negation *nevey*.

(22) koləmbə giyē gunē^F nevey.
Colombo.DAT go.PAST.E Gune *nevey*.
"It is not Gune who went to Colombo."

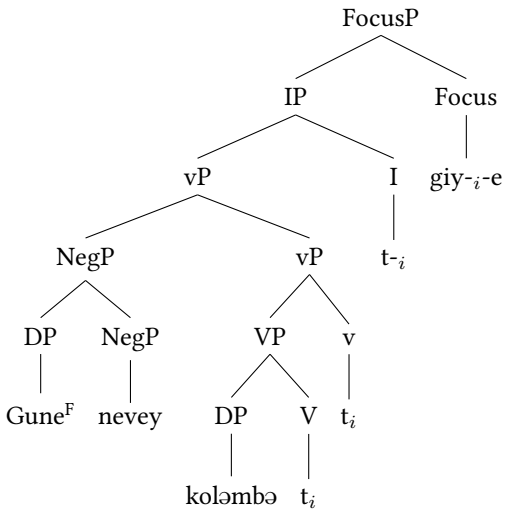
Example (22) asserts that Gune did not go to Colombo and presupposes that some person did in fact go to Colombo.

I propose the following denotation for constituent negation:

(23) $[\![\text{NEG}_{\text{CONSTITUENT}}]\!]^g$ = λPλPλQ:∃y ∈ P[Q(y)=1].∀y∈P[Q(y)=1 → [P(λx.y≠x)]]

Assuming that (22) has the structure of (24),[4]

(24)



we can derive the desired semantics for (22) as follows.

The focus semantic value of **Gune** is given in (25); according to the felicity condition in (26), the value of the pragmatically-determined variable C_2 must be a subset of (25). For ease of exposition, let us assume that $g(C_2)$ is the set {Gune, Ranjit}, as shown in (27).

(25) $[\![\text{Gune}^F]\!]^{g,F} = \{x \in D_e\}$

(26) $g(C_2) \subseteq \{x \in D_e\}$

(27) $g(C_2) = \{\text{Gune, Ranjit}\}$

The derivation of the ordinary semantic value of vP is given below in (28). I here assume that the selectional requirements of the negation induce type-shifting of the nominal *Gune*, which is "lifted" from **Gune** to λP[P(**Gune**)].[5]

[4] As discussed in Chapter 4, I assume that the *-e* base-generated in the head of FocusP and the verb raises from V to v to I to Focus; or else there is a late (morpho)phonological rule which brings about a restructuring such that *-e* is concatenated with the verb.

[5] On type-shifting, see Partee (1986).

(28) [nevey C_2 [$\sim C_2$ Gune^F]] koləmbə giy- =

a. $[\![\text{NEG}_{\text{CONSTITUENT}}]\!]^g(g(C_2))([\![\text{Gune}]\!]^g)([\![\text{[koləmbə giy-]}]\!]^g)$ =

b. $[\![\text{NEG}_{\text{CONSTITUENT}}]\!]^g(g(C_2))([\![\text{Gune}]\!]^g)(\lambda x.x \text{ went to Colombo})$ =

c. $[\![\text{NEG}_{\text{CONSTITUENT}}]\!]^g(g(C_2))(\lambda S.S(\text{Gune}))(\lambda x.x \text{ went to Colombo})$=

d. $\lambda P \lambda P \lambda Q{:}\exists z \in P[Q(z)=1].\forall y \in P[Q(y)=1 \rightarrow [P(\lambda x.y \neq x)]](g(C_2))(\lambda S.S(\text{Gune}))$
$(\lambda x.x \text{ went to Colombo})$ =

e. $\lambda P \lambda Q{:}\exists z \in g(C_2)[Q(z)=1].\forall y \in g(C_2)[Q(y)=1 \rightarrow [P(\lambda x.y \neq x)]](\lambda S.S(\text{Gune}))$
$(\lambda x.x \text{ went to Colombo})$ =

f. $\lambda Q{:}\exists z \in g(C_2)[Q(z)=1].\forall y \in g(C_2)[Q(y)=1 \rightarrow [\lambda S.S(\text{Gune})(\lambda x.y \neq x)]](\lambda x.x \text{ went to Colombo})$ =

g. $\lambda Q{:}\exists z \in g(C_2)[Q(z)=1].\forall y \in g(C_2)[Q(y)=1 \rightarrow [\lambda x.y \neq x](\text{Gune})](\lambda x.x \text{ went to Colombo})$ =

h. $\lambda Q{:}\exists z \in g(C_2)[Q(z)=1].\forall y \in g(C_2)[Q(y)=1 \rightarrow y \neq \text{Gune}](\lambda x.x \text{ went to Colombo})$ =

i. $\forall y \in \{\text{Gune, Ranjit}\}[\lambda x.x \text{ went to Colombo}(y)=1 \rightarrow y \neq \text{Gune}]:$
$\exists z \in \{\text{Gune, Ranjit}\}[\lambda x.x \text{ went to Colombo}(z)=1]$ =

j. $\forall y \in \{\text{Gune, Ranjit}\}[y \text{ went to Colombo}=1 \rightarrow y \neq \text{Gune}]:$
$\exists z \in \{\text{Gune, Ranjit}\}[z \text{ went to Colombo}=1]$

The ordinary semantic value of the IP is given in (29). The focus semantic value of the IP is given below in (30); let us assume that the pragmatically-determined value of C_1 is identical to the focus semantic value of IP, as in (31).

(29) $[\![\text{IP}]\!]^g = \lambda w.\forall x \in \{\text{Gune, Ranjit}\}[x \text{ went to Colombo in } w=1 \rightarrow [x \neq \text{Gune}]]:$
$\exists z \in \{\text{Gune, Ranjit}\}[z \text{ went to Colombo in } w=1]$

(30) $[\![\text{IP}]\!]^{g,F} = \{\lambda w.\forall x \in \{\text{Gune, Ranjit}\}[x \text{ went to Colombo in } w=1 \rightarrow [x \neq \text{Gune}]]\}$

(31) $g(C_1) = [\![\text{IP}]\!]^{g,F} = \{\lambda w.\forall x \in \{\text{Gune, Ranjit}\}[x \text{ went to Colombo in } w=1 \rightarrow [x \neq \text{Gune}]]\}$

The derivation then continues as shown in (32), given some world w′.

(32) $[\![\text{FocusP}]\!]^g$ =

a. [-e C_1 [$\sim C_1$ [nevey C_2 [$\sim C_2$ Gune^F]] koləmbə giy-]] =

b. $\lambda w.\lambda P \lambda p{:}p \in P \;\&\; \exists q \in P[q(w)=1].p(w)(w')(g(C_1))([\![(29)]\!]^g)$ =

c. $\lambda P \lambda p{:}p \in P \;\&\; \exists q \in P[q(w')=1].p(w')(g(C_1))([\![(29)]\!]^g)$ =

d. $\lambda p{:}p \in g(C_1) \;\&\; \exists q \in g(C_1)[q(w')=1].p(w')([\![(29)]\!]^g)$ =

e. $[\![(29)]\!]^g(w')$: $[\![(29)]\!]^g \in g(C_1) \;\&\; \exists q \in g(C_1)[q(w')=1]$

Given that $g(C_1)$ is thus a singleton set containing only the ordinary semantic value of the IP, the first part of the definedness condition of (32e) will be satisfied; let us assume also that it is true that there is some member of the set {Gune, Ranjit} who did not go to Colombo, which satisfies the second part of the definedness condition. I thus dispense with the definedness condition of (32e) in the remainder of the derivation given below in (33) in order to render the semantics more easily comprehensible.

(33) $[\![(32e)]\!]^g$ =

a. $[\![(29)]\!]^g(w')$ =

b. $\forall x \in \{\text{Gune, Ranjit}\}[x \text{ went to Colombo in } w'=1 \rightarrow [x \neq \text{Gune}]]:$
$\exists z \in \{\text{Gune, Ranjit}\}[z \text{ went to Colombo in } w'=1]$

(33) asserts that for all x such that x went to Colombo in w′, x is not Gune, and presupposes that some member of the set {Gune, Ranjit} did go to Colombo in w′. Thus the desired semantics and pragmatics for constituent negation are obtained.

5.5 Summary

In this chapter I have defined the semantics for the morpheme *-e*, the verbal ending used in focussing constructions in Sinhala, and for constituent negation; the basic analysis of focus I adopt is that developed by Rooth (1985, 1992, 1996). The focussing construction in modern colloquial Sinhala resembles the English cleft construction in that it involves an existential presupposition; I propose that this presupposition is carried by the *-e* morpheme.

Thus having established a plausible treatment of the syntax and semantics of focus in modern colloquial Sinhala, I turn now to the semantics of Q-particles in interrogatives. Not only do interrogatives and focus construction share in common the use of "alternative" semantic values, but—as discussed in Chapter 3—many interrogative constructions are in fact focus constructions. Hence the analysis of the semantics of focus developed in this chapter plays an important role when we turn to the analysis of the semantics of Sinhala interrogatives.

Chapter 6

The semantics of Q-particles in interrogatives

In this chapter I lay out an explicit formal semantic treatment of Q-particles in modern colloquial Sinhala interrogatives, adopting a Hamblin-type analysis of *wh*-words and a choice-functional analysis of Q-particles.

Given that in modern colloquial Sinhala alternative questions and most *wh*-interrogatives, as well as some yes/no-constructions, involve focussed elements, an understanding of the semantics of Sinhala focus constructions is a prerequisite for the analysis of the semantics of Sinhala interrogatives, hence the importance of the previous chapter.

I begin in Section 6.1 by discussing the details of a Hamblin semantics of *wh*-words, and then show in Section 6.2 how the adoption of a Hamblin-style treatment of *wh*-words, in combination with an analysis of Q-particles as denoting variables over choice functions, allows us to explain the fact that *wh*-words may scope out of islands, so long as their associated Q-particle originates outside of any islands. Section 6.3 discusses the alternative account of Cable (2007), who, following Beck (2006), attempts to unite the "alternative" semantics of focus and *wh*-words—and provide a semantically-grounded account of intervention effects—by supposing that *wh*-words differ from other elements in their lack of an ordinary semantic value (and thus bear only a focus semantic value); I show that this theoretical move has a number of undesirable consequences, and thus is to be dispreferred to the account developed here. Finally, Section 6.4 provides a detailed account of the semantic and pragmatic properties of Sinhala *wh*-, alternative-, and yes/no-interrogatives.

6.1 A Hamblin semantics of *wh*-words

I adopt an analysis of interrogatives as denoting sets of propositions (Hamblin 1973; Karttunen 1977), and more specifically adopt Hamblin's analysis of *wh*-words as denoting sets of individuals, so that $[\![\text{who}]\!]^g = \{x \in D_e \mid x \in \text{human}'\}$. For Hamblin, non-*wh* elements denote singleton sets, e.g. $[\![\text{came}]\!]^g = \{\lambda x.x \text{ came}\}$. And thus $[\![\text{who came}]\!]^g$ can be straightforwardly calculated via pointwise function application: i.e. each element in $[\![\text{came}]\!]^g$ applies to each element in $[\![\text{who}]\!]^g$ and the results are collected together into a set, resulting in $[\![\text{who came}]\!]^g = \{x \text{ came} \mid x \in D_e \wedge x \in \text{human}'\}$. I differ from Hamblin (1973) in that I do not treat non-*wh* elements as denoting singleton sets, but rather assign them the standard Montagovian-type denotations. This move requires a special set of function application rules, which following the designation given to a similar formulation in Hagstrom 1998, we may refer to as 'flexible function application' (see also Rooth 1985; Bittner 1994; Heim 1994; Rullman & Beck 1998; Sternefeld 2001).

(1) [α β] = F($[\![\alpha, \beta]\!]^g$), let a=$[\![\alpha]\!]^g$, b=$[\![\beta]\!]^g$, for any a,b,
F(a,b) =

a. a(b)
b. $\{c \mid \exists y \in b\ [c = a(y)]\}$
c. $\{c \mid \exists x \in a\ [c = x(b)]\}$
d. $\{c \mid \exists x \in a, \exists y \in b\ [c = x(y)]\}$

Whichever is defined.

For F(a,b), rule (1a) is the ordinary rule of function-application, applicable where both a and b are elements with Montagovian-type denotations; rule (1b) handles cases where a bears a Montagovian-type denotation and b a Hamblin-type denotation; rule (1c) accounts for cases where a bears a Hamblin-type denotation and b a Montagovian-type denotation; and rule (1d) is Hamblin's (1973) rule of pointwise function application, appropriate where both a and b bear Hamblin-type denotations. The following examples illustrate: given the denotations in (2), the semantic computations of *who saw John*, *John saw whom*, and *who saw whom* proceed as in (3).

(2) a. $[\![\text{saw}]\!]^g = \lambda x.\lambda y.y \text{ saw } x$
b. $[\![\text{John}]\!]^g = \text{John}$
c. $[\![\text{who(m)}]\!]^g = \{x \in D_e \mid x \in \text{human}'\}$

(3) a. $[\![\text{Who saw John}]\!]^g$ = (by (2a), (2b), (2c))
(i) $\lambda x.\lambda y.y \text{ saw } x(\text{John})(\{x \in D_e \mid x \in \text{human}'\})$ = (by (1a))
(ii) $\lambda y.y \text{ saw John}(\{x \in D_e \mid x \in \text{human}'\})$ = (by (1b))
(iii) $\{x \text{ saw John} \mid x \in D_e \wedge x \in \text{human}'\}$

b. $[\![\text{John saw whom}]\!]^g$ = (by (2a), (2b), (2c))
(i) $\lambda x.\lambda y.y \text{ saw } x(\{x \in D_e \mid x \in \text{human}'\})(\text{John})$ = (by (1b))
(ii) $\{\lambda y.y \text{ saw } x \mid x \in D_e \wedge x \in \text{human}'\}(\text{John})$ = (by (1c))
(iii) $\{\text{John saw } x \mid x \in D_e \wedge x \in \text{human}'\}$

c. $[\![\text{Who saw whom}]\!]^g$ = (by (2a), (2b), (2c))
(i) $\lambda x.\lambda y.y \text{ saw } x(\{x \in D_e \mid x \in \text{human}'\})(\{x \in D_e \mid x \in \text{human}'\})$ = (by (1b))
(ii) $\{\lambda y.y \text{ saw } x \mid x \in D_e \wedge x \in \text{human}'\}(\{x \in D_e \mid x \in \text{human}'\})$ = (by (1d))
(iii) $\{x \text{ saw } y \mid x,y \in D_e \wedge x,y \in \text{human}'\}$

Wh-words, at least in languages like Sinhala, Japanese, Tlingit, Malayalam, thus can be analysed as denoting Hamblin-type sets. The purpose of Q-particles is thus, I argue, to transform Hamblin-type sets back into ordinary Montagoid-type elements.

6.2 Choice-functions & the denotation of Q-particles

In this section I argue that Q-particles should be treating as denoting variables over choice functions. One fact which argues for such an analysis is the possibility of *wh*-words in Sinhala scoping out of islands—so long as the associated Q-particle itself is not inside of the island.

Consider the fact that in Sinhala complex NPs are movement islands, as shown by (4).

(4) a. *oyā [[Chitra t_i dunnə] potə] kieuwe Ranjit$_i$-ṭə.
you [[Chitra t_i given] book] read.E Ranjit$_i$-DAT
"It was Ranjit$_i$ that you read the book that Chitra gave to t_i."

b. *Chitra [[Ranjit t_i gatta kiənə] kaṭəkatāwə] æhuwe ē potə$_i$
Chitra [[Ranjit t_i bought.A that] rumour] heard.E that book
"It was that book$_i$ that Chitra heard the rumour that Ranjit bought t_i." (Kishimoto 2005: 27)

WH+*də* cannot appear inside of a complex NP, as shown by (5).

(5) a. *oyā [[Chitra kāṭə **də** dunnə] potə] kieuwe?
you [[Chitra who.DAT ***də*** given] book] read.E
"To whom$_i$ did you read the book that Chitra gave t_i?"

b. *Chitra [[Ranjit monəwa **də** gatta kiənə] kaṭəkatāwə] æhuwe?
Chitra [[Ranjit what ***də*** bought.A that] rumour] heard.E
"What$_i$ did Chitra hear the rumour that Ranjit bought t_i?"(Kishimoto 2005: 29)

Wh-words themselves may be internal to islands, so long as there are no islands between the particle *də* and the CP over which it takes scope, as shown by (6) and (7).

(6) a. [[Ranjit monəwa gatta kiənə] kaṭəkatāwə] **də** Chitra æhuve?
[[Ranjit what bought-A that] rumour] **Q** Chitra heard-E
'What did Chitra hear the rumour that Ranjit bought?'

b. *[[Ranjit monəwa **də** gatta kiənə] kaṭəkatāwə] Chitra æhuve?
[[Ranjit what **Q** bought-A that] rumour] Chitra heard-E

(7) a. [kauru liyəpu potə] **də** Ranjit gatte?
[who written book] **Q** Ranjit bought-E?
'Who wrote the book Ranjit bought?'

b. *[kau **də** liyəpu potə] Ranjit gatte?
[who **Q** written book] Ranjit bought-E?

One might suppose that in (6a) and (7a) the *wh*-word obtains matrix scope by moving covertly to a position within the Spec of the lower CP—an 'escape hatch' position from which it is still visible to syntactic operations of the next phase. However, example (8) shows that this cannot be the case, as the *wh*-word may in fact be inside of an island which itself is inside of an island, so long as the Q-particle has no island barriers between it and the CP.

(8) a. [[Ranjit [kauru liyəpu potə] gatta kieənə] kaṭəkatāwə] **də** Chitra æhuve?
[[Ranjit [who written book] bought-A that] rumour] **Q** Chitra heard-E
'Who is the person *x* such that Chitra heard the rumour that Ranjit bought the book that *x* wrote?'

b. *[[Ranjit [kauru liyəpu potə] **də** gatta kieənə] kaṭəkatāwə] Chitra æhuve?
[[Ranjit [who written book] **Q** bought-A that] rumour] Chitra heard-E

c. *[[Ranjit [kau **də** liyəpu potə] gatta kieənə] kaṭəkatāwə] Chitra æhuve?
[[Ranjit [who **Q** written book] bought-A that] rumour] Chitra heard-E

Therefore, *wh*-interrogatives in Sinhala cannot be analysed as taking scope via covert (LF) movement. I argue that a choice-functional analysis of the Q-particle *də* provides a natural account of how *wh*-interrogatives may semantically scope out of islands (cp. the similar treatments of Hagstrom 1998; Cable 2007; Yatsushiro 2001, 2009).

Thus, I treat Question-particles like Sinhala *də* as variables over choice functions (cf. Hagstrom 1998; Yatsushiro 2001, 2009; Cable 2007 on choice-functional analyses of Question-particles; on other uses of choice-functional

analyses in natural language semantics, see Reinhart 1997, 1998; Winter 1997; Sternefeld 2001); where choice functions are defined in (9) below.

(9) **Choice Function**:
A function $f_{\langle\tau,\alpha\rangle}$ is a choice function (i.e. $CH(f_{\langle\tau,\alpha\rangle})$ holds) iff for every non-empty set $\mathscr{S}_\tau$, $f_{\langle\tau,\alpha\rangle}(\mathscr{S}_\tau)$ is defined and $f_{\langle\tau,\alpha\rangle}(\mathscr{S}_\tau)$ is in the extension of $\mathscr{S}_\tau$ (i.e. $\mathscr{S}_\tau(f_{\langle\tau,\alpha\rangle}(\mathscr{S}_\tau))$ holds).

That is, a choice function is a function which when applied to a set returns a member of that set. The denotation of a Question-particle (henceforth Q) is given below in (10):

(10) $[\![Q_i]\!]^g = g(i) \in D_{cf}$

Recall that, if we adopt a Hamblin-type analysis of *wh*-words, then, by the rules for flexible function application given above in (1), the Hamblin-type set semantics of a *wh*-word will 'infect' any element the *wh*-word composes with. This results in the Hamblin-type semantics 'spreading' through the derivation, unless it is closed-off by an element which takes a set and returns a member of that set. A choice function is just such an element. Thus, the semantic role of Q (=a choice function variable) can be seen then as a sort of 'cure' for Hamblin-type semantics, returning an element with ordinary Montagovian semantics.

I assume that the choice function variables represented by Question-particles like *ni* can be existentially bound by the denotation of the interrogative COMP,[1] see (11).

(11) $[\![[\text{COMP}^{\text{INT}}_i \text{ XP}]]\!]^g = \lambda p\left[\exists f \in D_{cf}.p = [\![\text{XP}']\!]^{g^{[f/i]}}\right]$

This allows us to correctly predict the denotations of sentences like (7a), as follows:[2]

(12) $[\![(7a)]\!]^g =$

a. $[\![[_{\text{CP}}\ \text{C-INT}_i\ [_{\text{IP}}\ [_{\text{QP}}\ [_{\text{DP}}\ \text{kauru liyəpu potə}_i\ ]\ \textbf{də}\ ]\ \text{Ranjit gatte}\ ]\ ]]\!]^g =$

b. $\lambda p[\exists f \in D_{cf}.p = [\![[_{\text{IP}}\ [_{\text{QP}}\ [_{\text{DP}}\ \text{kauru liyəpu potə}\ ]\ \textbf{də}_i\ ]\ \text{Ranjit gatte}\ ]]\!]^{g^{[f/i]}}] =$

c. $\lambda p[\exists f \in D_{cf}.p = \lambda w.\text{Ranjit bought } [\![\text{də}_i]\!]^{g^{[f/i]}}\left([\![[_{\text{DP}}\ \text{kauru liyəpu potə}\ ]]\!]^{g^{[f/i]}}\right) \text{ in w}] =$

d. $\lambda p[\exists f \in D_{cf}.p = \lambda w.\text{Ranjit bought } f(\{\iota x.\text{book}'(x) \text{ in w} \ \&\ \text{y wrote x in w} \mid y \in \text{human}' \text{ in w}\}) \text{ in w}] =$

Thus the analysis of Q-particles as denoting variables over choice functions, combined with a Hamblin semantics for *wh*-words, allows for a natural explanation of the semantic properties of *wh*-words and Q-particles with respect to islands.

The following section demonstrates why the semantic analysis developed here is to be preferred to that of Cable (2007), who adopts Beck's (2006) proposal that *wh*-words lack ordinary semantic values, bearing only focus semantic values, which are thus interpretable only after the application of a Q-particle (the only element, they suggest, which makes sole reference to its complement's focus semantic value). This analysis—while attractive in its unification of the "alternative" semantics of *wh*-words and focus, and moreover in its semantically-motivated account of intervention effects—is shown to entail a number of insupportable theoretical and empirical consequences.[3]

[1] I use COMP rather than C in order to prevent confusion between the complementiser and Rooth's pragmatic variable C.

[2] I here set aside the semantics associated with the focus element *-e*.

[3] Section 6.3 does not advance the development of the current thesis, and may be safely skipped by readers who are uninterested in intervention effects and/or already persuaded of the correctness of the present proposal.

6.3 Interlude: The semantics of *wh*-words: Q-particles, Focus, & Intervention effects

6.3.1 Cable's analysis of *wh*-words

The analysis of Cable 2007 capitalises on the relationship between *wh*-words and focus semantic values—namely that both denote sets of "alternatives". Cable's analysis suggests, following Beck (2006), that *wh*-words lack ordinary semantic values, with their contribution to interrogatives then being their focus semantic value (which is a set of alternatives). This requires that Q-particles make reference solely their complement's focus semantic value. Both Cable (2007) and Beck (2006) suggests that this treatment of *wh*-words can account for the appearance of intervention effects. In the remainder of Section 6.3, I point out the difficulties that such an analysis creates, arguing that it is necessary to recognise that *wh*-words, like all other lexical elements, bear both ordinary and focus semantic values.

6.3.2 Overview of intervention effects

In many languages certain configurations of *wh*-phrases and quantificational or focussing elements result in ungrammaticality, a phenomenon known in the literature as an intervention effect. Typical examples of intervention effects are illustrated by the Korean and German examples provided below in (13) and (14), respectively.

(13) a. *Minsu$_F$-man nuku-lûl po-ss-ni?
Minsu$_F$-only who-ACC see-PAST-Q
'Who did only Minsu see?'

b. Minsu$_F$-nun nuku-lûl po-ss-ni?
Minsu$_F$-TOP who-ACC see-PAST-Q
'Who did Minsu see?'

c. nuku-lûl Minsu$_F$-man po-ss-ni?
who-ACC Minsu$_F$-only see-PAST-Q
'Who did only Minsu see?'

(14) a. *Wen hat niemand wo gesehen?
Who.ACC has nobody where seen
'Where did nobody see whom?'

b. Wen hat Luise wo gesehen?
Who.ACC has Luise where seen
'Where did Luise see whom?'

c. Wen hat wo niemand gesehen?
Who.ACC has where nobody seen
'Where did nobody see whom?'

Intervention effects are exhibited by (13a) and (14a). In (13a) the *wh*-word *nukulûl* 'whom' is c-commanded by the focus-sensitive operator *-man* 'only'; in (14a) the *wh*-word *wo* 'where' is c-commanded by the quantificational element *niemand* 'nobody'. That it is focussing or quantificational elements (henceforth 'interveners') which are responsible for intervention effects is shown by the fact that where these elements are absent, as in (13b) and (14b), no ungrammaticlity results. The particular syntactic configuration is crucial, for both an intervener and a *wh*-word may be present in the same clause so long as the *wh*-word is not c-commanded by the intervener, as in (13c) and (14c),

where the *wh*-word has been scrambled out of the c-command domain of the intervener.[4] See Beck & Kim (1997) and Beck (2006) for examples of intervention effects in other languages.

Following Beck (2006) and Kim (2002), we can adopt the following generalisation of the patterns in (13) and (14) above.

(15) **Generalisation:**
A quantificational or focussing element may not intervene on the c-command path between a *wh*-phrase and its licensing element.

In German the '*wh*-licensing element' is the head of the interrogative CP; in Korean it is the Question-particle *ni*. Thus the basic configuration for intervention effects can be represented as in (16).[5]

(16) *[Q_i [. . . [intervener [. . . *wh*-phrase$_i$ . . .]]]]

6.3.3 Beck's and Cable's accounts of intervention effects

Beck (2006) suggests that intervention effects can be motivated semantically if we assume that *wh*-phrases play the same role as focussed phrases, i.e. to introduce alternatives into the semantic computation, but that, unlike focussed phrases, they bear no ordinary semantic value. She capitalises on the connection noted above in Section 5.1 between a Hamblin-style analysis of *wh*-words and a Roothian treatment of focus semantics. That is, she proposes that *wh*-words like *who* resemble other DPs like *John*, except that they lack ordinary semantic values; cp. (17) and (18) below.

(17) a. $[\![\text{John}_{(F)}]\!]^g$ = John
b. $[\![\text{John}_F]\!]^{g,F} = \{x \in D_e\}$

(18) a. $[\![\text{who}_{(F)}]\!]^g$ = undefined
b. $[\![\text{who}_F]\!]^{g,F} = \{x \in D_e\}$

The proposal that *wh*-words lack ordinary semantic values thus requires that *wh*-words must enter the derivation bearing a focus-feature, otherwise the semantic computation will crash, as discussed below. This means in essence that *wh*-words are assumed to be obligatorily focussed.

Beck's idea that the normal semantic contribution of a *wh*-word is its focus semantic value is theoretically attractive, given the connection between focus semantic values and an Hamblin-style analysis of *wh*-words, as discussed above. Furthermore, the idea that *wh*-words are obligatorily focussed finds empirical support. For instance, the prosodic and syntactic properties of interrogative pronouns in English are consistent with the notion that *wh*-words obligatorily enter the derivation bearing a focus-feature.

Sinhala also offers morphological evidence pointing to *wh*-words as obligatorily focussed. In modern colloquial Sinhala there are special 'focussing' forms of verbs which occur whenever a focussed element (or a trace of a focussed element) occupies a position within the c-command domain of the verb (see above Chapter 4 as well as Gair 1970,

[4] There is some variation in the strength of these intervention effects. As Beck (2006: 3n2) remarks, the intervention effects reported for Korean examples like (13a) hold for most speakers, but there are some speakers who do not perceive the intervention effect as strongly. Likewise, German examples like (14a) appear ungrammatical for most speakers, but not all. I concentrate here on those varieties in which the intervention effects are clearly perceived.

[5] In the remainder of the paper I concentrate on intervention effects where the intervener is clearly a focus-sensitive element, as in the Korean example (13a). Note that, crosslinguistically, the core set of interveners are indeed the focussing-operators 'only', 'even', and 'also' (Kim 2002; Beck 2006). Cases where the intervener is a quantificational element, as in example (14a), are more difficult to account for semantically, but I tentatively accept Beck's suggestion (2006: 24–27) that quantification elements involve a $\sim$ operator and thus that these too could ultimately reduce to the incompatibility of certain configurations of focus-sensitive elements and *wh*-words.

1983[1998], 1986[1998]b; Kariyakarawana 1998).[6] Compare the 'neutral' sentence in (19), where the default *-a* ending appears on the verb, with the sentence in (20), where the *-e* ending appears on the verb due to the presence of a focussed element in its scope.

(19) mamə [gaməṭə (*tamay)] yanna (/*yanne)
I.NOM [village.DAT (*EMPH)] go.PRES.A (/*go.PRES.E)
'I go to the village.'

(20) mamə [gaməṭə$_F$ (tamay)] yanne (/*yanna)
I.NOM [village.DAT$_F$ (EMPH)] go.PRES.E (/*go.PRES.A)
'It is to the village I go.'

The data in (19) and (20) show that in Sinhala there is a bidirectional dependency between the appearance of elements bearing focus and the appearance of E-verbal forms: if a focussed-element appears within the c-command domain of a verb, that verb obligatorily appears in the E-form; and, if a verb appears in an E-form, then there must be a focussed element within its c-command domain.

In *wh*-questions we also find the obligatory occurrence of the 'focussing' E-form of the verb,[7] as shown below in (21).

(21) [kau də] gaməṭə yanne (/*yanna)
[who.NOM *də*] village.DAT go.PRES.E (/*go.PRES.A)
'Who goes to the village?'

Thus here again we find empirical evidence supporting Beck's idea that *wh*-words are obligatorily focussed.

Despite these data, I shall demonstrate that in fact the analysis of Beck (2006) has a number of undesirable consequences. Before doing so, I present in brief the key features of Beck's proposal that intervention effects follow from focus semantic interpretation.

Beck's (2006) actual analysis of focus semantics differs from that of Rooth (1992) in certain respects, including the adoption of a set of distinguished variables (from Wold 1996). Here $[\![\alpha]\!]^g$ is the ordinary semantic interpretation and $[\![\alpha]\!]^{g,h}$ represents the focus semantic interpretation, where h represents the function which assigns values to distinguished variables.

Using this formalisation, the focussed and unfocussed versions of *John* receive the following ordinary and focus semantic interpretations:

(22) a. $[\![\text{John}]\!]^g$=john
b. $[\![\text{John}]\!]^{g,h}$=john
c. $[\![\text{John}_{F_1}]\!]^g$=john
d. $[\![\text{John}_{F_1}]\!]^{g,h}$=h(1) if 1∈dom(h), =john otherwise

More importantly, a *wh*-word like *who* receives the following interpretations.

(23) a. $[\![\text{who}]\!]^g$ is undefined
b. $[\![\text{who}]\!]^{g,h}$ is undefined
c. $[\![\text{who}_{F_1}]\!]^g$ is undefined
d. $[\![\text{who}_{F_1}]\!]^{g,h}$=h(1) if 1∈dom(h), undefined otherwise

[6] Again, following Kishimoto (2005), I gloss the 'neutral' *-a*/*-ā* endings as -A, the 'focussing' *-e*/*-ē* endings as -E.
[7] Except in a very restricted set of special circumstances, see Kishimoto (2005: 6–14) and Chapter 2.1 for details.

Beck's use of distinguished variables for focus semantic interpretations is motivated by the existence of examples where Q operators apparently must be able to selectively bind (*wh*-)variables.[8] However, Cable (2007: 245) notes that, though this analysis is successful in accounting for the relevant data, Beck's theory is weakened by her use of this Wold-style system for interpreting focus. Since Beck (2006) argues that intervention effects occur whenever a focus-sensitive operator other than the Q-operator attempts to compute the ordinary semantic value of *wh*-words, this account obviously rests on the assumption that, in general, focus-sensitive operators unselectively bind all variables in their scope. Cable (2007) argues that if we posit Q-particles which are distinct from interrogative COMP heads (a position I adopt as well) then the interrogative COMP head need not be treated as a focus-sensitive operator, but rather simply as a normal quantifier, binding ordinary variables within the IP.

Given the theoretical advantages of Cable's revisions, it would seem preferable to recast (23) in more usual (and more straightforward) Roothian terms, as below in (24).

(24) a. $[\![wh\text{-}\mathrm{XP}_\mathrm{F}]\!]^{g}$ = undefined
b. $[\![wh\text{-}\mathrm{XP}]\!]^{g,F}$ = undefined
c. $[\![wh\text{-}\mathrm{XP}_\mathrm{F}]\!]^{g,F} = \{x \in D_e\}$

Based on the assumptions underlying the assignment of ordinary and focus semantic interpretations to *wh*-words like *who* as in (23) or the equivalent (24) above, Beck (2006) shows that a semantic account of intervention effects is plausible. Consider the Korean examples of (13), repeated below as (25).

(25) a. *Minsu$_\mathrm{F}$-man nuku-lûl po-ss-ni?
Minsu$_\mathrm{F}$-only who-ACC see-PAST-Q
'Who did only Minsu see?'
b. Minsu$_\mathrm{F}$-nun nuku-lûl po-ss-ni?
Minsu$_\mathrm{F}$-TOP who-ACC see-PAST-Q
'Who did Minsu see?'
c. nuku-lûl Minsu$_\mathrm{F}$-man po-ss-ni?
who-ACC Minsu$_\mathrm{F}$-only see-PAST-Q
'Who did only Minsu see?'

Example (25) shows that when a focus-sensitive operator like *-man* "only" intervenes on the c-command path between the Q operator (which Beck posits to be in the head of CP) and a *wh*-word the result is ungrammatical. Compare (25a) against (25b), where in the latter the clause contains no focus-sensitive operator; (25c) shows that it is not simply the presence of a focus-sensitive operator which creates an intervention effect, rather such effects occur only when the focus-sensitive operator occurs between the Q operator and the *wh*-word.

Beck's account rests then on two proposals. The first is that *wh*-words are special in that—unlike all other lexical items—they bear only focus semantic values, but no ordinary semantic values (i.e. they are undefined with respect to ordinary semantic values). This further requires that any *wh*-word which receives an interpretation MUST be focussed, given that both the ordinary and focus semantic value of an unfocussed *wh*-word is undefined; see above (24).

The second aspect of Beck's account involves the proposal that Q is the only focus sensitive element which makes reference exclusively to its complement's focus semantic value. Since $\sim$ makes reference not only to the focus semantic value of its complement but also to its ordinary semantic value, the presence of any *wh*-words, not already 'closed off' by a Q, in the scope of $\sim$ will result in the entire clause becoming undefined. Beck (2006) predicts

[8] E.g., 'Baker Ambiguities' (Baker 1970) and cases where a focussed element inside of a question semantically associates with a focus-sensitive operator outside the question.

therefore that a *wh*-phrase must have Q as its closest c-commanding focus-sensitive operator; otherwise stated in (26).

(26) Generalisation: A *wh*-phrase may not have a $\sim$ operator as its closest c-commanding potential binder. (Beck 2006: 17)

Therefore Beck (2006) can account for the intervention effect in (25a), which would bear the structure shown below in (27).

(27) *[$_{CP}$ [$_{IP_2}$ man$_C$ [$_{IP_1}$ Minsu$_F$ nuku$_2$-lûl poss]] ni$_2$]
*[$_{CP}$ [$_{IP_2}$ only$_C$ [$_{IP_1}$ Minsu$_F$ who$_2$-ACC see.PAST]] Q$_2$]

$[\![IP_1]\!]^g$ is undefined for any g since the *wh*-word's ordinary semantic value is undefined; thus $[\![IP_2]\!]^g$ is undefined, as then is $[\![IP_2]\!]^{g,F}$. IP3 and CP inherit the undefined value of the lower phrases, thus $[\![IP_3]\!]^{g,F}$ is undefined and so is $[\![CP]\!]^g$, which results in the whole structure of (25a) being uninterpretable and therefore ungrammatical—given the principle of interpretability as stated below in (28).

(28) *Principle of Interpretability:* (Beck 2006: 16)
An LF must have an ordinary semantic interpretation.

In (25c), on the other hand, the *wh*-word has been scrambled out of the c-command domain of the focus-sensitive operator *man* "only" and thus no intervention effect is generated.

Beck 2006 therefore offers an attractive semantic account of intervention effects. However, a number of the proposals underlying this account create empirical difficulties, including the inability to distinguish between different *wh*-words (e.g. between *who* and *what*), the incorrect prediction that *wh*-words must always be focussed, and the inability to distinguish between ordinary *wh*-words and contrastively focussed *wh*-words, as discussed in more detail in the following subsections.

6.3.4 Distinguishing between *wh*-words

At the most basic level, Beck's proposal that *wh*-words' only possible contribution is their focus semantic value is problematic since it results in the prediction that all *wh*-words bear the same denotation. Consider the ordinary and focus semantics values Beck's proposal assigns to *who* and *what*, shown below in (29) and (30), respectively.

(29) a. $[\![who_F]\!]^g$ = undefined
b. $[\![who]\!]^{g,F}$ = undefined
c. $[\![who_F]\!]^{g,F} = \{x \in D_e\}$

(30) a. $[\![what_F]\!]^g$ = undefined
b. $[\![what]\!]^{g,F}$ = undefined
c. $[\![what_F]\!]^{g,F} = \{x \in D_e\}$

Who and *what* are thus not semantically distinguished in Beck's (2006) account. This is an unavoidable consequence, under Rooth's theory of alternative semantics, of treating *wh*-words' normal semantic contribution as their focus semantic value. That is, $[\![John_F]\!]^{g,F}$ must be $\{x \in D_e\}$ and NOT $\{x \in D_e \mid x \in human'\}$. Alternatives in Rooth's system are simply entities of the same semantic type as the ordinary semantic value of the focussed element, and are not otherwise semantically restricted. This is a crucial feature of Rooth's system: we cannot restrict $[\![John_F]\!]^{g,F}$ to

$\{x \in D_e \mid x \in \text{human}'\}$ since the set of humans may not contain all of the entities relevant for a particular pragmatic situation; consider (31) below, where **rover** is a member of **dog**′, not **human**′.

(31) I didn't see Rover, I saw JOHN.

The failure to distinguish semantically between *who* and *what* of course creates numerous problems, including the incorrect prediction that examples like (32) are contradictions, since $[\![\text{who John saw}]\!]^g$ and $[\![\text{what John saw}]\!]^g$ will thus bear the same denotation.[9]

(32) I asked who John saw, not what John saw.

Cable (2007), adopting many aspects of Beck's (2006) account, also inherits many of its problems. The difficulty Beck's account faces in analysing the focus semantic values of *wh*-words appears however in a different form in Cable 2007. Cable (2007: 136) successfully distinguishes between *who* and *what*, as shown in (33).

(33) a. $[\![\text{who}_{(F)}]\!]^g$ = undefined
b. $[\![\text{who}_F]\!]^{g,F} = \{x \in D_e \mid x \in \text{human}'\}$
c. $[\![\text{what}_{(F)}]\!]^g$ = undefined
d. $[\![\text{what}_F]\!]^{g,F} = \{x \in D_e \mid x \notin \text{human}'\}$

Cable (2007) does not offer a discussion of how the focus semantic values of *wh*-words in (33) are calculated—though he does review Rooth's theory (Cable 2007: 130–134), including the general formula (= (34)) for calculating focus semantic values—but it is obvious that he avoids the difficulties Beck faces in distinguishing *wh*-words only at a high cost: his assignment of focus semantic values to *who* and *what* is entirely stipulative.

Assuming the assignments in (33) above, the appeal to Rooth's alternative semantics for focus becomes rather tenuous, given that Cable (2007) must assume an idiosyncratic rule for assigning focus semantic values to each *wh*-word. In other words, the focus semantic values Cable (2007) posits for *wh*-words do not conform to the rule for assignment of focus semantic values, (1), repeated below in (34).

(34) $[\![X_F]\!]^{g,F} = \{x \in D_\tau \mid [\![X]\!]^g \in D_\tau\}$

Thus, Beck's (2006) account suffers from the very basic problem that, in following Rooth's general formula for the assignment of focus semantic values, all *wh*-words end up with identical denotations. Cable (2007), apparently realising that Beck's account is problematic in this respect (though he fails to discuss it), distinguishes between different *wh*-words, but only by abandoning Rooth's transparent formula for calculating focus semantic values, and assigning stipulative focus semantic values to *wh*-words.

In contrast, the Hamblin-semantics approach to *wh*-words adopted here can assign the desired denotations to *wh*-words without any problematic assumptions, as shown above in Chapter 6.1.

6.3.5 Not all *wh*-words are focussed: evidence from Sinhala & German

In many languages, indefinite pronouns are connected in some way to *wh*-pronouns (see Haspelmath 1997), in some cases being form-identical to *wh*-interrogative pronouns, as is the case in Japanese, discussed by Kuroda (1965), who dubs such forms 'indeterminate' pronouns. Beck's proposal, which implies that all *wh*-words are obligatorily focussed, encounters difficulties with respect to the analysis of this type of *wh*-indefinite.[10] The properties of *wh*-based

[9] See further Section 6.3.6 below.
[10] Cable (2007: 136n79) does realise this problem, and comments briefly upon it.

indefinites crosslinguistically argues against the prediction of Beck 2006 that *wh*-words must bear a focus-feature in order to be semantically computatable, given that such indefinites in many languages (including Sinhala and German) appear to be obligatorily UNfocussed.

Recall from Section 6.3.3 above that Sinhala verbs take a special ending when a focussed element appears within their c-command domain (e.g. in (20) above), and that this same focus-associated ending also obligatorily appears on the verb in the presence of interrogative pronouns (as in example (21) above). In contrast to *wh*-interrogatives, the form-identical *wh*-indefinite pronouns in Sinhala do not trigger the 'focussing' E-verbal forms—see example (35) below—indicating that in this case we cannot treat the *wh*-word as obligatorily focussed (see Gair & Sumangala 1991: 104).

(35) [kau də] gaməṭə yanna (/*yanne)
[who.NOM *də*] village.DAT go.PRES.A (/*go.PRES.E)
'Someone goes to the village.'

Not only is the indeterminate pronoun *kau* in (35) not obligatorily focussed, it is obligatorily unfocussed, as shown by the fact that the verb cannot appear in the E-form. According to Beck's (2006) proposal, (35) is predicted to be ungrammatical since the *wh*-word *kau* is unfocussed, and thus should remain undefined on Beck's analysis.

The Sinhala evidence may not appear to be entirely conclusive given that the type of focus associated with the -E verbal ending in Sinhala is specifically exhaustive/identificational focus (in the sense of É. Kiss 1998), whereas Beck's analysis requires only that *wh*-words bear some sort of focus-feature; non-exhaustive focus in Sinhala (e.g., in the case of focussed elements associated with EVEN, cf. Horvath 2007) does not trigger the -E ending, as shown by example (36).

(36) Gunapala$_F$-t gamaṭa giyā (/*giyē)
Gunapala$_F$-also/even village.DAT go.PAST.A (/*go.PAST.E)
'Even Gunapala$_F$ went to the village.'

However, Sinhala *wh*-pronouns exhibit a consistent pattern: where they are focussed, they are interpreted as interrogative pronouns, where they are unfocussed, they are interpreted as indefinite pronouns:

(37) a. [kau$_F$ də] gaməṭə yanne
[who$_F$.NOM *də*] village.DAT go.PRES.E
(i) "Who goes to the village?"
(ii) *"Someone goes to the village."

b. [kau də] gaməṭə yanna
[who.NOM *də*] village.DAT go.PRES.A
(i) *"Who goes to the village?"
(ii) "Someone goes to the village."

Further, the apparent incompatibility of focus with *wh*-indefinites is found in other languages as well. In certain varieties of German we also find indefinite pronouns which are form-identical with interrogative pronouns (Zaefferer 1991), e.g. *wer* "who", "someone", *was* "what", "something" etc., with focussed *wh*-words receiving the interrogative interpretation and their unfocussed counterparts the indefinite interpretation, as shown by the possible readings of (38a) and (38b)—which differ only in the placement of main sentence stress.

(38) (Examples from Haida 2007: 180–182)

a. Wer mag WAS?
   (i) 'Who likes what?'
   (ii) *'Who likes something?'

b. Wer MAG was?
   (i) *'Who likes what?'
   (ii) 'Who likes something?'

Here, as for Sinhala, we see that *wh*-based indefinites are obligatorily unfocussed, which is problematic for an account like Beck 2006 which requires *wh*-words to bear a focus feature in order to be interpretable.

Since the Hamblin-semantic approach adopted in this paper does not require that *wh*-words be focussed in order to receive an interpretation, that Sinhala and German *wh*-based indefinites are unfocussed is unproblematic on the current account.

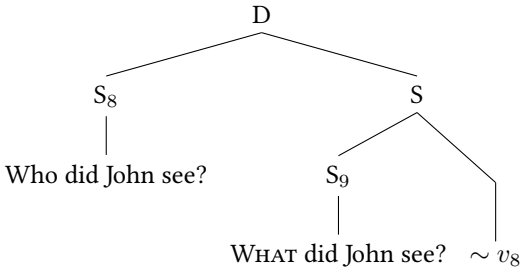
6.3.6 Contrastively-focussed *wh*-words

Though Beck (2006) treats all *wh*-words as obligatorily focussed, there are data which suggest that we need to recognise the existence of a distinction between focussed and unfocussed *wh*-words, in the sense that *wh*-words can participate in contrastive discourses which parallel the structure of (7). Consider the following context and dialogue, (39).

(39) Speaker A believes that John saw someone. Speaker B believes that John saw some sort of humanoid, but non-human, creature.
A: Who did John see?
B: You mean "WHAT did John see?"

The discourse in (39) can be given the following simplified structure, (40).

(40)



Though (39) is parallel in structure to (6), we cannot give a parallel account if we maintain either of the following two parts of the proposal in Beck 2006: (1) that *wh*-words lack ordinary semantic values, and (2) that all *wh*-words are obligatorily focussed.

In order for the dialogue in (39) to be felicitous, it must meet the condition given above in (10), repeated below as (41).

(41) $[\alpha \sim \mathrm{C}]$ is felicitious if $[\![\mathrm{C}]\!]^{g} \in [\![\alpha]\!]^{g,F}$

To meet this condition, we must derive the following semantic interpretations for $[\![[\mathrm{What}]_F \text{ did John see?}]\!]^{g}$, $[\![\text{Who did John see?}]\!]^{g}$, and $[\![[\mathrm{What}]_F \text{ did John see?}]\!]^{g,F}$.

(42) a. $[\![[\text{What}]_F \text{ did John see?}]\!]^g = \{\lambda w.\text{John saw x in w} \mid x \in D_e \wedge x \in \text{non-human}' \text{ in w}\}$
b. $[\![\text{Who did John see?}]\!]^g = \{\lambda w.\text{John saw x in w} \mid x \in D_e \wedge x \in \text{human}' \text{ in w}\}$
c. $[\![[\text{What}]_F \text{ did John see?}]\!]^{g,F} = \{\{\lambda w.\text{John saw x in w} \mid x \in P\} \mid P \subseteq D_e\}$

The denotation of $[\![[\text{What}]_F \text{ did John see?}]\!]^{g,F}$, given above in (42c), is essentially equivalent to what Karttunen (1977: 12n7) refers to as a 'second order question', that is, it is equivalent to a set of questions (in other words, a set of sets of propositions).[11]

Beck's analysis does not allow for the proper derivation of the required semantic values shown above in (42) since it does not allow for a distinction between focussed and non-focussed *wh*-words. In fact, in Beck's account (39) is predicted to be a contradiction, as her analysis assigns $[\![\text{who}_F]\!]^{g,F}$ and $[\![\text{what}_F]\!]^{g,F}$ the same semantic value, namely $\{x \in D_e\}$ (see Section 6.3.4), and thus $[\![\text{who}_F \text{ John saw}]\!]^g$ and $[\![\text{what}_F \text{ John saw}]\!]^g$ will both be equivalent to $\{\lambda w.\text{John saw x in w} \mid x \in D_e\}$.

On the account advocated here—since the ordinary semantic value of *wh*-words is defined as a set of (semantically-restricted) alternatives—we may calculate the focus semantic interpretation of a *wh*-word in the normal fashion, as in (1) above, repeated below as (43).

(43) $[\![X_F]\!]^{g,F} = \{x \in D_\tau \mid [\![X]\!]^g \in D_\tau\}$

Given the ordinary semantic interpretations for *who* and *what* as sets of individuals, i.e. as in (44), we predict that their focus semantic values will be sets of sets of individuals, as shown in (45) below.

(44) a. $[\![\text{who}_{(F)}]\!]^g = \{x \in D_e \mid x \in \text{human}'\}$
b. $[\![\text{what}_{(F)}]\!]^g = \{x \in D_e \mid x \notin \text{human}'\}$

(45) a. $[\![\text{who}_F]\!]^{g,F} = \{P \subseteq D_e\}$
b. $[\![\text{what}_F]\!]^{g,F} = \{P \subseteq D_e\}$

These focus semantic values allow us to derive the felicity of the dialogue in (39) above, repeated below as (46), as they predict the desired denotations of (42) above, repeated as (47).[12]

[11] Karttunen (1977) discusses second order questions in the context of echo questions, e.g. (i).

(i) a. Who did Mary see?
b. Who did WHO see?

For further discussion of the semantics of such questions see Artstein (2002), and also Cohen (2009).

[12] I omit choice functions from the denotation in (47) since there is some question regarding how choice functions variables enter the derivation in English, given that there are at least no overt Q-particles in English. It may be that English *wh*-words are morphologically complex, in that they include a choice function variable as part of their denotation, e.g. $[\![\text{who}_i]\!]^g = f_i \in D_{cf}.f_i(\{x \in D_e \mid x \in \text{human}'\})$, or that Q-particles are separate elements in English as in Korean, but that English Q-particles are morphologically unrealised; see Cable 2007 for a more fully fleshed-out version of the latter solution. Either solution would suffice for my purposes here, resulting in the full denotation of a question like *[What]$_F$ did John see?* as something along the lines of (i).

(i) a. $[\![[\text{COMP}_i^{\text{INT}}]([[\text{What}]_F \text{ did John see?}])]\!]^{g,F} =$
b. $\{\lambda p[\exists f \in D_{cf}.p{=}\lambda w.\text{John saw f(x) in w} \mid x \in P] \mid P \subseteq D_e\}$

(46) Speaker A believes that John saw someone. Speaker B believes that John saw some sort of humanoid, but non-human, creature.
A: Who did John see?
B: You mean "WHAT did John see?"

(47) a. $[\![[\text{What}]_F \text{ did John see?}]\!]^g = \{\lambda w.\text{John saw x in w} \mid x \in D_e \wedge x \in \text{non-human}' \text{ in w}\}$
b. $[\![\text{Who did John see?}]\!]^g = \{\lambda w.\text{John saw x in w} \mid x \in D_e \wedge x \in \text{human}' \text{ in w}\}$
c. $[\![[\text{What}]_F \text{ did John see?}]\!]^{g,F} = \{\{\lambda w.\text{John saw x in w} \mid x \in P\} \mid P \subseteq D_e\}$

Thus, assuming the felicity condition of (10), repeated as (48),

(48) $[\alpha \sim C]$ is felicitious if $[\![C]\!]^g \in [\![\alpha]\!]^{g,F}$

since $[\![\text{Who John saw}]\!]^g \in [\![[\text{What}]_F \text{ John saw}]\!]^{g,F}$, we predict the felicity of (46), as desired. Note that this result obtains naturally from the assumption that the ORDINARY semantic value of a *wh*-word is a set of alternatives, and that therefore the focus semantic value of a *wh*-word can thus be calculated like that of any other element. The contrastively-focussed *wh*-word data, e.g. as (39)—which is problematic for Beck (2006) as discussed above—thus can be handled in a straightforward manner by the account proposed here.

6.3.7 Conclusion

While the severity of the empirical problems of Beck's account obviously create numerous difficulties for her analysis, the basic idea that intervention effects are instances where the semantic computation is impeded in some way is theoretically very attractive. I believe that such an analysis can be formulated within the choice-functional analysis of Q-particles advocated herein; however, the details of the development of such an analysis lie beyond the scope of the present study.

It is clear, however, that *wh*-words must be treated as bearing both ordinary and focus semantic values, and that therefore the analysis of the *ordinary* semantic value of *wh*-words being that of a Hamblin-type set, as proposed above in Section 6.1, is preferable to that proposed by Cable 2007.

The following sections present explicit semantic derivations of *wh*-, alternative, and yes/no-questions in modern colloquial Sinhala.

6.4 The semantics and pragmatics of Sinhala *wh*-, alternative-, and yes/no-questions

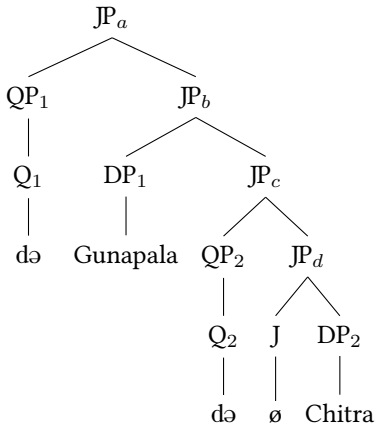
The syntactic and semantic analyses of Q-particles proposed in previous chapters, combined with the proposed analyses of disjunction and constituent negation, allows us to provide a fine-grained account of *wh*-, alternative, and yes/no-questions in Sinhala. I examine both normal and "non-presupposing" *wh*-questions in Section 6.4.1, alternative questions in Section 6.4.2, and finally yes/no-questions in Section 6.4.3.

The semantic treatment of of alternative- and yes/no-question requires a semantics for (dis)junction. A full discussion of the structure of disjunction is postponed until Chapter 10, to which the reader should refer for details and further argumentation.

In brief, recall from Chapter 3.5 that I assume that an alternative question like (49) bears the syntactic structure shown in (50).

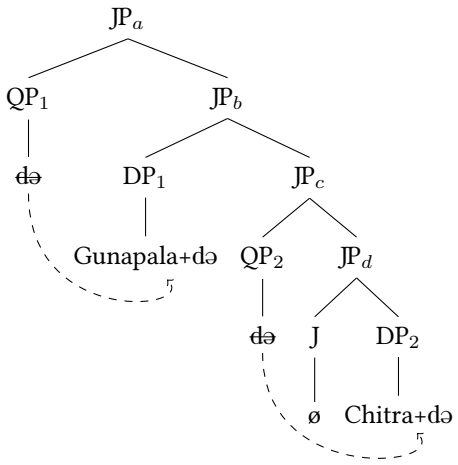
(49) Gunəpālə **də** Chitra **də** gamətə giyē?
Gunapala ***də*** Chitra ***də*** village.DAT go.PAST.E
'Was it Gunapala or Chitra who went to the village?'

(50)



Here QPs left-adjoin to the minimal JP containing both the head of JP and the disjunct with which the Q-particle is associated. The surface ordering of the Q-particles with respect to the disjuncts is the result of a later "PF-Structure" rule (Lowering), as shown in (50).

(51) "PF-Structure" of (50):



In Chapter 10 I arrive at the following semantic translation of J, which I adopt in the remainder of this chapter.

(52) 'Junction rule' :
$[\![\mathrm{J}]\!]^{g} = \lambda \mathrm{X}_{\langle \tau,t\rangle}\lambda \mathrm{Z}_{\langle cf,t\rangle}\lambda \mathrm{Y}_{\langle \tau,t\rangle}.\{\mathrm{Y}\} \cup \{\mathrm{Z}(\{\lambda \mathrm{P}_{\langle \tau,t\rangle}.\mathrm{P}\}(\mathrm{X}))\}$

Here J takes three arguments: two XPs of the same semantic type (e.g. two DPs) and a choice-function variable (i.e. a Q-particle). The Hamblin-type identity function $\{\lambda \mathrm{P}_{\langle \tau,t\rangle}.\mathrm{P}\}$ guarantees, by the rules of flexible function application, that the lower disjunct is a Hamblin-type element.[13] The basic function of rule (16) is to perform a union operation over a set containing the higher disjunct and a set containing the result of the choice-function variable applied to the lower disjunct. Again, refer to Chapter 10 for further justification of this analysis of disjunction.

[13]Essentially, the identity function transforms the lowest of a sequence of disjuncts into a Hamblin-type element, while leaving all subsequent disjuncts unaltered.

Before examining the semantic and pragmatic properties and analysis of alternative and yes/no-questions, let us first consider those of *wh*-questions, including those in which we find the "neutral" *-a* ending on the verb, rather than the focus-associated *-e* suffix usually present in *wh*-questions.

6.4.1 Semantics and pragmatics of Sinhala *wh*-questions

Recall from Chapter 2 above that under special circumstances in *wh*-questions the Q-particles *də* may occur clause-finally rather than following the *wh*-word; see examples (53) and (54)—here I am concerned with the placement of the bolded ***də*** within its clause; the position of the *də* of the matrix clause, which appears clause-finally since it is a yes/no question, is of no relevance here.

(53) a. Q: oyaa [[kauru aawa **də**] kiyəla] dannəwa də?
you [[who came-A ***də***] that] know-A *də*
'Do you know who came?'

b. A: oo. kauru-wat aawe næ.
yes anyone came-E not
'Yes. No-one came.'

(54) a. Q: oyaa [[kau **də** aawe] kiyəla] dannəwa də?
you [[who ***də*** came-E] that] know-A *də*
'Do you know who came?'

b. A: # oo. kauru-wat aawe næ.
yes anyone came-E not
'Yes. No-one came.'

I assume that in normal *wh*-questions (including the embedded clause of (54)) that the Q-particle *də* is focussed.[14]

Recall also that the existence presupposition is absent in configurations like (53). The semantic derivation for the embedded clause of (53) is given in (55) below.

(55) $[_{CP}\ [_{IP_a}\ [_{IP_b}\ \sim C\ [_{QP}\ [_{DP}$ kau] də^F] aaw-] -e C]] =

a. $[\![\text{COMP}_i^{\text{INT}}\ \text{IP}'_a]\!]^g$ =

b. $\lambda p\left[\exists f_1 \in D_{cf}.p{=}\lambda w.\left([\![\text{IP}'_a]\!]^{g^{[1/i]}}\right)\right]$ =

c. $\lambda p.\exists f_1 \in D_{cf}.p{=}\lambda w.[\![\text{-e}]\!]^g\left([\![\text{IP}'_b]\!]^{g^{[1/i]}}\right)$ =

d. $\lambda p.\exists f_1 \in D_{cf}.p{=}\lambda w.\lambda P.\lambda q{:}q \in P\ \&\ \exists r \in P[r(w){=}1].q(w)(g(C))\left([\![\text{IP}'_b]\!]^{g^{[1/i]}}\right)$

e. $\lambda p.\exists f_1 \in D_{cf}.p{=}\lambda w.\lambda q{:}q \in g(C)\ \&\ \exists r \in g(C)[r(w){=}1].q(w)\left([\![\text{IP}'_b]\!]^{g^{[1/i]}}\right)$

f. $\lambda p.\exists f_1 \in D_{cf}.p{=}\lambda w{:}[\![\text{IP}'_b]\!]^{g^{[1/i]}} \in g(C)\ \&\ \exists r \in g(C)[r(w){=}1].[\![\text{IP}'_b]\!]^{g^{[1/i]}}(w)$

The ordinary and focus semantic values of IP_b are given below:

(56) $[\![\text{IP}_b]\!]^g = \lambda w.f_1(\{y \in D_e \mid y \in \text{human}'\})$ came in w

(57) $[\![\text{IP}_b]\!]^{g,F} = \{\lambda w.f_\alpha(\{y \in D_e \mid y \in \text{human}'\})$ came in w $\mid f_\alpha \in D_{cf}\}$

[14] Assuming that *də* is focussed in normal *wh*- questions accounts for the fact that the *wh*-word+*də* may appear to the right of the verb, the position where we find other focussed elements. See above Chapter 4.

For the sake of exposition, let us assume that assignment by g to the pragmatic variable C here is identical to the focus semantic value of IP_b:

(58) $g(C) = [\![IP_b]\!]^{g,F}$

The derivation of (55) thus continues as below in (59).

(59) $\lambda p.\exists f_1 \in D_{cf}.p=\lambda w.f_1(\{y \in D_e \mid y \in human'\})$ came in w
iff $[\lambda w.f_1(\{y \in D_e \mid y \in human'\})$ came in w$] \in \{\lambda w.f_\alpha(\{y \in D_e \mid y \in human'\})$ came in w $\mid f_\alpha \in D_{cf}\}$
& $\exists r \in \{\lambda w.f_\alpha(\{y \in D_e \mid y \in human'\})$ came in w $\mid f_\alpha \in D_{cf}\}[r(w)=1]$

Setting aside the first part of the definedness condition (which is clearly satisfied here), we obtain:

(60) $\lambda p.\exists f_1 \in D_{cf}.p=\lambda w.f_1(\{y \in D_e \mid y \in human'\})$ came in w
iff $\exists r \in \{\lambda w.f_\alpha(\{y \in D_e \mid y \in human'\})$ came in w $\mid f_\alpha \in D_{cf}\}[r(w)=1]$

Thus, as desired, the question (60) presupposes that there is some member x of the set of humans such that x came.

In (54), on the other hand, I assume that *də* is unfocussed; the semantic derivation for the embedded clause of (54) is given below in (61).

(61) $[_{CP}\ [_{QP}\ [_{IP}\ [_{DP}$ kau] aawa] də]] =
a. $[\![COMP^{INT}_i\ QP']\!]^g$ =
b. $\lambda p[\exists f_1 \in D_{cf}.p=f_1(\{\lambda w.x$ came in w $\mid x \in human'\})]$

The absence of the *-e* morpheme in Focus here results in there being no existence presupposition, as the derivation in (61) correctly predicts.

6.4.2 The semantics and pragmatics of Sinhala disjunctions and alternative questions

I treat yes/no-questions as a special subtype of alternative questions. Before considering either alternative or yes/no questions, let us first consider the semantic derivation of a non-interrogative disjunction like example (62), since this derivation is simpler than that of an alternative or yes/no-question, but shares with those constructions the same basic syntax and semantics.

(62) Gunəpālə **hari** Chitra **hari** gaməṭə giyā.
Gunapala ***hari*** Chitra ***hari*** village.DAT go.PAST.A
'Gunapala or Chitra went to the village.'

Recall from Chapter 3.5 that the structure of (62) prior to the application of PF rules (i.e. the structure relevant to semantics) is that shown in (63).

(63) $[_{IP}\ [_{JP_1}\ [_{QP_1}$ hari$_1$] $[_{JP_2}\ [_{DP_1}$ Gunəpālə] $[_{JP_3}\ [_{QP_2}$ hari$_2$] $[_{JP_4}$ J $[_{DP_2}$ Chitra]]] gaməṭə giyā]]

The semantic derivation of (62) can then proceed as in (64).

(64) $[\![(62)]\!]^g$ =

a. $[_{IP}\ [_{JP_1}\ [_{QP_1}\ \text{hari}_1\ ]\ [_{JP_2}\ [_{DP_1}\ \text{Gunəpālə}\ ]\ [_{JP_3}\ [_{QP_2}\ \text{hari}_2\ ]\ [_{JP_4}\ J\ [_{DP_2}\ \text{Chitra}\ ]\ ]\ ]\ ]\ \text{gamətə giyā}\ ]\ ]$ =

b. $\exists f_1, f_2 \in D_{cf}.\lambda w.\lambda x.x$ went to the village in $w(f_1([\![[JP_{1_b}]]\!]^g))$

c. $\exists f_1, f_2 \in D_{cf}.\lambda w.\lambda x.x$ went to the village in w
$(f_1(\lambda X.\lambda Z.\lambda Y.\{Y\}\cup\{Z(\{\lambda P.P\}(X))\}([\![DP_2]\!]^g)([\![Q_2]\!]^g)([\![DP_1]\!]^g)))$

d. $\exists f_1, f_2 \in D_{cf}.\lambda w.\lambda x.x$ went to the village in w
$(f_1(\lambda X.\lambda Z.\lambda Y.\{Y\}\cup\{Z(\{\lambda P.P\}(X))\}(\text{Chitra})(f_2)(\text{Gunəpālə})))$

e. $\exists f_1, f_2 \in D_{cf}.\lambda w.\lambda x.x$ went to the village in $w(f_1(\{\text{Gunəpālə}\}\cup\{f_2(\{\lambda P.P\}(\text{Chitra}))\}))$

f. $\exists f_1, f_2 \in D_{cf}.\lambda w.\lambda x.x$ went to the village in $w(f_1(\{\text{Gunəpālə}\}\cup\{f_2(\{\text{Chitra}\})\}))$

g. $\exists f_1, f_2 \in D_{cf}.\lambda w.\lambda x.x$ went to the village in $w(f_1(\{\text{Gunəpālə}, f_2(\{\text{Chitra}\})\}))$

h. $\exists f_1, f_2 \in D_{cf}.\lambda w.f_1(\{\text{Gunəpālə}, f_2(\{\text{Chitra}\})\})$ went to the village in w ≡

i. λw.[Gunapala went to the village in w] ∨ [Chitra went to the village in w]

An alternative question like example (65), bears the LF structure of (66).

(65) Gunəpālə **də**^{F} Chitra **də**^{F} gamətə giyē?
Gunapala ***də***^{F} Chitra **də**^{F} village.DAT go.PAST.E
'Was it Gunapala or Chitra who went to the village?'

(66) $[_{CP}\ [_{FocusP}\ [_{IP}\ [_{JP_{1_a}}\ [_{QP_1}\ \text{də}^F\ ]\ [_{JP_{1_b}}\ [_{DP_1}\ \text{Gunəpālə}\ ]\ [_{JP_{2_a}}\ [_{QP_2}\ \text{də}^F\ ]\ [_{JP_{2_b}}\ J\ [_{DP_2}\ \text{Chitra}\ ]\ ]\ ]\ ]\ ]\ ]\ \text{gamətə giy-}\ ]\ \text{-e}\ ]\ ]$

The semantic derivation of (65) may the proceed as in (67).

(67) $[\![(65)]\!]^g$ =

a. $[_{CP}\ [_{FocusP}\ [_{IP}\ [_{JP_{1_a}}\ [_{QP_1}\ \text{də}^F\ ]\ [_{JP_{1_b}}\ [_{DP_1}\ \text{Gunəpālə}\ ]\ [_{JP_{2_a}}\ [_{QP_2}\ \text{də}^F\ ]\ [_{JP_{2_b}}\ J\ [_{DP_2}\ \text{Chitra}\ ]\ ]\ ]\ ]\ ]\ ]\ \text{gamətə giy-}\ ]$
$\text{-e}\ ]\ ]$ =

b. $[\![\text{COMP}^{\text{INT}}_{i,j}\ \text{FocusP}]\!]^g$ =

c. $\lambda p\left[\exists f_1,f_2 \in D_{cf}.p=\lambda w.\left([\![\text{FocusP}']\!]^{g^{[1/i,2/j]}}\right)\right]$ =

d. $\lambda p.\exists f_1,f_2 \in D_{cf}.p=\lambda w.[\![\text{-e}]\!]^g\left([\![\text{IP}']\!]^{g^{[1/i,2/j]}}\right)$ =

e. $\lambda p.\exists f_1,f_2 \in D_{cf}.p=\lambda w.\lambda P.\lambda q\text{: } q \in P \ \&\ \exists r \in P[r(w)=1].q(w)(g(C))\left([\![\text{IP}']\!]^{g^{[1/i,2/j]}}\right)$ =

f. $\lambda p.\exists f_1,f_2 \in D_{cf}.p=\lambda w.\lambda q\text{:}q \in g(C)\ \&\ \exists r \in g(C)[r(w)=1].q(w)\left([\![\text{IP}']\!]^{g^{[1/i,2/j]}}\right)$ =

g. $\lambda p.\exists f_1,f_2 \in D_{cf}.p=\lambda w[\![\text{IP}']\!]^{g^{[1/i,2/j]}}(w)\text{: } [\![\text{IP}']\!]^{g^{[1/i,2/j]},F} \in g(C)\ \&\ \exists r \in g(C)[r(w)=1]$

The ordinary semantic value of the IP is provided below in (68).

(68) $[\![\text{IP}]\!]^g$ =

a. $\lambda w.\lambda x.x$ went to the village in $w\left([\![\text{JP}'_1]\!]^{g^{[1/i,2/j]}}\right)$ =

b. $\lambda w.\lambda x.x$ went to the village in $w\left(f_1\left([\![\text{J}']\!]^{g^{[1/i,2/j]}}\left([\![\text{DP}'_2]\!]^{g^{[1/i,2/j]}}\right)(f_2)\left([\![\text{DP}'_1]\!]^{g^{[1/i,2/j]}}\right)\right)\right)$ =

c. $\lambda w.\lambda x.x$ went to the village in w
$\left(f_1\left(\lambda X.\lambda Z.\lambda Y.\{Y\}\cup\{Z(\{\lambda P.P\}(X))\}\left([\![\text{DP}'_2]\!]^{g^{[1/i,2/j]}}\right)(f_2)\left([\![\text{DP}'_1]\!]^{g^{[1/i,2/j]}}\right)\right)\right)$ =

d. $\lambda w.\lambda x.x$ went to the village in w
$(f_1(\lambda X.\lambda Z.\lambda Y.\{Y\}\cup\{Z(\{\lambda P.P\}(X))\}(\text{Chitra})(f_2)(\text{Gunapala})))$ =

e. $\lambda w.\lambda x.x$ went to the village in $w(f_1(\lambda Z.\lambda Y.\{Y\}\cup\{Z(\{\lambda P.P\}(\text{Chitra}))\}(f_2)(\text{Gunapala})))$ =

f. $\lambda w.\lambda x.x$ went to the village in $w(f_1(\lambda Y.\{Y\}\cup\{f_2(\{\lambda P.P\}(\text{Chitra}))\}(\text{Gunapala})))$ =

g. $\lambda w.\lambda x.x$ went to the village in $w(f_1(\{\text{Gunapala}\}\cup\{f_2(\{\lambda P.P\}(\text{Chitra}))\}))$ =

h. $\lambda w.f_1(\{\text{Gunapala}\}\cup\{f_2(\{\lambda P.P\}(\text{Chitra}))\})$ went to the village in w =

i. $\lambda w.f_1(\{\text{Gunapala}\}\cup\{f_2(\{\text{Chitra}\})\})$ went to the village in w

Which entails that the focus semantic value of the IP is:

(69) $[\![\text{IP}]\!]^{g,F}$ = $\{\lambda w.f_\alpha(\{\text{Gunapala},f_\beta(\{\text{Chitra}\})\})$ went to the village in w $\mid f_\alpha, f_\beta \in D_{cf}\}$

Again, for the sake of exposition, I assume that here the assignment of g to the pragmatically-determined variable C is identical to the focus semantic value of the IP.

(70) $[\![g(C)]\!]^g = [\![\text{IP}]\!]^{g,F}$

Given this assignment to C, the first part of the definedness condition may be dispensed with, and the derivation continues as shown below:

(71) $\lambda p.\exists f_1,f_2 \in D_{cf}.p=\lambda w.f_1(\{\text{Gunapala},f_2(\{\text{Chitra}\})\})$ went to the village in w:
$\exists r \in \{\lambda w.f_\alpha(\{\text{Gunapala},f_\beta(\{\text{Chitra}\})\})$ went to the village in w $\mid f_\alpha, f_\beta \in D_{cf}\}[r(w)=1]$

That is, (71) presupposes that either Chitra or Gunapala went to the village in w: the desired outcome.

6.4.3 The semantics and pragmatics of Sinhala yes/no-questions

Recall that in addition to "neutral" yes/no-questions like (72a), Sinhala also employs "focussed" yes/no-questions like (72b).[15]

(72) a. Ranjit aawa də?
Ranjit come.PAST.A *də*
"Did Ranjit come?"

b. Ranjit^F də aawe?
Ranjit^F *də* come.PAST.E
"Was it Ranjit who came?"

As discussed in Chapter 3.6, I treat yes/no-questions as a special subtype of alternative question, with an elided *or not X* constituent. Thus I assume that the LF of (72a) is as in (73).

[15]Ginsburg (2010) provides a syntactic analysis of Sinhala "neutral" and "focussed" yes/no questions which, like the analysis advanced here, recognises that the difference between these two constructions is based (in part) on whether the Q-particle (*də*) bears a focus feature or not.

(73) $[_{CP}\ [_{JP_{1_a}}\ [_{QP_1}$ də] $[_{JP_{1_b}}\ [_{IP}$ Ranjit aawa] $[_{JP_{2_a}}\ [_{QP_2}$ də] $[_{JP_{2_b}}$ J $[_{IP}$ not Ranjit aawa]]]]]]

Since this derivation is more complicated than the derivations shown in the preceding sections, I provide the derivation in smaller pieces before showing the derivation of the entire clause. The denotations for DP_1, DP_2 are given below.

(74) $[\![DP_1]\!]^g$ = λw.Ranjit came in w

(75) $[\![DP_2]\!]^g$ = λw.¬[Ranjit came in w]

The denotation of JP_{1_a} is:

(76) $[\![JP_{1_a}]\!]^g$ =

a. $f_i(\lambda X.\lambda Z.\lambda Y.\{Y\}\cup\{Z(\{\lambda P.P\}(X))\})([\![DP_2]\!]^g)([\![Q_2]\!]^g)([\![DP_1]\!]^g))$ =

b. $f_i(\lambda X.\lambda Z.\lambda Y.\{Y\}\cup\{Z(\{\lambda P.P\}(X))\}(\lambda w.\neg[\text{Ranjit came in w}])(f_j)(\lambda w.\text{Ranjit came in w}))$ =

c. $f_i(\lambda Z.\lambda Y.\{Y\}\cup\{Z(\{\lambda P.P\}(\lambda w.\neg[\text{Ranjit came in w}]))\}(f_j)(\lambda w.\text{Ranjit came in w}))$ =

d. $f_i(\lambda Y.\{Y\}\cup\{f_j(\{\lambda P.P\}(\lambda w.\neg[\text{Ranjit came in w}]))\}(\lambda w.\text{Ranjit came in w}))$ =

e. $f_i(\{\lambda w.\text{Ranjit came in w}\}\cup\{f_j(\{\lambda P.P\}(\lambda w.\neg[\text{Ranjit came in w}]))\})$ =

f. $f_i(\{\lambda w.\text{Ranjit came in w}\}\cup\{f_j(\{\lambda w.\neg[\text{Ranjit came in w}]\})\})$ =

g. $f_i(\{\lambda w.\text{Ranjit came in w}, f_j(\{\lambda w.\neg[\text{Ranjit came in w}]\})\})$

Given the derivation of the above pieces, the denotation of the entire CP will thus be:

(77) $[\![COMP^{INT}_{i,j}\ JP_1]\!]^g$ =

a. $\lambda p\left[\exists f_1,f_2 \in D_{cf}.p=[\![JP'_1]\!]^{g^{[1/i,2/j]}}\right]$ =

b. $\lambda p[\exists f_1,f_2 \in D_{cf}.p=f_1(\{\lambda w.\text{Ranjit came in w}, f_2(\{\lambda w.\neg[\text{Ranjit came in w}]\})\})]$

The resulting denotation in (77) is the sort of denotation we would expect for a yes/no question.

The focussed yes/no-question in (72b), repeated below as (78), I assume to bear the LF structure of (79).

(78) Ranjit^F də aawe?
Ranjit^F *də* come.PAST.E
"Was it Ranjit who came?"

(79) $[_{CP}\ [_{FocusP}\ [_{IP}\ \sim C_1\ [_{JP_{1_a}}\ [_{QP_1}$ də^F] $[_{JP_{1_b}}\ [_{DP}$ Ranjit^F] $[_{JP_{2_a}}\ [_{QP_2}$ də^F] $[_{JP_{2_b}}$ J $[_{NegP}\ [_{Neg}$ nevey C_2] [$\sim C_2$ $[_{DP}$ Ranjit^F]]] t_j]]]]] aaw-] -e C_1]]

Again, I break the derivation of (78) into smaller pieces for ease of exposition.

The focus semantic value of **Ranjit** is the set of all individuals of the same semantic type, namely all individuals of type e.

(80) $[\![\text{Ranjit}^F]\!]^{g,F} = \{x \mid x \in D_e\}$

For the sake of exposition, let us assume that the pragmatically determined variable C_2 is the set of **Ranjit** and **Gune**:

(81) $g(C_2)$ = {Ranjit, Gune} ($\subseteq [\![\text{Ranjit}^F]\!]^{g,F}$)

The denotations of DP_1 and DP_2 are given below:[16]

[16]See Chapter 5 for the semantics of constituent negation.

(82) $[\![\mathrm{DP}_1]\!]^g$ = Ranjit

(83) $[\![\mathrm{DP}_2]\!]^g$ =

a. $\lambda P \lambda \mathrm{P}.\lambda \mathrm{Q}{:}\exists \mathrm{z} \in P[\mathrm{Q(z)=1}].\forall \mathrm{y} \in P[\mathrm{Q(y)}\rightarrow[\mathrm{P}(\lambda \mathrm{x.y}\neq \mathrm{x})]](\mathrm{g}(\mathrm{C}_2))(\lambda \mathrm{S.S(Ranjit)})$ =

b. $\lambda \mathrm{P}.\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(z)=1}].\forall \mathrm{y} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(y)}\rightarrow[\mathrm{P}(\lambda \mathrm{x.y}\neq \mathrm{x})]](\lambda \mathrm{S.S(Ranjit)})$ =

c. $\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(z)=1}].\forall \mathrm{y} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(y)}\rightarrow[\lambda \mathrm{S.S(Ranjit)}(\lambda \mathrm{x.y}\neq \mathrm{x})]]$ =

d. $\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(z)=1}].\forall \mathrm{y} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(y)}\rightarrow[\lambda \mathrm{x.y}\neq \mathrm{x(Ranjit)}]]$ =

e. $\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(z)=1}].\forall \mathrm{y} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(y)}\rightarrow[\mathrm{y}\neq \mathrm{Ranjit}]]$

Given the denotations of these pieces, the derivation of (78) proceeds as follows:

(84) $[_{\mathrm{CP}}\ [_{\mathrm{FocusP}}\ [_{\mathrm{IP}}\ \sim \mathrm{C}_1\ [_{\mathrm{JP}_{1_a}}\ [_{\mathrm{QP}_1}\ \mathrm{d\ni}^{\mathrm{F}}\ ]\ [_{\mathrm{JP}_{1_b}}\ [_{\mathrm{DP}}\ \mathrm{Ranjit}^{\mathrm{F}}\ ]\ [_{\mathrm{JP}_{2_a}}\ [_{\mathrm{QP}_2}\ \mathrm{d\ni}^{\mathrm{F}}\ ]\ [_{\mathrm{JP}_{2_b}}\ \mathrm{J}\ [_{\mathrm{NegP}}\ [_{\mathrm{Neg}}\ \mathrm{nevey}\ \mathrm{C}_2\ ]\ [\ \sim \mathrm{C}_2\ [_{\mathrm{DP}}\ \mathrm{Ranjit}^{\mathrm{F}}\ ]\ ]\ ]\ \mathrm{t}_j\ ]\ ]\ ]\ ]\ ]$ aaw-] -e C_1]] =

a. $[\![\mathrm{COMP}^{\mathrm{INT}}_{i,j}\ \mathrm{IP}_a]\!]^g$ =

b. $\lambda \mathrm{p}\left[\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\left([\![\mathrm{FocusP}']\!]^{g^{[1/i,2/j]}}\right)\right]$ =

c. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.[\![\text{-e}]\!]^g(\mathrm{g}(\mathrm{C}_1))\left([\![\mathrm{IP}']\!]^{g^{[1/i,2/j]}}\right)$ =

d. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda P.\lambda \mathrm{q}{:}\mathrm{q} \in P\ \&\ \exists \mathrm{r} \in P[\mathrm{r(w)=1}].\mathrm{q(w)}(\mathrm{g}(\mathrm{C}_1))\left([\![\mathrm{IP}']\!]^{g^{[1/i,2/j]}}\right)$ =

e. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}\left([\![\mathrm{IP}']\!]^{g^{[1/i,2/j]}}\right)$ =

f. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].$
$\mathrm{q(w)}\left(\lambda \mathrm{w}'.\lambda \mathrm{x.x\ came\ in\ w}'\left([\![\mathrm{JP}'_1]\!]^{g^{[1/i,2/j]}}\right)\right)$ =

g. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].$
$\mathrm{q(w)}\left(\lambda \mathrm{w}'.\lambda \mathrm{x.x\ came\ in\ w}'\left([\![\mathrm{JP}'_1]\!]^{g^{[1/i,2/j]}}\right)\right)$ =

h. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}(\lambda \mathrm{w}'.\lambda \mathrm{x.x\ came\ in\ w}'$
$\left([\![\mathrm{J}]\!]^{g^{[1/i,2/j]}}\left([\![\mathrm{DP}'_2]\!]^{g^{[1/i,2/j]}}\right)(\mathrm{f}_2)\left([\![\mathrm{DP}_1]\!]^{g^{[1/i,2/j]}}\right)\right)$ =

i. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}(\lambda \mathrm{w}'.\lambda \mathrm{x.x\ came\ in\ w}'$
$\left(\mathrm{f}_1\left(\lambda \mathrm{X}.\lambda \mathrm{Z}.\lambda \mathrm{Y}.\{\mathrm{Y}\}\cup\{\mathrm{Z}(\{\lambda \mathrm{P.P}\}(\mathrm{X}))\}\left([\![\mathrm{DP}'_2]\!]^{g^{[1/i,2/j]}}\right)(\mathrm{f}_2)\left([\![\mathrm{DP}'_1]\!]^{g^{[1/i,2/j]}}\right)\right)\right)$ =

j. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}(\lambda \mathrm{w}'.\lambda \mathrm{x.x\ came\ in\ w}'$
$(\mathrm{f}_1(\lambda \mathrm{X}.\lambda \mathrm{Z}.\lambda \mathrm{Y}.\{\mathrm{Y}\}\cup\{\mathrm{Z}(\{\lambda \mathrm{P.P}\}(\mathrm{X}))\}(\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(z)=1}].$
$\forall \mathrm{y} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(y)}\rightarrow[\mathrm{y}\neq \mathrm{Ranjit}]])(\mathrm{f}_2)(\mathrm{Ranjit}))))$ =

k. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}(\lambda \mathrm{w}'.\lambda \mathrm{x.x\ came\ in\ w}'$
$(\mathrm{f}_1(\{\mathrm{Ranjit}\}\cup\{\mathrm{f}_2(\{\lambda \mathrm{P.P}\}(\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(z)=1}].\forall \mathrm{y} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(y)}\rightarrow[\mathrm{y}\neq \mathrm{Ranjit}]]))\}))))$ =

l. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}(\lambda \mathrm{w}'.\lambda \mathrm{x.x\ came\ in\ w}'$
$(\mathrm{f}_1(\{\mathrm{Ranjit}\}\cup\{\mathrm{f}_2(\{\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(z)=1}].\forall \mathrm{y} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(y)}\rightarrow[\mathrm{y}\neq \mathrm{Ranjit}]]\})\}))))$ =

m. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}(\lambda \mathrm{w}'.$
$\mathrm{f}_1(\{\mathrm{Ranjit}\}\cup\{\mathrm{f}_2(\{\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(z)=1}].\forall \mathrm{y} \in \mathrm{g}(\mathrm{C}_2)[\mathrm{Q(y)}\rightarrow[\mathrm{y}\neq \mathrm{Ranjit}]]\})\})$ came in w′) =

n. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}(\lambda \mathrm{w}'.$
$\mathrm{f}_1(\{\mathrm{Ranjit}\}\cup\{\mathrm{f}_2(\{\lambda \mathrm{Q}{:}\exists \mathrm{z}{\in}\{\mathrm{Ranjit,\ Gune}\}[\mathrm{Q(z)=1}].\forall \mathrm{y}{\in}\{\mathrm{Ranjit,\ Gune}\}[\mathrm{Q(y)}\rightarrow[\mathrm{y}\neq \mathrm{Ranjit}]]\})\})$
came in w′) =

o. $\lambda \mathrm{p}.\exists \mathrm{f}_1,\mathrm{f}_2 \in \mathrm{D}_{cf}.\mathrm{p}=\lambda \mathrm{w}.\lambda \mathrm{q}{:}\mathrm{q} \in \mathrm{g}(\mathrm{C}_1)\ \&\ \exists \mathrm{r} \in \mathrm{g}(\mathrm{C}_1)[\mathrm{r(w)=1}].\mathrm{q(w)}(\lambda \mathrm{w}'.$
$\mathrm{f}_1(\{\mathrm{Ranjit},\ \mathrm{f}_2(\{\lambda \mathrm{Q}{:}\exists \mathrm{z} \in \{\mathrm{Ranjit,\ Gune}\}[\mathrm{Q(z)=1}].\forall \mathrm{y} \in \{\mathrm{Ranjit,\ Gune}\}[\mathrm{Q(y)}\rightarrow[\mathrm{y}\neq \mathrm{Ranjit}]]\})\})$
came in w′)

Let us assume that the denotation of the pragmatically determined variable C_1 is identical to the focus semantic value of the IP, for the sake of exposition.

(85) $g(C_1) = [\![IP]\!]^{g,F}$

Given this assumption we can dispense with the first part of the definedness condition of the FocusP (namely, $q \in g(C_1)$), since by definition the $[\![IP']\!]^g \in [\![IP']\!]^{g,F}$.[17] The second part of the definedness condition of the FocusP ($\exists r \in g(C_1)[r(w)=1]$) simply requires that the proposition is true for some pair of choice functions; given that this is a yes/no question, the only two possibilities are "Ranjit came" and "Ranjit did not come". Logically, one of these two possibilities must be true, and so we can also dispense with the second part of the definedness condition of the FocusP.

Therefore the final denotation for the question can be given as in (86).

(86) $[\![(78)]\!]^g =$

a. $\lambda p.\exists f_1,f_2 \in D_{cf}.p=\lambda w.\lambda q.q(w)(\lambda w'.$
$f_1(\{\text{Ranjit}, f_2(\{\lambda Q:\exists z \in \{\text{Ranjit, Gune}\}[Q(z)=1].\forall y \in \{\text{Ranjit, Gune}\}[Q(y)\rightarrow[y\neq\text{Ranjit}]]\})\})$
came in $w') =$

b. $\lambda p.\exists f_1,f_2 \in D_{cf}.p=\lambda w.\lambda w'.$
$f_1(\{\text{Ranjit}, f_2(\{\lambda Q:\exists z \in \{\text{Ranjit, Gune}\}[Q(z)=1].\forall y \in \{\text{Ranjit, Gune}\}[Q(y)\rightarrow[y\neq\text{Ranjit}]]\})\})$
came in $w'(w) =$

c. $\lambda p.\exists f_1,f_2 \in D_{cf}.p=\lambda w.$
$f_1(\{\text{Ranjit}, f_2(\{\lambda Q:\exists z \in \{\text{Ranjit, Gune}\}[Q(z)=1].\forall y \in \{\text{Ranjit, Gune}\}[Q(y)\rightarrow[y\neq\text{Ranjit}]]\})\})$
came in w

The denotation in (86c) is equivalent to:

(87) $\lambda w.\left\{ \begin{array}{c} \text{Ranjit came in w,} \\ \forall y \in \{\text{Ranjit, Gune}\}[\text{y came in w} \rightarrow y \neq \text{Ranjit}] \end{array} \right\}$, defined iff $\exists x \in \{\text{Ranjit, Gune}\}[\text{x came in w}]$

The question (87) thus presupposes that SOMEBODY came, which is the desired outcome.

6.5 Summary

In this chapter I have argued that *wh*-words in Sinhala, Japanese, Tlingit, and Malayalam should be given a Hamblin-type analysis, under which such elements denote sets of individuals, e.g. $[\![$*kau* "who"$]\!]^g = \{x \in D_e \mid x \in \text{human}'\}$. I then treat Q-particles like Sinhala *də* as denoting variables over choice functions (where a choice function is a function that, when applied to a non-empty set, returns a member of that set), an analysis motivated in part by the fact that Sinhala *wh*-words are able to scope out of islands—so long as there are no island barriers between the Q-particle and the clause over which the *wh*-word takes scope. I assume that these choice function variables are, in interrogatives, bound by the denotation of the interrogative complementiser. This allows us to explain the behaviour of *wh*-words and Q-particles with respect to islands. I have shown also that an Hamblin-style analysis of *wh*-words is preferable to treating *wh*-words as lacking ordinary semantic values (as Beck 2006, Cable 2007) on both theoretical and empirical grounds.

[17] The focus semantic value of the IP would be:

(i) $[\![IP']\!]^{g,F} =$
$\{\lambda w'.f_\alpha(\{\text{Ranjit}, f_\beta(\{\lambda Q:\exists z \in \{\text{Ranjit, Gune}\}[Q(z)=1].\forall y \in \{\text{Ranjit, Gune}\}[Q(y)\rightarrow[y\neq\text{Ranjit}]]\})\})$
came in $w' \mid f_\alpha, f_\beta \in D_{cf}\}$

This chapter also provided complete derivations for (1) "normal" Sinhala *wh*-questions where the verb takes the focussing *-e* ending, (2) *wh*-questions which occur in special circumstances where the verb takes the "neutral" *-a* ending, (3) alternative questions, and (4) neutral and (5) focussed yes/no questions.

I analyse the *-e* morpheme which appears as a suffix on the verb as an element residing in the head of FocusP. Normal *wh*-questions, where the verb appears with the *-e* ending, involve the presence of an existence presupposition (carried by the *-e* morpheme). I assume that the Q-particle *də* in such constructions bears focus—an assumption which correctly predicts both the nature of the existence presupposition and the possibility of the QP moving overtly to SpecFocusP.

Under special circumstances—in the case of the *wh*-word being *kiidenek* "how many" or where the *wh*-word occupies an embedded clause which is the complement of a certain set of verbs including *dannəwa* "know"—the verb may optionally appear with the "neutral" *-a* ending rather than the *-e* focus-associated ending, in which case the question carries no existence presupposition. The absence of an existence presupposition in such questions is correctly predicted if we assume that the existence presupposition is carried by the morpheme *-e*—which is absent in such constructions.

In alternative questions the verb appears in the *-e* form and presupposes that the proposition is true for at least one of the alternatives. This again can be correctly predicted by positing that the existence presupposition is carried by the *-e* morpheme.

I argue that yes/no-questions can be treated as a special subtype of alternative question, with a hidden *or not* element. In "neutral" yes/no-questions, where the verb appears in the *-a* form, the disjunction takes place at the level of the IP, and there is no existence presupposition. In "focussed" yes/no-questions I also posit a hidden *or not* element, but in this case the disjunction occurs at a level lower than that of the IP, and thus the negation involved is constituent negation (on which see above, Chapter 5). The presupposition of the constituent negation projects up to the CP, resulting in the correct prediction that such questions presuppose the existence of some individual which satisfies the proposition.

Here we see that the analysis of syntax and semantics of focus, developed in Chapters 4, 5, is a necessary precondition for a complete understanding of the semantics of interrogatives in Sinhala.

In the following chapter I turn to one of the non-interrogative uses of Q-particles which has yet to be discussed in any detail, namely the appearance of Q-particles in the formation of *wh*-based indefinites. Crosslinguistically, indefinites thus formed display a tendency to be epistemic indefinites, which signal lack of knowledge concerning who or what satisfies the existential claim. The formal analysis proposed for these "pragmatic signals" will also turn out to be crucial for the account of the distribution of *də* and *hari* in disjunctive contexts.

Chapter 7

The semantics of Q-particles in indefinites in Sinhala & other languages

> '. . . So I found myself halfway between the perception of the concept "horse" and the knowledge of an individual horse. And in any case, what I knew of the universal horse had been given me by those traces, which were singular. I could say that I was caught at that moment between the singularity of the traces and my ignorance, which assumed the quite diaphanous form of a universal idea. If you see something from a distance, and you do not understand what it is, you will be content with defining it as a body of some dimension.. . . '
>
> —William of Baskerville to his pupil Adso of Melk,
> in Umberto Eco's *The name of the rose*

7.1 *Wh*-based indefinites using Q-particles

In all four of the languages which form the major basis for this study, to wit Sinhala, Japanese, Malayalam, and Tlingit, we find indefinites formed from *wh*-words combined with Q-particles (for an overview, see Chapter 2). In three of these languages, Sinhala, Japanese, and Malayalam, indefinites formed in this way are clearly epistemic indefinites, that is indefinites which explicitly signal a lack of further information about who or what satisfies the existential claim. In this chapter, I focus primarily on Sinhala, which presents a complex distribution involving two pragmatically-distinct types of epistemic indefinites.[1]

7.2 How are the Q-particles of *wh*-indefinites bound?

Both Hagstrom (1998) and Cable (2007) treat Q-particles like Sinhala *də* and Japanese *ka* as denoting choice functions, but Hagstrom (1998) suggests that Q-particles have inherent quantificational force, while for Cable (2007) Q-particles are bound either by existential closure (in the case of *wh*-indefinites) or derive their quantificational force from the denotation of the interrogative complementiser.

Hagstrom (1998: 129–134) argues that *wh*-indefinite pronouns in Japanese do not act like Heimian indefinites, that it, they exhibit no quantificational variability, and cannot "pick up" quantificational force from their environment, contrast (1) with (2).

(1) If an article$_i$ is published in *Linguistic Inquiry*, John **usually** reads it$_i$.
(≡ John reads **most** articles that are published in *LI*.)

[1] This chapter treats the same basic material as Slade (under review).

(2) *MIT Press-ga nani-ka$_i$-o syuppansureba John-ga **taitei** sore$_i$-o yomu.
MIT Press-NOM what-*ka*$_i$-ACC published-if John-NOM **usually** it$_i$-ACC read.
"If something$_i$ is published by MIT Press, John usually reads it$_i$."

Japanese bare-NP type indefinites, on the other hand, do behave like Heimian indefinites, as shown by example (3):

(3) MIT Press-ga ronbun$_i$-o syuppansureba John-ga **taitei** sore$_i$-o yomu.
MIT Press-NOM article$_i$-ACC published-if John-NOM **usually** it$_i$-ACC read.
"If MIT Press publishes an article$_i$, John usually reads it$_i$."

Note that in English, *if. . . then* conditionals, an element with inherent quantificational force in the protasis cannot be co-indexed with a pronoun in the apodosis, as shown by (4).

(4) *If everything$_i$ (submitted) is published in *LI*, John (usually) reads it$_i$.

Here *everything* is trapped inside of the *if* clause (a strong island) and thus cannot raise to a position from which it can c-command and thus bind the pronoun *it* in the apodosis.

Example (2) would thus seem to indicate that Japanese *wh*-indefinites do have their own quantificational force, just like English quantificational pronouns like *everything* (see example (4)). However, Saito (1998) and Hagstrom (1998: 132n14) note that if the case-marker *-o* is omitted from the *wh*-indefinite in example (2) then the example becomes grammatical; Hagstrom (1998: 132n15) further notes that some speakers find (2) perfectly grammatical even when the *wh*-indefinite bears an overt case-marker. Thus, it is not clear to me that there is sound evidence against treating Japanese *ka*-indefinites as 'Heimian indefinites' (see Yatsushiro (2009), who reports different grammaticality judgements and who thus argues that Japanese *wh*-indefinites do not have inherent quantificational force).

The situation in Sinhala is also complicated. While *də*-indefinites are ungrammatical in *if. . . then* constructions where the indefinite in the protasis is to be interpreted as binding a pronoun in the apodosis (6), *hari*-indefinites are perfectably grammatical in the same context, and thus appear to act like 'Heimian indefinites' (5).

(5) sāmānyayen kauru$_1$ **hari** maṭə īmel ekak evvot, mamə eyāṭa$_1$ uttarayak yavanavā
usually who$_1$ ***hari*** me.DAT e-mail one.INANIM.INDEF send.COND, I him$_1$ reply.INANIM.INDEF send.PRES
"Usually, if someone$_1$ sends me an e-mail, I send him$_1$ a response."

(6) *sāmānyayen kau$_1$ **də** maṭə īmel ekak evvot, mamə eyāṭa$_1$ uttarayak yavanavā
usually who$_1$ ***də*** me.DAT e-mail one.INANIM.INDEF send.COND, I him$_1$ reply.INANIM.INDEF send.PRES

However, it appears that *də*-indefinites are simply ungrammatical in *if. . . then* constructions, even where there is no element in the apodosis which must be bound by the indefinite in the protasis; as shown by the ungrammaticality of (8) (which, again, is grammatical if *hari* is employed rather than *də*, as shown by (7)).

(7) kauru **hari** maha rāə maṭə ṭeləpōn kalot nam, maṭə taraha yanava.
who ***hari*** much night me.DAT telephone do.COND COND.PTCP, me.DAT anger come.PRES
"If someone calls me on the phone in the middle of the night, I get angry."

(8) *kau **də** maha rāə maṭə ṭeləpōn kalot nam, maṭə taraha yanava.
who ***də*** much night me.DAT telephone do.COND COND.PTCP, me.DAT anger come.PRES

This suggests that the ungrammaticality of (6) is due to some incompatibility of *də* with conditionals, rather than evidence that *də*-indefinites have inherent quantificational force.

I will thus assume in the remainder of this study that Sinhala *wh*-indefinites bear no inherent quantificational force, and therefore that particles like Sinhala *də* are simple variables over choice-functions, bound by existential closure.

7.3 Indefinites signalling various levels of ignorance

In Sinhala we find three pragmatically-distinct indefinites: the NP-type indefinite of (9a); and two types of *wh*-based indefinites, the first having the form of *wh*-word + the particle *hari*, as in (9b); the second having the form *wh*-word + the particle *də*, as in (9c).[2,3]

(9) a. sanat deyak gatta.
Sanath thing.INDEF buy.PAST.A

b. sanat monəva hari gatta.
Sanath what *hari* buy.PAST.A

c. sanat monəva də gatta.
Sanath what *də* buy.PAST.A
"Sanath bought something."

The sentences in (9a)–(9c) all have the same truth-conditional semantics, but they differ in terms of the pragmatic conditions in which they are felicitous. The two *wh*-based indefinites (9b) and (9c) both pragmatically signal that the identity of the person in question is unknown.[4] The NP-type of (9a) is pragmatically-unspecified/'neutral', providing no information about whether or not the identity of the person in question is known; morphologically it involves a NP with an indefinite suffix (*-ek*). However, even the two *wh*-based indefinites involve different felicity conditions, discussed in detail below.

There appears to be a crosslinguistic tendency for *wh*-based indefinite pronouns to contrast with NP-based indefinite pronouns in just this way. Consider the Japanese examples in (10) (cf. Moore 2003).

(10) a. Hito-kara denwa atta
person-from phone exist.PAST
"There was a telephone call from someone."

b. Dare-ka-kara denwa atta.
who-*ka*-from phone exist.PAST.
"There was a telephone call from someone (unknown)."

[2]There is actually a fourth type of indefinite involving *wh*-word + the particle *vat*, which appears only in NPI contexts. Since *vat* appears to simply be the NPI-counterpart of *hari* I do not provide separate discussion.

[3]The alternation between *kauru* in (9b) and *kau* in (9c) is purely allomorphic: "who" appears as *kauru* unless immediately followed by *də*, in which case it appears as *kau* (see, amongst others, Fairbanks et al. 1968, Lalith Ananda 2008).

[4]Typically, these types of indefinites signal that the identity of the person or thing in question is unknown TO THE SPEAKER, but not necessarily, see further below at footnote 5.

Here again the NP-type indefinite of (10a) is pragmatically-unspecified with respect to the presence/absence of information about who or what satisfies the existential claim, whereas the use of the *dare* "who" plus the particle *ka* in (10b) signals that the identity of the person or thing in question is not known.[5]

In Malayalam too we find the same contrast between NP-based and *wh*-based indefinites, as shown in (11) below (cp. Jayaseelan 2001: 66), where again the *wh*-based indefinite in (11a) carries a pragmatic signal that the identity of the person/thing in question is unknown, a signal not carried by the NP-based indefinite of (11b).

(11) a. ñaan̲ innale aar-e-(y)oo paricayappeṭṭu
I yesterday who-ACC-*oo* met
"I met somebody (unknown) yesterday."

b. ñaan̲ innale oru aaḷ-e paricayappeṭṭu
I yesterday one person-ACC met
"I met a person/somebody yesterday."

The contrast shown by Malayalam is found in other Dravidian languages as well, such as Kannada (see Bhat 1981). Haspelmath (1997: 45–7) provides further examples of *wh*-based indefinites which signal that the identity of the referent of the indefinite is unknown in Russian (see also Geist & Gáspár 2007) and Lithuanian.

What is interesting, in the case of Sinhala as shown in Table 7.1, is that the two *wh*-based indefinites (see above examples (9b), (9c))—though they both signal that the referent's identity is unknown—differ subtly in their pragmatics.[6] As a first approximation, my Sinhala consultants describe *də*-indefinites as somehow "more unknown" than their *hari*-counterparts. Consider again the examples in (9a)–(9c), repeated below in (12a)–(12c) with slightly different translations.

(12) a. sanat deyak gatta.
Sanath thing.INDEF buy.PAST.A
"Sanath bought a thing."

b. sanat monəva hari gatta.
Sanath what *hari* buy.PAST.A
"Sanath bought some thing."

c. sanat monəva də gatta.
Sanath what *də* buy.PAST.A
"Sanath bought some thing-or-other."

Now consider the following scenarios:

[5] Again, usually with respect to the speaker's knowledge, but not necessarily. So (10b) can be continued either as in (ia) or (ib) (cp. Haspelmath 1997: 312, Moore 2003: 608).

(i) a. . . . kedo dare-kara da ka wakar-anai
. . . though who-*ka* COP *ka* know-NEG
". . . I don't know who it was from."

b. . . . kedo dare-kara da ka atete goran
. . . though who-*ka* COP guessing try.IMPV
". . . try to guess who it was from!"

In (ia) it is the speaker who lacks knowledge of the identity of the person/thing in question; in (ib) it is the addressee who (is presumed) to lack knowledge about who or what satisfies the existential claim.

[6] Alonso-Ovalle & Menéndez-Benito (2003); Alonso-Ovalle (2006) discuss a similar pragmatic differences between English *some* NP and Spanish *algún* NP.

(13) a. I saw Sanath buy Jean-Baptiste Greuze's painting *The White Hat*.
b. I saw Sanath buy some piece of artwork, but I don't really know what it was. (I.e., I might be able to describe it, but I don't know what it's called, who painted it, etc.)
c. Sanath told me that he bought something, but I have no direct experience of the event.

Table 7.1 indicates which utterances from (12a)–(12c) are felicitous in contexts (13a)–(13c).

	(12a) is felicitous	**(12b) is felicitous**	**(12c) is felicitous**
In scenario (13a)	✓		
In scenario (13b)	✓	✓	
In scenario (13c)	✓	?[7]	✓

Table 7.1: Felicity conditions for indefinite constructions in Sinhala

The distinction between WH+*hari* and WH+*də* is reminiscent of differences between *some* NP in English and *algún* NP in Spanish. English *some* NP also signals that the speaker lacks further knowledge about the identity of the referent, as shown by the contrast between *a philosopher* and *some philosopher* in examples (14) and (15) below.

(14) John, Bill, and Bill's sister, Mary—who are all linguists—are attending a party where all of the other guests are philosophers. All three of the linguists share a great disdain for philosophers and are only attending the party because there's an open bar. They don't know any of the philosophers by name, and don't plan to get to know them either. John spots Mary kissing one of the philosophers. Shocked, he tugs on Bill's coat-sleeve and exclaims:
a. Look, your sister is kissing a philosopher!
b. Look, your sister is kissing some philosopher!

Here John may felicitously utter either (14a)—since *a(n)* NP is a pragmatically-unmarked indefinite—or (14b), with the pragmatically-marked *some* NP indefinite since the identity of the philosopher is unknown to John (and Bill).

(15) John, Bill, and Bill's sister, Mary—who are all linguists—are attending a party where all of the other guests are philosophers. All three of the linguists share a great disdain for philosophers and are only attending the party because there's an open bar. They don't know any of the philosophers by name—and don't plan to get to know them either—except for one: George, whose office happens to be next door to the office John and Bill share. John spots Mary kissing George. Shocked, he tugs on Bill's coat-sleeve and exclaims:
a. Look, your sister is kissing a philosopher!
b. #Look, your sister is kissing some philosopher!

With regard to the scenario in (15), John may felicitously utter (15a)—since *a(n)* NP is a pragmatically-unmarked indefinite which can be used regardless of whether the identity of the person in question is known or not, but not (15b), since the identity of the philosopher Mary kissed is known to John (and Bill).

[7]Since *hari*-indefinites are "less unknown" then their *də* counterparts, we might expect that *hari*-indefinites should be felicitous wherever *də* indefinites are. However, my Sinhala consultants appear to disfavour *hari* indefinites in contexts like (13c). This may be the result of a sort of "blocking effect", where the appropriateness of a *də* indefinite in a context "blocks" a *hari* indefinite from occurring in that context.
In fact, *hari* and *də* indefinites show additional differences beyond the level of ignorance that they convey, and it may be one of the other properties of *hari* which leads to speakers disfavouring it in contexts like (13c).
One additional way in which the two types of Sinhala differ is that *də* indefinites are "specific", while *hari* indefinites are "non-specific", as discussed in Section 7.8. In this study I am concerned, however, primarily with the differences between *hari* and *də* indefinites which involve the ignorance component, and I leave the examination of the differences in specificity and other properties for future study.

Alonso-Ovalle & Menéndez-Benito (2003) discuss the properties of Spanish *algún. Algún* resembles English *some* in its behaviour in certain contexts: like English *some*, *algún* signals that the speaker has no further information about the identity of person or thing in question. Thus, just as the identity of an *a(n)* NP indefinite may be queried (16) but not that of a *some* NP indefinite (17) in English, so too the identity of a *un(a)* NP indefinite may be further queried (18) but not that of an *algún* NP indefinite (19) in Spanish.

(16) a. A: A cabinet minister has been shot.
b. B: Who?

(17) a. A: Some cabinet minister has been shot. (Strawson 1974)
b. B: #Who?

(18) a. A: María está tomando **una** clase de lingüística
Mary is taking **a** class of linguistics
b. B: ¿Cuál?
Which one?

(19) a. A: María está tomando **alguna** clase de lingüística
Mary is taking ***algún*** class of linguistics
b. B: #¿Cuál?
Which one?

However, *some* differs from *algún* in that *algún* is only possible when the speaker cannot identify the person in question in any way (by name or descriptively), while *some* is possible when the speaker could visually identify the person in question (e.g. pick him out of a police line-up) but not identify him by name. Thus in the context that the speaker sees a man—who he knows to be a professor but whose name he does not know—dancing on a table, the following sentence is felicitous in English:

(20) Look! Some professor is dancing the lambada on his table.

But *algún* cannot be used in this context:[8]

(21) #¡Mira! **algún** profesor está bailando la lambada encima de la mesa
Look! ***algún*** professor is dancing the lambada on the table
(Alonso-Ovalle & Menéndez-Benito 2010)

Note that the pragmatic contexts in which English *some* occurs are similar to those appropriate to Sinhala *hari* indefinites and those in which Spanish *algún* occurs are similar to those appropriate to Sinhala *də* indefinites. Thus the Sinhala distinction between WH-*hari* and WH-*də* has parallels elsewhere, but the question remains of how to formalise this distinction.

In the next section, Section 7.4, I summarise the formalisation Alonso-Ovalle & Menéndez-Benito (2010) provide for Spanish *algún*, pointing out that it does not allow for a distinction between English *some* and Spanish *algún*, nor between Sinhala WH-*hari* and WH-*də*. This study argues that treating Q-particles as denoting variables over choice functions allows for a unified semantic analysis across the various syntactic contexts in which Q-particles appear, and in Section 7.5 I discuss the pragmatic distinctions between *wh+hari* and *wh+də* indefinites in Sinhala and propose a formalisation which relies on the notions of intensional choice functions (Romero 1999) and important predicates. Section 7.5.1 provides an overview of the distinction between the two epistemic indefinites; Section 7.5.2 discusses

[8]Though it could be used, for instance, if the speaker was told by a third person that a professor is dancing the lambada on the table.

what it really means to "know" who or what someone or something is and introduces the notion of "Important Predicates"; Section 7.5.3 suggests how Important Predicates and intensional choice functions may be related, and provides a formalisation of the pragmatics of *wh+ɦari* and *wh+dɔ* indefinites in terms of intensional choice functions. Section 7.5.4 demonstrates that the proposed denotations correctly predict the environments in which the two types of epistemic indefinites are felicitous. In Section 7.6 I argue that the pragmatic "signal" of unknown indefinites can be treated as a presupposition, *pace* Alonso-Ovalle & Menéndez-Benito (2010). Section 7.8 provides a summary and concluding remarks.

7.4 Spanish *algún*: an anti-singleton analysis

Alonso-Ovalle & Menéndez-Benito (2010) compare Spanish *algún* to German *irgendein*, which Kratzer & Shimoyama (2002) analyse as a domain widener; in contrast, Alonso-Ovalle & Menéndez-Benito (2010) take *algún* to be an 'anti-singleton' indefinite.

Alonso-Ovalle & Menéndez-Benito (2010) note the similarity of the ignorance component of *algún* to the ignorance effect triggered by Hindi *-bhī* correlatives and English *-ever* free relatives (Dayal 1997; von Fintel 2000b).

(22) jo bhī laṛkī mehnat kar rahī hai vah safal hogī
RP *bhī* girl effort make CONT is he/she successful be.SG.FEM.FUT
"Whichever girl is making an effort will be successful." (Dayal 1997: 9)

(23) There's a lot of garlic in whatever (it is that) Arlo is cooking. (von Fintel 2000b)

Building on Dayal (1997), von Fintel (2000b) provides an analysis for the ignorance component of *whatever*: where F is the modal base for *whatever*, a set of worlds (usually the set of worlds epistemically-accessible to the speaker):[9]

(24) *whatever*(w)(F)(P)
a. denotes: ιx.P(w)(x)
b. presupposes: $\exists w', w'' \in F$: ιx. $P(w')(x) \neq$ ιx.$P(w'')(x)$

Given the translation rule in (24) a sentence like (23) presupposes that the thing that Arlo is cooking is not the same in all of the worlds in the modal base (e.g., not the same in all of the worlds epistemically-accessible to the speaker, if the modal base F is that of the speaker).

Based on the fact that both *whatever* and *algún* involve an 'ignorance component', Alonso-Ovalle & Menéndez-Benito (2010) consider the possibility that Spanish *algún* might involve a presupposition of a similar sort. Given that *algún* is an indefinite rather than a definite like *whatever*, they suggest the possible formulation show in (25).

(25) *algún*(w)(F)(P)(Q), where F is a modal base, and P and Q are predicates,
a. denotes: $\{x \mid P(w)(x) \,\&\, Q(w)(x)\} \neq \emptyset$
b. presupposes: $\exists w', w'' \in F$: $\{x \mid P(w')(x) \,\&\, Q(w')(x)\}$
$\neq \{x \mid P(w'')(x) \,\&\, Q(w'')(x)\}$

However, Alonso-Ovalle & Menéndez-Benito (2010: 12) point out that this does not seem to be the correct analysis for *algún*, based on the felicity of sentences like (26).

[9]The indifference reading of *whatever* is also discussed by von Fintel (2000b), but I am not concerned with this reading here.

(26) No es verdad que Juan salga con **alguna** chica del departamento de lingüística
not is true that Juan goes-out with ***algún*** girl from the department of linguistics
"Juan is not dating any of the girls in the linguistics department."

Alonso-Ovalle & Menéndez-Benito (2010) assume that assertions involve a covert assertoric operator (which they represent as □) which quantifies over all worlds in the modal base. Alonso-Ovalle & Menéndez-Benito (2010: 7) posit that this assertoric operator occupies the highest LF position. In (27) and following examples, I spell this assertoric operator out as ∀w ∈ F, for the sake of clarity. The assertion and presupposition of (26) is provided below in (27).

(27) a. asserts: ∀w ∈ F.¬∃x[girl-from-linguistics-dept(w, x)].date(w, x, Juan)
b. presupposes: ∃w′, w″ ∈ F: {x | girl-from-linguistic-dept(w′, x) & date(w′, x, Juan)} ≠
{x | girl-from-linguistic-dept(w″, x) & date(w″, x, Juan)}

If the ignorance component of *algún* were a presupposition it should be able to project up to the matrix level in (26), which would result in a contradiction: it would presuppose that the set of girls that Juan is dating is not the same in all worlds in the modal base (assuming here that the modal base is the set of worlds epistemically-accessible to the speaker), while asserting that the set of girls that Juan is dating is in fact the same in all of the worlds epistemically-accessible to the speaker (namely, the empty set). But this contradiction does not arise; (26) simply means that Juan is not dating any girl in the linguistics department.[10]

Alonso-Ovalle & Menéndez-Benito (2010) also show that the ignorance component of *algún* cannot be a conventional implicature[11] since, as discussed above, the ignorance component of *algún*—like that of *whatever*—is not necessarily relative to the modal base of the speaker, and, furthermore, is cancellable (see (28)).

The analysis which Alonso-Ovalle & Menéndez-Benito (2010) arrive at is that the ignorance component of *algún* is a conversational implicature, given that (a) it is cancellable, as shown by (28), and (b) it disappears under negation (and under downward entailing operators more generally), as shown by (26), repeated below as (29).

(28) María se casó con **algún** estudiante de lingüística. De hecho, sé exactamente con quién.
María SE marry.3SG.PAST with ***algún*** student of linguistics. In fact, I know exactly with whom.
"María married a linguistics student. In fact, I know exactly who."

(29) No es verdad que Juan salga con **alguna** chica del departamento de lingüística
not is true that Juan goes-out with ***algún*** girl from the department of linguistics
"Juan is not dating any of the girls in the linguistics department."

Alonso-Ovalle & Menéndez-Benito (2010) argue that this conversational implicature arises from the fact that *algún* imposes an anti-singleton constraint. That is, they argue that the translation of *algún* is that of (30); whereas that of *un* is that of (31) (identical to (30) except that it carries no anti-singleton presupposition).[12]

[10] Even if we allow that presuppositions can be accomodated locally, this does not save the presupposition analysis, since in that case (26) would signal that the speaker knows which girl in the linguistics department Juan is dating, which would still result in a contradiction since the speaker asserts that Juan is not dating any girl in the linguistics department.

[11] In the sense of Potts (2007), that is, a non-deniable, lexical component which is invariably speaker-oriented. Potts (2007) suggests that appositive expressions involve a conventional implicature component.

(i) a. Sheila says that Chuck, a confirmed psychopath, is fit to watch the kids.
b. Shelia believes that Chuck, a psychopath, should be locked up. # But Chuck is not a psychopath.

The invariant speaker-oriented nature of conventional implicatures means that in (ia) it is the speaker, and not Sheila, who is committed to the claim that Chuck is a psychopath, which also accounts for the infelicity of (ib).

[12] In (30) and (31), the function f is a subset selection function (von Fintel 2000a).

(30) $[\![\text{algún}]\!]^g$

a. denotes: $\lambda f_{\langle et,\, et\rangle}\lambda P_{\langle e,\, t\rangle}\lambda Q_{\langle e,\, t\rangle}.\exists x[f(P)(x)\ \&\ Q(x)]$

b. presupposes: $|f(P)| > 1$

(31) $[\![\text{un}]\!]^g$

a. denotes: $\lambda f_{\langle et,\, et\rangle}\lambda P_{\langle e,\, t\rangle}\lambda Q_{\langle e,\, t\rangle}.\exists x[f(P)(x)\ \&\ Q(x)]$

Thus a sentence like (32) can be translated as in (33).

(32) María se casó con **algún** estudiante de lingüística
María SE marry.3SG.PAST with ***algún*** student of linguistics
"María married a linguistics student."

(33) a. asserts: $\forall w \in F.\exists x[x \in f(\text{student}')$ in w & marry(x,María) in w$]$

b. presupposes: $|f(\text{student}')| > 1$

The conversational implicature arises from the fact that the anti-singleton presupposition might cause the hearer to infer that *algún* is employed rather than *un* because the speaker either (a) wants to avoid making a false claim or (b) wants to avoid a false exhaustivity inference.

It is not, however, clear to me how the anti-singleton analysis distinguishes *algún* from *some*, since the latter would presumably have to be analysed as imposing an anti-singleton constraint. *Some* (like Sinhala *wh+hari* indefinites) differs from *algún*, as noted by Alonso-Ovalle & Menéndez-Benito (2003), in that it allows that the referent might be able to be uniquely identified via a definite description, e.g. *the Persian man who studies Basque semantics*. This distinction cannot be captured by the formalisation of the ignorance component simply as an anti-singleton constraint.[13]

Given that there is no clear way of extending the anti-singleton analysis proposed by Alonso-Ovalle & Menéndez-Benito (2010) to epistemic indefinites like English *some X* or Sinhala WH+*hari*, some alternative account of epistemic indefinites is highly desirable. Since the denotation of Q-particles like Sinhala *də* and *hari* has been argued in this study to be best analysed in terms of choice functions, it makes sense to consider an analysis of epistemic indefinites which makes use of the machinery of choice functions. In the next section I propose a formal analysis for the ignorance components borne by epistemic indefinites like Sinhala WH+*hari*, WH+*də* which is based on the treatment of Q-particles like *də* and *hari* as variables over choice functions.

7.5 Extensionally- and intensionally-unknown indefinites

Given the independent evidence favouring a choice-functional analysis of particles like *hari* and *də*, it makes sense to explore the possibility that the "ignorance component" of WH+*hari* and WH+*də* indefinites can also be formulated

[13] As observed in note 7, *də* indefinites, which resemble Spanish *algún* NP in terms of the level of ignorance they convey, are "specific". If we were to adopt the singleton indefinite analysis of specific indefinites proposed by Schwarzschild (2002), then it is obvious that it would thus not be possible to extend the anti-singleton analysis of Alonso-Ovalle & Menéndez-Benito (2010) to Sinhala *də* indefinites.

Whatever analysis of specific indefinites is adopted, however, this seems to be one aspect in which Sinhala *də* indefinites in fact differ from Spanish *algún NP* constructions, since Alonso-Ovalle & Menéndez-Benito (2010) provide examples where *algún* can trigger the interpretation that the speaker does not know how many individuals satisfy his existential claim, as in:

(i) Hay **alguna** mosca en la sopa.
There is *algún* fly in the soup
(Alonso-Ovalle & Menéndez-Benito 2010: 24)

Where (i) conveys that there is at least one fly (but possibly more) in the soup.

in terms of choice-functions. Specifically I suggest that sentences containing epistemic indefinites, such as Sinhala WH+*hari* and WH+*də*, assert the existence of a choice function which, when applied to the *wh*-pronoun (denoting a set of individuals), selects an individual which satisfies the existential claim of the speaker, but which pragmatically signals that the speaker[14] lacks certain information which would enable him to uniquely identify which individual satisfies the existential claim—and thus the speaker cannot identify a single choice function for which the existential claim is uniquely true.

It remains to be determined how exactly epistemic indefinites of the types represented by Sinhala WH+*hari* and WH+*də* can be formally distinguished. In the following subsection I shall suggest that the former type can be described as "extensionally-unknown" and the latter as "intensionally-unknown". "Extensionally-unknown" indefinites are felicitous where the speaker has no means of uniquely identifying an extension which satisfies his existential claim, and "intensionally-unknown" indefinites are felicitous where the speaker cannot even uniquely identify an individual concept which satisfies his existential claim.

7.5.1 An initial characterisation

In order to illustrate the difference between "extensionally-" and "intensionally-"unknown indefinites and the contexts in which these two types of indefinites are appropriate, I discuss the different pragmatic contexts under which we might find the English sentences in (34).

As suggested earlier, English *some NP* behaves similarly to Sinhala WH+*hari* in terms of the degree of ignorance signalled (see Section 7.3), and English *some NP-or-other* behaves similarly to Sinhala WH+*də* (see footnote 7). Assuming that these equivalences are valid, at least in terms of the ignorance components involved, consider the sentences in (34).

(34) a. John is kissing a girl.
b. John is kissing some girl.
c. John is kissing some girl-or-other.

Example (34a) is neutral, in the sense that the indefinite does not convey any additional information about the state of the speaker's knowledge. Example (34b) can be used to signal the speaker's lack of knowledge about the precise identity of the girl that John is kissing. However, it does not signal complete lack of knowledge, as shown by the possible paths of the dialogue in (35).

(35) A: John is kissing some girl.
a. B: #Who? (cf. Strawson 1974)
b. B: Which one?
(i) A: That one over there. [pointing]
(ii) A: The blonde. (in case there is only one pragmatically-salient blonde)
(iii) A: #Sally Bloggs, the daughter of our department head.

So while, A having uttered "John is kissing some girl", B may not ask "Who?" (35a), he may ask "Which one?" (35b); with A's possible replies (35b-i), (35b-ii), demonstrating that A may in fact have additional knowledge which uniquely identifies the girl John is kissing. However, A's possible felicitous replies are limited: response (35b-iii) is pragmatically odd, given A's initial use of *some girl*. The felicity difference between (35b-i), (35b-ii) and (35b-iii) can be explained as follows.

[14] Here and in the remainder of this discussion I assume that the modal base F is that of the speaker.

It is plausible that the extension of the individual concept corresponding to "that one over there" or "the (pragmatically-salient) blonde" is non-identical in those worlds which are epistemically-accessible to speaker A (i.e. possible worlds consistent with speaker A's beliefs about the actual world). On the other hand, it seems likely the individual concept "Sally Bloggs, our department head's daughter" will have the same extension in all worlds epistemically-accessible to speaker A.

The situation is otherwise in case a *some NP-or-other* indefinite is employed, as shown by (36).

(36) A: John is kissing some girl-or-other.

a. B: #Who?

b. B: #Which one?

c. B: Do you know anything about this girl?

(i) A: #Yes, she is the blonde standing over there.

(ii) A: Well, she's in the philosophy department.

(iii) etc.

Here both (36a) and (36b) are infelicitous, since they presuppose that speaker A has some means of uniquely identifying the girl in question. Speaker B could felicitously ask questions along the lines of (36c), since it is still consistent with the pragmatic signal of *some girl-or-other* that the speaker has additional knowledge about the girl John is kissing—so long as this knowledge cannot serve to UNIQUELY identify the girl in question (either extensionally or intensionally).

There remains the question of how to formalise the notions "extensionally-unknown" and "intensionally-unknown" with respect to speaker knowledge. Such a formalisation requires that we refine our notion of what it means to "know" who or what someone or something is, as discussed in the next subsection.

7.5.2 What is it "to know"? – Important Predicates

Though proper names are frequently treated as rigid designators (cf. Kripke 1972, 1980), i.e. as denoting the same individual in all possible worlds, in practice knowing a name is often not sufficient for the purpose of actually KNOWING who or what someone or something is in a meaningful way. I may know that some person named "Aishwariya Rai" exists, without knowing much of anything else about her, such as the fact that she was Miss World in 1994.

Further, one may, in a sense, simultaneously know and not know who some person *N* is. For example, I may know who Noam Chomsky is, in the sense that I know him to be one of the founders of modern generative linguistic theory, to be the author of *Syntactic Structures* etc. But, at the same time, I may be completely ignorant of any of Chomsky's political criticism, not know him to be the author of *Manufacturing Consent* etc. Likewise, I might know who Chomsky is in terms of his contributions to linguistic and/or political theory, but still have no idea what he looks like, where he lives etc.

Or consider the case of Professor Moriarty: in Sir Arthur Conan Doyle's story "The Valley of Fear", the great private detective Sherlock Holmes knows who Moriarty is—in the sense that Holmes knows that Professor Moriarty is the criminal mastermind behind a vast and subtle crime ring, while Inspector MacDonald knows only that Moriarty is a seemingly harmless academic (despite the fact that MacDonald interviews Moriarty in the professor's study where hangs a painting by Jean-Baptiste Greuze—a piece of artwork which, as Holmes points out, Moriarty would not have been able to afford on a professor's salary).

In order to account for these sorts of facts, Boër & Lycan (1975) suggest that "knowing" should be evaluated

relative to some purpose or project.[15] Relative to the purpose of solving crimes, Holmes knows who Moriarty is, while MacDonald does not. With respect to knowing "who's who" in the realm of academia, both Holmes and MacDonald know who Moriarty is. Likewise, a politically disinterested linguist may know "who Chomsky is", without knowing anything about Chomsky's political writings, and a political scientist could know "who Chomsky is", in terms of his political theories, without knowing anything about Chomsky's status in the world of linguistics.

Evaluation of speaker knowledge of a person or thing in terms of a particular purpose or perspective also helps to address one problem regarding the definition of "intensionally-unknown" indefinites, which I have defined as pragmatically signalling that the speaker lacks knowledge of ANY means of uniquely identifying the person in question, including in terms of individual concepts (whose extensions may vary from world to world). Taking English "someone or other" to be an "intensionally-unknown" indefinite:

(37) John met someone-or-other yesterday.

The sentence in (37) should pragmatically signal that the speaker lacks any intensional concept which serves to uniquely identify the person John met yesterday with respect to the speaker's epistemically-accessible worlds. However, one (potentially uniquely-identifying) individual concept is trivially available: namely the individual concept corresponding to "the person John met yesterday".

Obviously, it would be desirable to be able to systematically exclude such individual concepts from consideration: they are parallel to knowing that "Noam Chomsky is Noam Chomsky", which of course is not sufficient grounds for claiming to really know WHO Noam Chomsky is. In any case, we are still left with the question: what does "knowing" who someone is really mean?

Boër & Lycan (1975: 304–306) establish several criteria which they deem desirable for any formal semantic account of "knowing", of which perhaps the most important is that such a theory should provide a general answer to the question of what it is to know who someone is consistent with actual linguistic usage of ⟦know⟧. After considering various approaches, they argue that a crucial component of the answer is the notion of an "important predicate" (Imppred), which they define formally as in (38).

(38) $\text{Imppred}(\phi, \text{N}, \text{P}) \leftrightarrow \exists F\Big[\text{⚠}\big[\ulcorner\lambda x.\phi(x)\urcorner, \lambda x[x{\in}C(P) \,\&\, F(x)]\big]$
$\&\ \big[\lambda x[x{\in}C(P) \,\&\, F(x)]\big]!(N) \,\&\, \big[\exists G[\text{Det}(\lambda x.G(x), N, C(P)) \rightarrow \text{Det}(\lambda x.F(x), \text{in } N, C(P))]\big]\Big]$

That is, an Important Predicate is a predicate ϕ of which N is its sole member relative to identifying N for some specific purpose, project, or perspective P. P defines a category C(P), which can be understood as a set whose membership is fixed by P. Given this definition, knowing who someone is relative to P consists of knowing which member of C(P) he is. A particular individual N may be so identified by means of some property F, which is unique to N. For example, P might be "who is Noam Chomsky (with respect to the field of linguistics)", in which case C(P) would be something like a set of linguists.

That knowing who someone is in this sense means being able to know which member of a category C(P) that person is can be used to rule out trivial cases of "knowing", like that discussed for (37) above. Here, knowing "who" the person John met yesterday was must be considered relative to some purpose, project, or perspective. If the only knowledge the speaker of (37) possesses about the person John met yesterday which could serve to uniquely identify that individual consists of knowing that he was the person John met yesterday, then there are vanishingly few pragmatically likely projects, purposes, or perspectives for which the speaker could be said to "know" who the person John met yesterday is.

[15] To which we might add "perspective" in order to also include non-telic situations (i.e. situations which are not well described as involving a "purpose" or "project").

Boër & Lycan (1975: 329–330) additionally suggest, that in general, F should additionally be related in some way to being a member of C(P). Thus, in our example, one reasonable F might be the property of having written *Syntactic Structures*, since this property is connected to being a person who has authored a linguistic study; as opposed to, say, having written *Manufacturing Consent*, which is not related in any direct way to being a member of the set of people who have authored linguistic studies. Boër & Lycan (1975) refer to Fs which are related to the category C(P) as C(P)-determinables.

As Boër & Lycan (1975: 330–331) point out, certain C(P)s may not have any C(P)-determinables. For instance, if C(P) is a set of blondes, it may be that none of the unique properties of the members of C(P) are directly related to membership in C(P). Identifying N in C(P) may rely on a non-category related property like "having a gold tooth" (which has nothing to do with being blonde).

Finally, Boër & Lycan (1975: 314) adopt from Kripke (1972: 270) the notion of a rigid designator, where "a designator rigidly designates a certain object iff it designates that object wherever [in all possible worlds in which] the object exists". In symbols, $\triangle[\alpha, A]$ means "α rigidly designates A"; that is, where I is the set of possible worlds, $\forall w \in I.[\![\alpha]\!](w)=A$.

With these definitions in place, (38) may be paraphrased as stating that if ϕ is an Important Predicate for N with respect to purpose/perspective P, then there exists a property F such that $\ulcorner\lambda x.\phi(x)\urcorner$ rigidly designates $\lambda x[x \in C(P)$ & $F(x)]$ (i.e. ϕ is a property which is true of some member of C(P)); N is the sole member of F; and if there are any C(P)-determinables, then F is a C(P)-determinable.

Given this definition of Imppred, a speaker S may be said to know who N is if (where K_s means S knows that):

(39) $\exists\phi\exists\beta\left[\text{Imppred}(\phi, N, P) \,\&\, \triangle(\beta, N) \,\&\, K_s\ulcorner\phi!(\beta)\urcorner\right]$

The definition in (39) is relevant where the speaker referentially knows who N is. (39) states that for some predicate ϕ and some individual β, ϕ is an Important Predicate for N with respect to purpose/perspective P and β rigidly designates N and S knows that β is a member of the singular predicate ϕ. Returning to our earlier example, S may be said to know who N, the author of *Syntactic Structures*, is if there is some Important Predicate ϕ—say the property of being the author of *Aspects of the Theory of Syntax*—of which N is a member with respect to purpose/perspective P; and $[\![\text{Chomsky}]\!]$ rigidly designates N, and S knows that $[\![\text{Chomsky}]\!]$ is the unique member of the predicate ϕ.

Instead of knowing referentially who someone is, a speaker may know attributively who someone is (see Donnellan 1966).[16] For example, a detective may know "who the murderer is" referentially (e.g. he may know that "the murderer is Hannibal Lector"), in which case the formula in (39) is relevant. However, the detective might know instead that the murderer is the person whose fingerprints are on the murder weapon. A different logical form is necessary to handle knowing attributively; this is given in (40).

(40) $\exists!x\exists\alpha\exists\psi\exists\phi\left[F(x) \,\&\, \triangle[\alpha, x] \,\&\, \triangle[\ulcorner\lambda y.\psi(y)\urcorner, \lambda y.F(y)] \,\&\, \text{Imppred}(\phi, x, P) \,\&\, K_s\ulcorner\psi(\alpha) \,\&\, \phi!(\alpha)\urcorner\right]$

The formal notation in (40) states that there exists an α, a (unique) x, and two predicates, ϕ and ψ, such that (1) F is a property of x; (2) α rigidly designates x and $\ulcorner\lambda y.\psi(y)\urcorner$ rigidly designates $\lambda y.F(y)$; (3) ϕ is an Important Predicate for x

[16] Donnellan (1966) contrasts "attributive" use of definite descriptions, e.g. in a sentence like (i) below, in additional to the "referential" use of a definite description (ia)—where the particular individual picked out by the description is known to the speaker, there is also an "attributive" use of the definite description (ib), where the particular individual picked out by the description is unknown to the speaker.

(i) Smith's murderer is insane.
  a. Smith's murderer, namely Jones, is insane.
  b. Smith's murderer, whoever he may be, is insane.

See Kripke (1977); Wilson (1991); Abbott (2000) for further discussion. See also Evans (1982), Ludlow & Neale (1991) on descriptions used like demonstratives rather than as names, cp. Russell (1905, 1910–1911).

with respect to purpose/perspective P; and (4) S knows that α is a member of ψ and α is the unique member of ϕ.

In terms of our detective example, the detective might be said to "know who the murderer is" attributively if (1) x is the murderer ("the F"); (2) α rigidly designates x and ψ rigidly designates the property of being the murderer; (3) ϕ is an important predicate for x with respect to purpose/perspective P (say, "identifying the person who committed the murder"), for instance, ϕ might be the property of being the person whose fingerprints are on the murder weapon; and (4) S knows that it is α who committed the murder and that α is the unique individual whose fingerprints are on the murder weapon. In other words, the detective knows that someone committed the murder and whoever that person is, it is his fingerprints and no-one else's on the murder weapon.

7.5.3 Putting it together

Given that β and α have been defined as rigid designators (see the definitions in (39), (40)), in all possible worlds the extension of β is N (if N exists) and the extension of α is x (if x exists). But, by the definition of "Important Predicate" (see (38) above), $\ulcorner\lambda x.\phi(x)\urcorner$ also rigidly designates a particular property F such that N is the unique element which satisfies F. This seems to imply that a predicate can be an Important Predicate for some N in all possible worlds. But since unique properties of individuals can be restricted to a subset of possible worlds, it seems preferable to relativise Important Predicates to particular worlds, as in (41).[17]

(41) $\text{Imppred}(\phi, \text{N}, \text{P}, \text{w}) \leftrightarrow \exists F\Big[\text{A}\big[\ulcorner\lambda x.\phi(x)\urcorner, \lambda x[x{\in}\text{C(P)(w)} \;\&\; \text{F(w)(x)}]\big] \;\&\; \big[\lambda x[x{\in}\text{C(P)(w)} \;\&\; \text{F(w)(x)}]\big]!(\text{N}) \;\&\; \big[\exists \text{G}[\text{Det}(\lambda x.\text{G(w)(x)}, \text{N}, \text{C(P)(w)}) \rightarrow \text{Det}(\lambda x.\text{F(w)(x)}, \text{N}, \text{C(P)(w)})]\big]\Big]$

Then, we can say that a speaker knows who or what some individual is if he knows that (42) (for the referential type of knowing; the attributive type of knowing could be similarly modified).

(42) $\lambda w.\exists\phi\exists\beta\big[\text{Imppred}(\phi, \text{N}, \text{P}, \text{w}) \;\&\; \text{A}(\beta, \text{N}) \;\&\; \text{K}_{s_{(w)}}\ulcorner\phi!(\beta)\urcorner\big]$

In another words, if there is some Important Predicate ϕ for N w.r.t. purpose/perspective P in w, and β rigidly designates N, and S knows in world w that β is the unique individual satisfying the predicate ϕ, then S can be said to know who or what N is.

Thus if (42) (or the relativised version of (40)) is true, then the speaker S cannot be said to be ignorant of the identity of the person or object in question. Therefore, if either (39) or (40) is true, then the speaker's use of either an "extensionally-unknown" or an "intensionally-unknown" indefinite will be pragmatically infelicitous. Specifically, S's knowledge that $\ulcorner\phi!(\beta)\urcorner$ or $\ulcorner\psi(\alpha) \;\&\; \phi!(\alpha)\urcorner$ guarantee that the speaker can identify the person or object in question intensionally. Adding to this the fact that β rigidly designates N and α rigidly designates x guarantees that the speaker can identify the person or object in question extensionally as well, since the extensions of β will be constant in all possible worlds, as will the extensions of α. How can we connect the definitions of "knowing" from Boër & Lycan (1975) with a choice functional analysis of epistemic indefinites?

[17] C(P)(w) denotes the set of individuals established by purpose/perspective P relative to world w.

Firstly, we need choice functions which allow for different individuals to be selected in different possible worlds. This type of choice function has been proposed in Romero 1999:[18]

(43) **Basic intensional choice function**:
A function f $\in D_{\langle\langle\langle s,\tau\rangle,t\rangle\langle s,\tau\rangle\rangle}$ is a basic intensional choice function (BASICH(f)) iff for all Q.Q$\neq\emptyset$ in the domain of f: f(Q) $\in$ Q. (adapted from Romero 1999)

Intensional choice functions then operate over intensions rather than extensions, which permits the possibility that the member of Q which a choice function f selects may have different extensions in different possible worlds. This allows for the possibility that the individual chosen may vary from world to world, as desired.

Secondly, there exists an intuitive connection between Important Predicates (as defined above) and choice functions: both "choose" a unique individual from a set. In order to be able to properly characterise epistemic indefinites, this intuitive connection must be more precisely formalised. Therefore, we can define the notion of basic intensional choice functions which APPROXIMATE epistemically-accessible Important Predicates, where "approximate" should be understood as meaning that there exists some Important Predicate ϕ such that in all relevant worlds the choice function picks out an individual concept whose extension is the same as that of the unique individual which satisfies the Important Predicate ϕ.[19] Formally, this can be expressed as follows:

[18]Romero's account was developed to provide the correct characterisation of the semantics of *which* phrases which require intensional answers, such as:

(i) Which pet of his$_1$ did at most one boy$_1$ mistreat?

where an extensional choice function fails to properly deal with the local presupposition accommodation readings of possible answers. The definition of BASIC intensional choice functions allows for the correct predictions regarding *which* phrases which are "transparent" with respect to whether the extension of an individual concept has a particular property in a particular world. This is relevant for the following type of contexts:

(ii) Scenario for (iii):
Petra, an innocent four-year-old who sympathises with old dogs, wants every friend of hers to play with the oldest dog in that friend's neighbourhood, whatever dog that may be. Unbeknownst to her, all of the dogs, cats, and monkeys in the surrounding neighbourhoods have contracted rabies.

(iii) a. Q: Which animal that may give him$_i$ rabies does Petra want every friend$_i$ of hers to play with?
b. A: The oldest dog in his$_i$ neighbourhood—whichever dog that is. (Romero 1999: 10)

Here Petra obviously doesn't want her friends to play with rabies-infected dogs. Her desire is directly solely to the concept "the oldest dog (in that neighbourhood)".

[19]Assuming some purpose/perspective P such that the set of individuals defined by P (i.e. C(P)) is a subset of the set Q to which the choice function applies.

(44) **Important Predicate Approximating Basic Intensional Choice Functions**:
Given a set of epistemically-accessible worlds F, a purpose or perspective P, a predicate $Q_{\langle\langle s\tau\rangle,t\rangle}$, such that Q≠∅, a basic intensional choice function f is an Important Predicate Approximate, i.e. $\text{BASICH}_{\sim\text{ImpPred}}$(f), if:
$\exists\phi[\forall w' \in F.\text{ImpPred}(\phi,N,P,w') \;\&\; C(P)(w') \subseteq \{x(w') \in D_\tau \mid x \in Q\} \;\&\; f(Q)(w')=N]$

Given this definition of Important Predicate approximating basic intensional choice functions, we may define *hari* and *də* as follows in (45) and (46).

Hari-type indefinites, i.e. extensionally-unknown indefinites, signal that the speaker knows of no intensional concept whose extension both satisfies the existential claim AND is identical in all epistemically-accessible worlds. We can formally express this notion as follows:

(45) $hari_i(g)(w)(F)(G)(Q)(S)$, where g is an assignment function, w is a world, F is a modal base supplied by context, G is a set of epistemically-accessible Important Predicates, Q, S are predicates
   a. denotes: g(i)(w).BASICH(g(i))
   b. "signals": $\neg\left[\exists f.\text{BASICH}_{\sim\text{ImpPred}}(f)\ \neg[\exists w', w'' \in F: S(w')(f(Q)(w'))=1 \;\&\; S(w'')(f(Q)(w''))=1 \;\&\; f(Q)(w') \neq f(Q)(w'')]\right]$

That is, there is no Important Predicate approximating basic intensional choice function f such that it is not that case that there are at least two epistemically-accessible worlds where the extensions of the element chosen by f from Q are non-identical and both satisfy the predicate S.

Də-type indefinites, i.e. intensionally-unknown indefinites, signal that the speaker knows of no intensional concept whose extension satisfies the existential claim in all epistemically-accessible worlds. Using the concepts developed here, this can be formally stated as:

(46) $də_i(g)(w)(F)(G)(Q)(S)$, where g is an assignment function, w is a world, F is a modal base supplied by context, G is a set of epistemically-accessible Important Predicates, Q, S are predicates
   a. denotes: g(i)(w).BASICH(g(i))
   b. "signals": $\exists w' \in F \neg\left[\exists f.\text{BASICH}_{\sim\text{ImpPred}}(f): S(w')(f(Q)(w'))=1\right]$

In other words, there is at least one epistemically-accessible world w′ for which it is not the case that there is an Important Predicate approximating basic intensional choice function f such that the extension in w′ of the individual concept selected by f is a member of S in w′.

Consider the Sinhala sentences in (47), (48).

(47) Sanat monəva hari gatta.
Sanath what *hari* buy.PAST.A
"Sanath bought something." [extensionally-unknown]

(48) Sanat monəva də gatta.
Sanath what *də* buy.PAST.A
"Sanath bought something." [intensionally-unknown]

Given the denotations above in (45) and (46), we can translate these sentences as follows:

(49) $[\![(47)]\!]^g$ =

a. asserts: $\forall w \in F[\exists f.\text{BASICH}(f).\text{Sanath bought } f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w) \text{ in } w]$

b. 'signals': $\neg[\exists f.\text{BASICH}_{\sim \text{ImpPred}}(f)\ \neg[\exists w', w'' \in F$:
$\text{Sanath bought}(f(x \in D_{\langle se \rangle} \mid x \in \text{non-human}')(w')) \text{ in } w'=1\ \&$
$\text{Sanath bought}(f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w'')) \text{ in } w''=1\ \&$
$f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w') \neq f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w'')]]$

(50) $[\![(48)]\!]^g$ =

a. asserts: $\forall w \in F[\exists f.\text{BASICH}(f).\text{Sanath bought } f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w) \text{ in } w]$

b. 'signals': $\exists w' \in F \neg[\exists f.\text{BASICH}_{\sim \text{ImpPred}}(f)$:
$\text{Sanath bought } (f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w')) \text{ in } w'=1]$

The above semantic/pragmatic assignments correctly predict the distribution of *də* and *hari* indefinites: *hari* indefinites are felicitous where the speaker is uncertain of the extension for which the proposition is true; *də* indefinites are felicitous where the speaker is uncertain of the intensional concept for whose extension the proposition is true. Furthermore, these assignments can be extended to other epistemic indefinites: specifically, the assignment given to *hari* can be extended to English *some* (in *some* NP) and the assignment given to *də* can be extended to Spanish *algún*.

7.5.4 Evaluating the proposed analysis

To show that the meanings assigned to the Sinhala Q particles *hari* and *də* in (49) and (50) above succeed in capturing the actual linguistic data, let us consider the pragmatic conditions under which (49) and (50) predict *hari* and *də* indefinites to be felicitous or infelicitous.

In (51) I provide Sinhala sentences containing the three types of indefinites discussed herein, "plain" indefinites (with no additional pragmatic signal), "extensionally"-unknown, and "intensionally"-unknown indefinites.

(51) a. sanat deyak gatta.
Sanath thing.INDEF buy.PAST.A
"Sanath bought a thing." ["plain" indefinite]

b. Sanat monəva hari gatta.
Sanath what *hari* buy.PAST.A
"Sanath bought some thing." [extensionally-unknown indefinite]

c. Sanat monəva də gatta.
Sanath what *də* buy.PAST.A
"Sanath bought some thing-or-other." [intensionally-unknown indefinite]

The English examples in (52) are intended to be roughly equivalent to the Sinhala examples in (51).

(52) a. Sanath bought a painting. ["plain" indefinite]

b. Sanath bought some painting. [extensionally-unknown indefinite]

c. Sanath bought some painting-or-other. [intensionally-unknown indefinite]

In case the speaker knows that Sanath bought Jean-Baptiste Greuze's painting *The White Hat*, both of the Sinhala sentences, (51b) and (51c), and both of the English sentences, (52b) and (52c), are predicted to be infelicitous.

Scenario 1: Let us assume a scenario in which the speaker knows that Sanath bought the *White Hat* painting, and can identify that painting (i.e. he knows the extension of the intensional concept *White Hat* painting) in all epistemically-accessible worlds; more formally stated in (53).

(53) $\exists x \in D_e.\forall w \in F$[bought(x,Sanath) in w & Greuze's *The White Hat* in w=x]

The use of either (52b) or (52c) is infelicitous in the case in which the speaker knows (53).

A speaker who used (52b) would pragmatically signal that if there is any Important Predicate approximating basic intensional choice function f such that there exists a w' where the extension of the individual concept chosen by f satisfies the proposition in w', then there will also exist another epistemically-accessible world w'' where the extension of the individual chosen by f also satisfies the proposition in w'' and where the two extensions are non-identical:

(54) $\neg\big[\exists f.\text{BASICH}_{\sim\text{ImpPred}}(f)\ \neg[\exists w', w'' \in F$:
Sanath bought$(f(x \in D_{\langle se\rangle} \mid x \in \text{non-human}')(w'))$ in w'=1 &
Sanath bought$(f(\{x \in D_{\langle se\rangle} \mid x \in \text{non-human}'\})(w''))$ in w''=1 &
$f(\{x \in D_{\langle se\rangle} \mid x \in \text{non-human}'\})(w') \neq f(\{x \in D_{\langle se\rangle} \mid x \in \text{non-human}'\})(w'')]\big]$

(54) thus contradicts (53), since (53) requires that Sanath bought the same (extensional) thing in all epistemically-accessible worlds, while (54) requires that there two distinct epistemically-accessible worlds in which Sanath bought different things.

A speaker who used (52c) would pragmatically signal that there is at least one epistemically-accessible world w' in which for no Important Predicate approximating basic intensional choice function f selects an individual concept whose extension in w' satisfies the proposition. This is stated formally in (55).

(55) $\exists w' \in F \neg\big[\exists f.\text{BASICH}_{\sim\text{ImpPred}}(f)$:
Sanath bought $(f(\{x \in D_{\langle se\rangle} \mid x \in \text{non-human}'\})(w'))$ in w'=1$\big]$

(55) also contradicts (53), for (53) requires that the speaker know what Sanath bought in all epistemically-accessible worlds.

Scenario 2: If we consider a scenario in which the speaker knows that Sanath bought a painting, by a French artist, of a girl in a white hat, which hangs in the Boston Museum of Fine Arts (assuming for the sake of exposition that there is only one painting matching that description), but where he could not say whether this painting is Jean-Baptiste Greuze's *The White Hat* or Pierre Renoir's *Young woman in a white hat*, etc.[20] That is, while it may be true that the speaker's intensional concept maps to a unique individual in all epistemically-accessible worlds, for which it is true of that individual that Sanath bought it, the actual individual (painting) is not itself constant between worlds. I.e., the speaker's knowledge is compatible with Sanath buying different paintings. Formally, this means that the speaker knows only that:

(56) $\exists x \in D_{\langle s,e\rangle}$[$\forall w \in F$.bought(x(w),Sanath) in w & x=the French painting of a girl in a white hat in the Boston Museum of Fine Arts]

Here the use of (52b) is appropriate, since (56) is compatible with (54).

But (52c) is not felicitous in this context, since (55) contradicts (56): the latter states that in all epistemically-accessible worlds w the speaker knows of an individual concept whose extension in w satisfies the proposition—

[20] Renoir's painting does not in fact hang in the Boston Museum of Fine Arts, but let us here pretend that it does.

namely the individual concept "the French painting of a girl in a white hat in the Boston Museum of Fine Arts", while the former states that there is at least one epistemically-accessible world w′ in which there is no individual concept whose extension in w′ satisfies the proposition.

Scenario 3: Sentences like (52c) are only felicitous in contexts where the speaker lacks ANY means of uniquely identifying an individual who satisfies the proposition in all epistemically-accessible worlds. One such context would be where the speaker knows only that Sanath bought a painting of a girl in a white hat by a French artist, and where there are multiple paintings compatible with this knowledge. In this scenario the speaker's knowledge does not serve to uniquely identify either an individual or an individual concept, as indicated by its formalisation in (57).

(57) $\forall w \in F[\forall x \in D_e.\text{bought}(x,\text{Sanath})$ in w $\rightarrow$ painted-by-french-artist′(x) in w & painting-of-girl-in-white-hat′(x) in w]

In this context, (51c) and (52c) are appropriate, since (54) does not contradict (57). (54) requires that there is no intensional concept whose extension satisfies the proposition in all epistemically-accessible worlds, but it does permit that the speaker may have certain knowledge about the entity in question—just not uniquely identifying knowledge.

Thus, the proposed denotations for extensionally- and intensionally-unknown epistemic indefinites are shown to correctly predict the contexts in which they are appropriate. In the following section I address the question of the nature of the ignorance component involved in epistemic indefinites.

7.6 What sort of pragmatic signal is the ignorance component?

I argue that the ignorance component of particles like *hari*, *də*, as well as quantificational adjectives like *some* and *algún* is, in fact, a presupposition, rather than a conversation implicature, as argued by Alonso-Ovalle & Menéndez-Benito (2010).

The ignorance component of WH+*hari* and WH+*də* is—as is expected for a presupposition—cancellable, as shown by (58) and (59), as is that of both English *some*, as shown by (60), and Spanish *algún*, as shown by (28), repeated below as (61).

(58) Gunəpālə īye kauru hari kasada-benda. Ettaṭama mama dannava kau də īya bende kiyala.
Gunapala yesterday who *hari* marriage-tied.A. In fact I know.A who *də* yesterday marriage did.A
"Gunapala married someone yesterday. In fact, I know who he married yesterday."

(59) Gunəpālə īye kau də kasada-benda. Ettaṭama mama dannava kau də īya bende kiyala.
Gunapala yesterday who *də* marriage-tied.A. In fact I know.A who *də* yesterday marriage did.A
"Gunapala married someone yesterday. In fact, I know who he married yesterday."

(60) John married some girl yesterday. In fact, I know which girl it was that he married.

(61) María se casó con algún estudiante de lingüística. De hecho, sé exactamente con quién.
María SE marry.3SG.PAST with *algún* student of linguistics. In fact, I know exactly with whom.
"María married a linguistics student. In fact, I know exactly who."

Recall that the argument in Alonso-Ovalle & Menéndez-Benito 2010 against treating the ignorance component as a presupposition is the felicity of sentences like (26), repeated below as (62), since the ignorance component seems to disappear under negation, whereas presuppositions usually persist under negation.

(62) No es verdad que Juan salga con alguna chica del departamento de lingüística
not is true that Juan goes-out with *algún* girl from the department of linguistics
"Juan is not dating any of the girls in the linguistics department."

In Sinhala as well, equivalent sentences with either *kau hari* or *kau də* are possible, as shown by (63a) and (63b), respectively.

(63) a. Gunəpālə īye kauru hari kasada-benda kiyanne boru
Gunapala yesterday who *hari* marriage-tied-A that false
"It is false that Gunapala married someone yesterday."

b. Gunəpālə īye kau də kasada-benda kiyanne boru
Gunapala yesterday who *də* marriage-tied-A that false
"It is false that Gunapala married someone yesterday."

Given the analysis proposed above in Section 7.5, (63a) translates as:

(64) $[\![(63a)]\!]^g$ =

a. asserts: $\forall w \in F[\exists f.\text{BASICH}(f).\text{Gunapala married } f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w) \text{ in } w]$

b. 'signals': $\neg[\exists f.\text{BASICH}_{\sim\text{ImpPred}}(f)\ \neg[\exists w', w'' \in F$:
Gunapala married$(f(x \in D_{\langle se \rangle} \mid x \in \text{human}')(w'))$ in $w'=1$ &
Gunapala married$(f(\{x \in D_{\langle se \rangle} \mid x \in \text{human}'\})(w''))$ in $w''=1$ &
$f(\{x \in D_{\langle se \rangle} \mid x \in \text{human}'\})(w') \neq f(\{x \in D_{\langle se \rangle} \mid x \in \text{human}'\})(w'')]]$

and (63b) as:

(65) $[\![(63b)]\!]^g$ =

a. asserts: $\forall w \in F[\exists f.\text{BASICH}(f).\text{Gunapala married } f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w) \text{ in } w]$

b. 'signals': $\exists w' \in F \neg[\exists f.\text{BASICH}_{\sim\text{ImpPred}}(f)$:
Gunapala married $(f(\{x \in D_{\langle se \rangle} \mid x \in \text{non-human}'\})(w'))$ in $w'=1]$

Neither of the pragmatic "signals" contradicts the assertion, since both signals are negatively stated. That is, the "signal" does not disappear, but since the signals are negatively stated, neither signal requires the existence of an invididual who satisfies the proposition, and so neither creates a contradiction..

Furthermore, as is expected if the signal is a presupposition, we find that the ignorance component of epistemic indefinites in embedded clauses projects to the matrix level: contrast the felicity of (66b) in context 1 with the infelicity of (67b) in context 2.

(66) Context 1: John, Bill, and Bill's sister Mary—who are all linguists—are attending a party where all of the other guests are philosophers. All three of the linguists share a great disdain for philosophers and are only attending the party because there's an open bar. They don't know any of the philosophers. John spots Mary kissing one of the other guests. Shocked he tugs on Bill's coat-sleeve and exclaims: "Look, your sister is kissing some philosopher!" Bill, sighing, replies:

a. "Well, if Mary is kissing a philosopher, she must be drunk."

b. "Well, if Mary is kissing some philosopher, she must be drunk."

(67) Context 2: John, Bill, and Bill's sister Mary—who are all linguists—are attending a party where all of the other guests are philosophers. All three of the linguists share a great disdain for philosophers and are only attending the party because there's an open bar. They don't know any of the philosophers—except for one: George, whose office happens to be next door to the office John and Bill share. John spots Mary kissing George. Shocked he tugs on Bill's coat-sleeve and exclaims: "Look, your sister is kissing George!" Bill, sighing, replies:

a. "Well, if Mary is kissing a philosopher, she must be drunk."

b. #"Well, if Mary is kissing some philosopher, she must be drunk."

If the ignorance component of *some philosopher* did not project up to the matrix level, we would not expect any felicity contrast between (67a) and (67b).[21]

However, there exists one other potential objection to treating the ignorance component as a presupposition. Condoravdi (2005) points out that—though the existential presupposition of a sentence like (68a) is deniable as in (68b)—the ignorance component of *whatever* is not deniable, as shown by (69) and (70).

(68) a. The king of France is bald.

b. The king of France is NOT bald because there IS NO king of France!

(69) A: I want you to empty whatever is in your bag.

a. B: I will NOT empty whatever is in my bag because there is NOTHING in my bag!

b. B: I will NOT empty whatever is in my bag because you KNOW what is in my bag! (does not deny the ignorance implication of A's utterance)

(70) A: Will you welcome whoever is visiting?

a. B: No, I will NOT welcome whoever is visiting because NO-ONE is visiting!

b. B: No, I will NOT welcome whoever is visiting because you KNOW who is visiting! (does not deny the ignorance implication of A's utterance)

Both (69b) and (70b) are odd utterances given that the subordinate clauses do not result in a denial of the ignorance implication of A's utterance; (69b) asserts that B will not empty the contents of B's bag because A knows what is in B's bag, and (70b) asserts that B will not welcome those people who visit due to the fact that A knows who those people are.

We can construct equivalent examples which show that the ignorance component of *some* is not deniable in this way either, thus:

(71) a. A: I met some girl yesterday.

b. B: No, you DIDN'T meet some girl yesterday because you know very well you met MARY, your hitherto secret mistress! (does not deny the ignorance implication of A's utterance)

Yet denials like (68b) rely on the fact that the falsity of the presupposition (i.e. the king of France exists) entails that the assertion (I met the king of France) is false (cp. Cohen 2006). On the other hand, the denial of the presupposition of (71a) (i.e. I don't know which girl I met yesterday) does not entail that I didn't meet some girl yesterday. Therefore, the fact that plain negations of sentences containing indefinites which carry ignorance components do not deny the presupposition is unsurprising.

Note that metalinguistic negation (Horn 1989) can, in fact, deny the ignorance component, as shown by the

[21]Note further that embedding an epistemic indefinite does not result in any contribution of epistemic certainty to the assertion of the sentence. If it did, then (66b) and (67b) should mean something like "If we know that there is a philosopher that Mary is kissing. . . "

exchange in (72).

(72) a. A: I met some girl yesterday.
b. B: No, you DIDN'T meet (just) "some girl" yesterday; you know very well you met MARY, your hitherto secret mistress! (does deny the ignorance implication of A's utterance)

Of course, metalinguistic negation can be used to object to many aspects of an utterance, so perhaps this is also unsurprising. But consider the fact that neither standard negation (73b) nor metalinguistic negation (73c) can be used to deny a conventional implicature:

(73) a. A: I met Edna, a fearless leader, yesterday.
b. B: #No, you DIDN'T met Edna, a fearless leader, yesterday, because Edna is not a fearless leader!
c. B: #No, you DIDN'T met Edna, "a fearless leader", yesterday, because Edna is not a fearless leader!

In contrast, metalinguistic negation is in fact a possible strategy for denying existence presuppositions, as shown by:

(74) a. A: I met the king of France yesterday.
b. B: No, you DIDN'T meet "the king of France" yesterday because there is NO KING of France—you met JOHN! Just because he goes round calling himself the king of France doesn't make it so!

In (74), Speaker B knows exactly who Speaker A refers to by the phrase "the king of France", and does not deny that Speaker A met that person but rather denies that the person Speaker A met is the king of France.

Thus, given the translations suggested above for *hari* and *də*, the ignorance component of epistemic indefinites can be reasonably treated as a presupposition.

7.7 Impact of presuppositions on the use of Q-particles in other contexts

At this point, one might well raise the question of whether the presuppositions posited for *də* and *hari* affect the analysis of Q-particles in other contexts. That is, does the pragmatic analysis proposed here create difficulties for the analysis of Q-particles proposed thus far?[22]

Let us first consider the case of *hari* in non-interrogative disjunctions like (75).

(75) Gunəpalə hari Chitra hari gaməṭə giyā.
Gunapala *hari* Chitra *hari* village.DAT go.PAST.A
"Gunapala or Chitra went to the village."

Here the speaker asserts that for all epistemically-accessible worlds w, it is true that either Gunapala or Chitra (or both) went to the village in w. The speaker also presupposes that there are at least two worlds w′ and w″ such that Gunapala went to the village is true in w′ and Chitra went to the village is true in w″. In other words, (75) is only felicitous in case there are epistemically-accessible worlds in which "Gunapala went to the village" is true and worlds in which "Chitra went to the village" is true. Thus the posited presupposition is not only consistent with the use of *hari* in non-interrogative disjunctions, but in fact correctly predicts the fact that a (co-operative) speaker would not use (75) if he was certain about whether it was Gunapala or Chitra who went to the village.

In the case of alternative questions like (76), we also find that the presupposition proposed for *də* is consistent

[22] See further Chapter 11.2.3 for speculations concerning the relationship of the presuppositions of *də* and *hari* with their historical development.

with the data.[23]

(76) Gunəpalə də Chitra ə gamə̤ṭə giyē?
Gunapala *də* Chitra *də* village.DAT go.PAST.E
"Was it Gunapala or Chitra who went to the village?"

Since (76) is a question, there is no assertion. The speaker does presuppose that there is at least one world w′ for which he has no means of uniquely identifying a intension x such that it is true of the extension of x in w′ that x went to the village in w′, but this presupposition creates no contradiction. Rather this presupposition is consistent with the speaker's lack of knowledge about who went to the village, which itself is signalled by his use of a question.[24] The same holds true to for *wh*-questions like (77).

(77) Kau də gamə̤ṭə giyē?
who *də* village.DAT go.PAST.E
"Who went to the village?"

In (77) there is no assertion. The presupposition is the same as for (76), namely that there is some world w′ in which he has no means of uniquely identifying a intension x such that it is true that the extension of x in w′ went to the village in w′. Again, the lack of knowledge thus signalled by *də* is consistent with the speaker's use of a question.

In sum, the presuppositions proposed in this chapter for *də* and *hari* do not create difficulties for any aspect of the analysis of Q-particles in other (non-indefinite) environments; in fact, as we shall see in Chapter 9.4, the proposed presuppositions will in fact play a crucial role in accounting for the distribution of *də* and *hari* in disjunctive contexts.

7.8 Conclusions and Implications

The felicity conditions on the two types of 'unknown' indefinites in Sinhala have been demonstrated to be correctly predicted by an analysis employing intensional choice-functions combined with the notion of Important Predicates. Both of the Q-particles which occur in *wh*-based indefinites, namely *hari* and *də*, can be treated as denoting variables over choice-functions—an analysis which is supported by the fact that this denotation is also appropriate to the other contexts in which these particles appear, specifically yes/no, alternative, and *wh*-questions in the case of *də*, and declarative disjunctions in the case of *hari*.[25]

The WH+*hari* indefinite is felicitous in case the speaker knows of no individual which uniquely satisfies his existential claim in all epistemically-accessible worlds, the WH+*də* indefinite is felicitous in case the speaker knows of no individual concept whose extension uniquely satisfies his existential claim in all epistemically-accessible worlds. English *some NP* appears to be similar to Sinhala WH+*hari* indefinites in terms of its epistemic properties, and English *some NP-or-other* and Spanish *algún NP* resemble Sinhala WH+*də* indefinites. Put differently, WH+*hari* indefinites and English *some NP* are extensionally-unknown and WH+*də* indefinites, English *some NP-or-other*, and Spanish *algún NP* are intensionally-unknown.

I argued that this distinction can be formalised by treating Q-particles as denoting variables over intensional choice functions and appealling to the notion of Important Predicates (Boër & Lycan 1975). "Extensionally-unknown" indefinites like WH+*hari* presuppose that any Important Predicate approximating basic intensional choice function

[23]The presupposition carried by *də* will, however, correctly rule out its use in non-interrogative disjunctions, as discussed in detail in Chapter 9.4.

[24]On the presupposition carried by *-e*, see above Chapters 5 and 6.4.2.

[25]Again, see Chapter 9.4 for discussion of the effect the presuppositions of *də* and *hari* have on the distribution of these particles in disjunctive contexts.

which selects any individual concept whose extension satisfies the existential claim in some epistemically-accessible world has a non-identical extension in some other epistemically-accessible world which also satisfies the existential claim. "Intensionally-unknown" indefinites like WH+*də* presuppose that there is some epistemically-accessible world for which there is no Important Predicate approximating basic intensional choice function which selects an individual concept whose extension satisfies the existential claim.

Given the highly restrictive nature of the pragmatics of 'intensionally'-unknown indefinites like Sinhala WH+*də* and Spanish *algún* and the resultant narrow distribution, it is perhaps unsurprising that this type of element is seemingly rare crosslinguistically. However, in part, this rarity may be more apparent than actual. The English construction *some X or other*, for example, seems to have properties reminiscent of Spanish *algún* and Sinhala WH+*də* indefinites. Thus, a close examination of individual languages may reveal that 'intensionally'-unknown elements are actually not so uncommon.

In terms of the ignorance component, extensionally-unknown and intensionally-unknown indefinites are classes which are coherent at least across English, Sinhala, and Spanish. However, it is clear that there are important semantic/pragmatic differences between Sinhala WH+*də* indefinites, English *some NP-or-other*, and Spanish *algún NP*, as regards specificity (see footnotes 7, 13), but in terms of the ignorance component, extensionally-unknown and intensionally-unknown indefinites are classes which are coherent at least across English, Sinhala, and Spanish.

As mentioned briefly in note 7, epistemic indefinites may be specific or non-specific (and this is an area in which we observe crosslinguistic variation), as shown by the contrast in the possible interpretations of the following Sinhala examples:

(78) a. Hāmə pirimi-laməyek-mə dakka hāmə gahanu-laməyek-mə kāva **hari** iṁbinawa.
every boy.INDEF-PART see.PAST.A every girl.INDEF-PART who.ACC ***hari*** kiss.PAST.A

b. Hāmə pirimi-laməyek-mə dakka hāmə gahanu-laməyek-mə kāva **də** iṁbinawa.
every boy.INDEF-PART see.PAST.A every girl.INDEF-PART who.ACC ***də*** kiss.PAST.A
"Every boy saw every girl kiss someone."

Here the only available interpretation of (78b) is that there is a specific (though unknown) person *x* such that every boy saw every girl kiss *x*—an interpretation which is not available for (78a). The interpretation of (78b) stands in sharp contrast to Spanish *algún*, which can bear non-specific interpretations, as in (79)— but is however, like Sinhala WH+*də*, an "intensionally-unknown" epistemic indefinite.

(79) Hay **alguna** mosca en la sopa.
There is *algún* fly in the soup
(Alonso-Ovalle & Menéndez-Benito 2010: 24)

Here (79) conveys that there is at least one fly (but possibly more) in the soup.

That *hari* indefinites are non-specific is also apparent in the following example:

(80) maṭə kāva hari iṁbinnə onə.
I.DAT who.ACC *hari* kiss.INF want.PRES
"I want to kiss someone(, anyone)."

As noted above, in some respects, Sinhala *hari* indefinites appear to resemble English *some NP* and *də* indefinites English *some NP-or-other*.[26] The latter, like *də* indefinites, seem to behave like specific indefinites:

(81) John wants to kiss some girl-or-other.

Example (81) can bear the interpretation that there is some specific girl (unknown to the speaker) who John wants to kiss, but it cannot mean that John wants to kiss a girl, where any girl will do.

The resemblance between the Sinhala and English constructions is not perfect, however. Amongst other differences, note that English *some NP-or-other*—unlike Sinhala *də* indefinites—is not required to take the widest available scope, as demonstrated by the fact that *some man-or-other* in the following example may take either wide or intermediate scope:

(82) Every boy saw every girl kiss some man-or-other.

a. *(every boy) > (every girl) > (a man)
b. (every boy) > (a man) > (every girl)
c. (a man) > (every boy) > (every girl)

These two properties of indefinites—specificity and the ignorance component—are not, in fact, entirely disassociated. Consider the sentence below in (83): it has two possible readings, as indicated. Under the analysis proposed here, the English *some NP* type indefinite (sharing many if not all properties of the Sinhala WH+*hari* indefinite) presupposes that there is no choice function f corresponding to an epistemically-accessible Important Predicate such that the individual chosen by f both satisfies the proposition and is identical in all epistemically-accessible worlds.

(83) John wants to kiss some girl.

a. (want) > (some girl) [non-specific]
b. (some girl) > (want) [specific, unknown]

Some girl can take narrow scope, as in (83a), or wide scope, as in (83b). Note, however, on reading (83b), there is a presupposition that the identity of the girl in question is (existentially)-unknown to the speaker. Thus the presupposition proposed for *some NP* indefinites can be cached out in two different ways: either the speaker cannot uniquely identify the girl because John's desire is not directed towards any specific girl (reading (83a)) or else, though there is a specific girl that John wants to kiss, the speaker lacks the knowledge necessary to uniquely identify the girl in question (reading (83b)).

There thus remain a number of avenues for further investigation, including the relationship between the ignorance component and specificity: especially as this appears to be a locus of crosslinguistic variation, e.g. Sinhala WH+*də* and Spanish *algún* NP are both intensionally-unknown indefinites, but the former is always specific, while the latter may behave as a non-specific indefinite in certain contexts (see note 13).

The origin and development of epistemic indefinites (of both extensionally- and intensionally-unknown types) is another area worthy of more research. Though Haspelmath's study (1997) performs yeoman's service with regard to studying the origins of indefinites, further study of the source and evolution of these two different types of epistemic indefinites is still warranted and may well help to explain some of the crosslinguistic differences in properties such as specificity, as well as the connection in some languages (e.g. Sinhala) of epistemic indefinites with both interrogatives and disjunctions. In Chapter (60) I provide an examination of the development of the "ignorance components" of

[26] Becker (1999) examines the specificity properties of English *some NP* and *some NP-or-other*, but she does not identify any particular contrast between the two—though she notes that differences may exist (211).

Sinhala WH+*də* and WH+*hari* indefinites.

The close examination of the pragmatics of the Sinhala Q-particle *də* and *hari* developed in this chapter will be shown to play a crucial role in explaining the overall distribution of Q-particle in Sinhala, as it allows us to account for the inability of *də* to appear in declarative disjunctions (see Chapter 9.4). The next chapter provides a brief discussion on Q-particles in relative clauses, the final syntactic environment in which Q-particles sometimes occur. The chapter following, Chapter 9, pulls together all of the various aspects of Q-particles examined thus far in order to provide a full account of the crosslinguistic and diachronic distribution of Q-particles.

Chapter 8

Excursus: A brief note on Q-particles in Sinhala and Dravidian relatives clauses

As mentioned briefly in Chapter 1, in some stages of Sinhala we also find Q-particles appearing in the formation of relative clauses. This type of relative clause is formed by using *yam*[1] (possibly modifying a noun), co-referring to a correlative demonstrative (expressed or implied) in the correlative main clause. Additionally, either the Q-particle *da* or the conditional particle *nam* must occur at the end of the relative clause. Such constructions are well attested in Classical Sinhala and in modern literary Sinhala, while modern colloquial Sinhala forms relative clauses only with prenominal modifying participles, as in example (1).

(1) [[guruwərəyek wenə] mahattəya] huňgak dannəwa.
[[teacher.INDEF.NOM become.PRES.ADJ] man.DEF] much know.PRES
"The man who is becoming a teacher knows a lot." (cited from Gair 1995[1998]: 245) [MCS]

In modern literary Sinhala, where we find relative clauses formed using the relative pronoun *yam* and either the Q-particle *da* or the conditional particle *nam*, such constructions always seem to have the sense of generalising relatives, i.e. of the type "whoever speaks thus is a fool". Example (2) is representative.

(2) [**yam** kumariyak ohu duṭuvā] **da** [oo ohu kerehi piḷiňda sit ætikara
[**REL-PRON** princess.INDEF him see.PAST.3SG.FEM] ***da*** [she him towards connected mind developed
gattāya]
get.PAST.3SG.FEM]
"**Whatever princess** saw him fell in love with him." (cited from Gair & Karunatilaka 1974: 295) [LS]

In Classical Sinhala too *yam-da/nam* relatives tend to have the free relative generalising sense, as in example (3).

(3) [**yamak'hu** paḷamu diṭim] **da** [ohu marā gaṇan sapurami]
[**REL-PRON**.MSC.SG.ACC firstly see.1SG] ***da*** [him kill.CONV number complete.PRES.1SG]
"**Whichever person** I see first, I shall kill him and complete the number." (*Ama.* 133, cited from Wijemanne 1984: 210) [CS]

This is apparently not always the case, as evidenced by examples such as (4), where the relative appears to refer to a specific individual.

(4) [mam **yamak'hu-ge** savuyem] **da** [ohu-ge guṇa asava]
[I **REL-PRON**.SG.GEN follow(?)] ***da*** [his virtues listen.IMP]
"Listen to the virtues of the **person whose** follower I am." (*Ama.* 93, cited from Wijemanne 1984: 210) [CA]

[1] *Yam*, as mentioned in Chapter 11.2, in fact descends from the Old Indo-Aryan relative pronoun base. Also, as previously noted, it can act as an indefinite pronoun as well.

Similar to the relative clause constructions of Classical and modern literary Sinhala are those of modern Dravidian languages like Tamil, where we find examples like (5).

(5) [yār aṅkē mutalil vantu ceru-v-ār]$_{RC}$-**oo** [avar ṭikeṭṭu vaṅkalām]$_{CC}$
[who.NOM there first come.CONV arrive.FUT.3PL]$_{RC}$-***oo*** [they.NOM.PL ticket.NOM buy.PERM]$_{CC}$
"Let whoever reaches there first buy the tickets." [Modern Tamil]
(cited from Annamalai & Steever 1998)

In both Old Sinhala (see Paranavitana 1956: clxvii) and Old Tamil (see Hock 2008, referring to p.c. w/ Th. Lehmann) we find vanishingly few examples of relative-correlative structures (four examples in Old Sinhala; five to ten examples in Old Tamil). In Old Sinhala, *da* is absent from all *yam* relative clauses, the conditional particle *nam* occurs in a single example (graffito 251), the remaining three examples employ no particle. In all of the Old Tamil examples, we in fact find that no particle follows the relative clause, as in example (6). Hock (1988, 1989, 2008) points out that the same is true for Old Malayalam (see example (7)).

(6) [e-vaṟi nall-avar āṭavar]$_{RC}$ [a-vaṟi nall-ai]$_{CC}$
[which-place good.MASC.3PL men.MASC.3PL]$_{RC}$ [that-place good.2SG]$_{CC}$
"At which place men are good, at that place you are good." (cited from Lehmann 1998: 94) [Old Tamil]

(7) [yātoṉṯu mahārājaniyōgam]$_{RC}$ [atu a-vaṇṇam]$_{CC}$
[what.NEUT maharaja-order]$_{RC}$ [that.NEUT in-that-manner]$_{CC}$
"What is the king's order, (let) that (be done) in that manner." (cited from Pillai 1973: 165) [Old Malayalam]

Also lacking in post-relative clause particles are Old Kannada (Hock 2008), and a number of modern "northern" Dravidian languages (Pengo, Kuvi, Kolami, Parji, Kurukh), on which see Hock (1988, 1989, 2008). Hock (2008) also notes that even in modern Malayalam the post-relative clause *-oo* is optional (cf. Asher & Kumari 1997: 53).

It is beyond the scope of this study to provide a complete analysis of relative clauses and their connection with Q-particles. However, the basic connection is clear enough: relative pronouns such as *yam*, like indefinite pronouns, seem to create Hamblin-type sets, e.g. in (2) the set {x | princess′(x) & saw-him′(x)},[2] and choice functions are appropriate in contexts containing elements denoting Hamblin-type sets: thus the possibility for the appearance of *da* in this context. However, instead of *da*, we also find—in both Classical Sinhala and modern literary Sinhala—relatives formed with *nam* in place of *da*. The modern literary Sinhala example in (2) may have its *da* replaced by *nam* without change in meaning. An example of a *nam*-type relative from Classical Sinhala appears below in (8).

(8) **yam** gihi minisek ovun-ge vāda maṅḍanaṭa nisi vī **nam**
REL-PRON householder person.INDEF their talks trample.INF.DAT suitable be.PAST.3SG **COND.PTCP**
ohaṭa pādaparicārikā veti.
him.MASC.DAT wed.PRES.3SG
"They become the wives of **any layman who** may be able to refute their arguments." (*Ama.* 150, cited from Wijemanne 1984: 212) [CS]

The connection between the conditional construction and generalising relatives is transparent: (8) can also be interpreted as meaning "If a layman is able to refute their arguments, then they become his wives".

Less clear is how the choice function denoted by *da* in *yam-da* relatives are bound. The possibility of existential binding (via existential closure, as in the case of indefinites involving Q-particles) seems to be ruled out by the semantics, given that all of the relatives examined either appear to denote a definite (though possibly

[2] As noted in Chapter 11.2, it is relevant to observe that *yam* itself can also function as an indefinite pronoun.

unspecific/unknown) individual (as in examples (3), (4), (8)) or be akin to a universal (as in example (2)).

There is an additional concern regarding the status of such constructions in Classical Sinhala. As noted above, the Old Sigiri graffiti contain vanishingly few examples of *yam* relative constructions and none which involve the Q-particle *da*. The Classical Sinhala texts are largely translations of or commentaries upon Pāli Buddhist texts. And Wijemanne (1984: 212) remarks that "[a]lmost all the relative constructions in the Amāvatura [the text from which her examples are drawn–BMS] are exact renderings of Pali relative constructions". This raises the question of the status of such constructions in Classical Sinhala: would they have been possible in ordinary language or do they represent a construction which is only part of the specialised "translation/commentary" register?

However, both this issue and the question of the proper formal analysis of such constructions fall outside the bounds of the present study.[3]

[3]Amongst other issues, there is the question of the relationship of the use of particles following relatives clauses with "finiteness" constraints in some SOV languages, on which see Hock (2008, to appear) and references therein.

Chapter 9

Accounting for the distribution of Q-particles crosslinguistically & diachronically

The previous chapters have laid the groundwork for an explanation of the distribution of Q-particles both crosslinguistically and diachronically, which is the subject of this chapter.

I offer an account of the crosslinguistic distribution of Question-particles [Q-particles] relying on four components: (1) compositional semantics (Q-particle occur in environments containing an element with Hamblin-type semantics); (2) lexico-semantic differences (*wh*-words may be semantically simple, or semantically-complex); (3) differences in formal syntactic features (feature valuation); (4) fine-grained pragmatic differences between Q-particles (Q-particles may be associated with presuppositions).

This analysis thus provides an argument for the necessity of evaluating complex linguistic data by considering the role played by each of the modules of universal grammar. In other words, the distribution of Q-particles is not susceptible to a purely syntactic or a purely semantic analysis, but requires the consideration of the interaction of various components of the grammar, including semantics, pragmatics, syntax, and morphology.

9.1 Defining Q-particle environments I: Hamblin-semantics restrictions

As discussed extensively in this study, what has come to be known as a "Question" or "Q-particle" (cf. Baker 1970; Cable 2007)—nomenclature notwithstanding—surfaces crosslinguistically in a wide variety of syntactic environments, occurring not only in interrogative contexts, but also in disjunctions and in the formation of certain types of indefinites.[1] See Table 9.1.[2]

	Old Sin	Class Sin	Lit Sin	Colloq Sin	Old Mal	Mod Mal	Tlin	Jap
y/n-ques.	(da)	(da)	da	də	-oo	-oo	gé	ka, no, kai, kadooka
wh-ques.	(da)	(da)	da	də	-oo	—	sá	ka, no, ndai
wh-indef.	—	—	hō (aff. & neg.), vat (neg.)	də (aff.), hari (aff.), vat(neg.)	-oo	-oo	sá	ka
decl. disj.	hō, heva(-t)	hō, heva(-t)	hō (aff. & neg.), vat (neg.)	hari (aff. & neg.), vat (neg.),	-oo	-oo	khach'u	ka
interr. disj.	da	da	da	də	-oo	-oo	gé... gwáa	[ka]

Table 9.1: Distribution of Q particles in various stages of Sinhala & Malayalam; Tlingit, Japanese

As argued throughout, the occurrence of Q-particles in this seemingly disparate set of environments can be accounted for in terms of compositional semantics, if Q-particles are treated as denoting variables over choice

[1] In Japanese, for instance, the Q-particle *ka* occurs in all of these environments (Hagstrom 1998).

[2] Square brackets indicate some additional complication. Round brackets indicate optionality/variation.

functions (cf. Hagstrom 1998; Cable 2007), as in (2); where a choice function is a function which, when applied to a non-empty set, returns a single member of that set, as in (1). Interrogatives are often analysed as sets of propositions (Hamblin 1973; Karttunen 1977), and so it is perhaps unsurprising that Q-particles often appear in interrogative environments. And the appearance of Q-particles with indefinites and disjunction seems to be motivated for similar reasons—given that both indefinites (Reinhart 1997, 1998; Winter 1997; Hagstrom 1998; Cable 2007) and disjunctions (Alonso-Ovalle 2006) can also be analysed in terms of quantification over Hamblin-type sets, where an element with Hamblin-type semantics denotes a set (possibly a singleton set) of elements. Namely all of these environments involve an element with Hamblin-type semantics.

(1) **Choice function**:
A function f is a choice function (i.e. CH(f) holds) iff for every non-empty predicate P, f(P) is defined.

(2) $[\![Q_i]\!]^g = g(i) \in D_{cf}$

As discussed above in Chapter 6.1, the semantics I adopt are a mixed system: for most items I use standard Montagovian compositional semantics (Montague 1970a,b, 1973). But I assume that some items inherently have a Hamblin-type semantics (Hamblin 1973), following in the vein of a number of earlier studies (Ramchand 1997; Hagstrom 1998; Sternefeld 2001; Kratzer & Shimoyama 2002; Alonso-Ovalle 2006)—see above Chapter 6. Hamblin (1973) treats most elements as denoting singleton sets containing what would be their ordinary Montagovian denotation, thus the denotation of an ordinary declarative for Hamblin is a singleton set containing a proposition. However, Hamblin treats *wh*-words as denoting non-singleton sets, which compose with other elements via pointwise composition. A sentence containing a *wh*-word will thus end up being a set of propositions—which Hamblin takes to be the proper semantic type of an interrogative.

Thus, in part, the distribution of Q-particles can be explained semantically in that Q-particles may only occur in environments in which a choice function may apply, i.e. environments involving sets. However, it must be noted that technically in normal Montagovian semantics, even functions like (3) are sets (cf. Partee et al. 1990, Heim & Kratzer 1998).

(3) $[\![\text{human}]\!]^g = \lambda x.x \text{ is a human}$

The function in (3) is formally equivalent to the set A such that A contains all humans, i.e. $A = \{x \mid x \in \text{human}'\}$.

However, though we find Q-particles appearing obligatorily with *wh*-indefinites in a number of languages, as in the Malayalam example in (4)—where *aar-* "who" denotes $\{x \in D_e \mid x \in \text{human}'\}$—

(4) ñaan̲ innale aar-e-(y)**oo** paricayappeṭṭu
I yesterday who-ACC-**Q** met
"I met somebody yesterday." (Jayaseelan 2001: 66)

the appearance of a Q-particle with an ordinary NP like *aal̤-* "person" is ungrammatical—whether a determiner like *oru* "one" is present (5a) or absent (5b).

(5) a. *ñaan̲ innale oru aal̤-e-(y)**oo** paricayappeṭṭu
I yesterday one person-ACC-**Q** met
"I met a person/somebody yesterday."

b. *ñaan̲ innale aal̤-e-(y)**oo** paricayappeṭṭu
I yesterday person-ACC-**Q** met
"I met a person/somebody yesterday."

I take then elements with Hamblin-type set semantics to differ from elements with Montagovian denotations—even though both can be expressed in terms of sets—with respect to how they combine with other elements. We may adopt the diacritic convention from Sternefeld (2001) in order to formally differentiate these two classes of elements: the semantic type of elements with ordinary Montagovian denotations will be expressed in the normal fashion, e.g. $[\![\text{human}]\!]^g$ is type $\langle e, t\rangle$ or $\langle et\rangle$; whereas an element with Hamblin-type semantics will be expressed using '//' as the separator, e.g. $[\![\text{who}]\!]^g$ is type $\langle e//t\rangle$.

And thus choice-functions can be restricted to apply only to elements with Hamblin-type ('//') semantics, and (1) can be reformulated more precisely as (6).

(6) **Choice function**:
A function f is a choice function (i.e. CH(f) holds) iff $f \in D_{\langle\langle \alpha//t\rangle,\alpha\rangle}$ and for every non-empty predicate $P \in D_{\langle \alpha//t\rangle}$, f(P) is defined and it is in the extension of P (i.e. f(P)∈P holds). (cf. Winter 1997: 410; Sternefeld 2001: 7.)

The definition in (6) restricts choice functions such that they may only compose with elements which denote Hamblin-type sets, i.e. with elements of type $\langle \alpha//t\rangle$. Thus the appearance of Q-particle in a context lacking an element with Hamblin-type semantics will result in a semantic crash.

(7) **First restriction on Q-particles**:
Q-particles can only apply to Hamblin-type ('//') elements.[3]

However, as can be seen by inspection of Table 9.1 above, even in languages which employ overt Q-particles, we do not find Q-particles appearing uniformly in all of these environments—and further, some languages appear to employ different Q-particles in different contexts (e.g. Sinhala *də*, *hari*; Tlingit *sá*, *gé*). Therefore (7) is a necessary (and universal) condition on the appearance of Q-particles, but in order to understand the distribution of Q-particles in specific languages further constraints are still required.

9.2 Defining Q-particle environments II: morpholexical restrictions

As above, I treat the overt Q-particles which appear in certain languages, e.g. Sinhala *də*, Korean *ni*, Japanese *ka*, Malayalam *-oo* etc., as denoting variables over choice-functions. The distribution of Q-particles can also be determined in part by the lexicon. That is, in Sinhala, for instance, the Q-particle *də* is an element of the lexicon; while other languages may lack Q-particles as free-standing morphemes. Additionally, languages may differ as to the semantic status of *wh*-words. In some languages, such as Sinhala and Japanese, *wh*-words simply denote Hamblin-type sets (e.g. Sinhala *kau* "who" denotes $\{x \mid x \in \text{human}'\}$). In other languages, like English, *wh*-words may be more semantically-complex, incorporating, in effect, Q-particles as part of their denotation.

(8) $[\![\text{who}]\!]^g = \exists f.f \in D_{cf}[f(\{x \mid x \in \text{human}'\})]$

In many languages, again including Sinhala and Japanese, such *wh*-words may appear not only as interrogative pronouns, but may also function as indefinite pronouns—following Kuroda (1965) I refer to such *wh*-words as indeterminate pronouns. It is these indeterminate pronouns which crosslinguistically tend to involve the appearance

[3]There are apparent cases where Q-particles apply to non-Hamblin-type sets, e.g. if English *some* in the *some NP* construction is analysed as a Q particle (as argued above in Chapter 7), then it apparently applies—not to a Hamblin-type set—but rather to a normal set of individuals (i.e. a predicate). The solution to this apparent difficulty is to define English *some* (and similar elements) as follows:

(i) $[\![\text{some}]\!]^g = \lambda P_{\langle e,t\rangle}.f(\{x{\in}P\}_{\langle e//t\rangle}).f \in D_{cf}$

of a Q-particle. In other languages, like English, indefinite pronouns are morphologically-distinct from, though often historically connected to, interrogative pronouns (cf. English *somehow*, *somewhere*)—and involve no separate Q-particle. Thus, as in the case of interrogative pronouns, English indefinite pronouns can be analysed as semantically-complex, again in effect incorporating a Q-particle, as in (9).[4]

(9) $[\![\text{someone}]\!]^g = \exists f.f \in D_{cf}[f(\{x \mid x \in \text{human}'\})]$

Yet we cannot directly equate indeterminate pronouns with what I shall refer to as "simple *wh*-words", i.e. *wh*-words which simply denote sets of individuals, as Sinhala *kau* "who" above—where simple *wh*-words require the presence of a Q-particle. That is, some languages, like Japanese and modern colloquial Sinhala, utilise simple *wh*-words both for interrogative and indefinite pronouns; while other languages utilise semantically-complex and morphologically-distinct interrogative and indefinite pronouns. However, the distinction between simple *wh*-words and semantically-complex interrogative and indefinite pronouns is not always manifested by a morphological distinction—thus in modern Malayalam indefinite pronouns and interrogative pronouns are homophonous but syntactically distinct in that the former but not the latter occurs with a Q-particle. Further, a language may employ simple *wh*-words for interrogative pronouns (requiring the presence of a Q-particle), but semantically-complex indefinite pronouns (incompatible with the presence of a Q-particle), or vice-versa.[5]

	Eng.	**Early Sin. (I)**	**Early Sin. (II)**	**Mod. Sin.**	**Old Mal.**	**Mod. Mal.**
morph.-distinct indef. & interrog. pronouns:	yes	yes	yes	no	no	no
interrog. pronoun is semantically:	complex	complex	simple	simple	simple	complex
indef. pronoun is semantically:	complex	complex	complex	simple	simple	simple

Table 9.2: Properties of interrogative and indefinite pronouns

The analysis of potential variation between semantically-simple and semantically-complex *wh*-words finds further support in the fact that even in modern colloquial Sinhala, we find both semantically-simple *wh*-interrogatives (which require the presence of an accompanying Q-particle) and semantically-complex *wh*-interrogatives (which obligatorily occur WITHOUT an accompanying Q-particle). As Kishimoto (2005: 41–43) points out, the Sinhala *wh*-adjunct *æi* "why"—in contrast to all other *wh*-words—obligatorily occurs without any accompanying Q-particle, as shown in (10).

(10) a. Chitra æi potə kieuwe?
Chitra why book read.PAST.E
"Why did Chitra read the book?"

b. *Chitra æi də potə kieuwe?
Chitra why *də* book read.PAST.E

This suggests that Sinhala *æi*, unlike other *wh*-words in Sinhala, is semantically-complex, in effect incorporating both a Hamblin-type set and a Q-particle within a single lexical item. This analysis is strengthened by the fact that the alternative to *æi*, namely *mokə də*—has an inseparable *də*. Unlike other *wh*-word+*də* constructions, *mokə də* always occurs as a single unit, compare (11) with (12).

[4] A similar proposal, positing crosslinguistic variation in whether the variable and binder of *wh*-interrogatives are lexically realised as an individual lexical item or separately, is found in Cole & Hermon (1998: 238–241), cp. Cheng (1991), Aoun & Li (1993), Watanabe (1993), Tsai (1994).

[5] "Early Sinhala" covers both Old Sinhala and Classical Sinhala; both stages show variation, idealised here as two distinct grammars, I and II. "Modern Sinhala" covers both formal and colloquial varieties of the modern language.

(11) a. Ranjit [Chitra monəwa də gatte kiyəla] dannəwa.
Ranjit [Chitra what *də* buy.PAST.E that] know.PRES.A
"Ranjit knows what Chitra bought."

b. Ranjit [Chitra monəwa gatta də kiyəla] dannəwa.
Ranjit [Chitra what buy.PAST.A *də* that] know.PRES.A
"Ranjit knows what Chitra bought."

(12) a. Ranjit [Chitra mokə də aawe kiyəla] dannəwa.
Ranjit [Chitra why *də* come.PAST.E that] know.PRES.A
"Ranjit knows why Chitra came."

b. *Ranjit [Chitra mokə aawa də kiyəla] dannəwa.
Ranjit [Chitra why come.PAST.A *də* that] know.PRES.A
"Ranjit knows why Chitra came."

In general, *də* may be separated from its associated *wh*-word, as shown by examples like (11b); however, the collocation *mokə də* is inseparable, as shown by (12b). This points to *mokə də* as forming a single, unanalysable lexical item, just like *æi*. These examples serve to bolster the argument that there is potential variation in whether the Hamblin-type set component and the Q-component of a *wh*-interrogative are realised as a single lexical item, as in English, and modern Malayalam, and in Sinhala "why", or else as two separate lexical items, as in the remainder of Sinhala *wh*-interrogatives, Japanese *wh*-interrogatives etc.

Lexical differences in whether interrogative and indefinite pronouns are semantically complex or simple is one way in which languages may differ—and since I take it that Q-particles appear with *wh*-words (both interrogative and indefinite) if and only if they are semantically simple—this provides us with one piece of the answer to the question of how to account for the crosslinguistic distribution of Q-particles.

It is interesting to note that the comparison of Sinhala and Malayalam shows that, unsurprisingly, change between semantically-simple and semantically-complex *wh*-words can occur in either direction. That is, the history of Sinhala exhibits a change from semantically-complex to semantically-simple *wh*-words (Early Sinhala I > Early Sinhala II > Modern Sinhala), while the history of Malayalam provides an example of a change from semantically-simple to semantically-complex *wh*-words.

In the next section I turn to the examination of the role that formal syntactic features play in the distribution of Q-particles. As I demonstrate there, the majority of crosslinguistic differences in the distribution of Q-particles can be accounted for by positing language-specific differences in formal feature specification.

9.3 Defining Q-particle environments III: syntactic feature valuation-based restrictions

The lexico-semantic distinction between semantically complex and semantically simple *wh*-words, however, can only account for differences in Q-particle distributions which are connected with *wh*-words. It cannot account for differences in the (non-)appearance of Q-particles in yes/no/alternative questions or disjunctions; further, it accounts only for *whether* a Q-particle is employed or not in *wh*-contexts, but does not aid us in the determination of *which* Q-particle is employed (in case of a language possessing multiple Q-particles).

The penultimate piece of the puzzle lies in language-specific differences in the formal syntactic features borne by the relevant linguistic entities (Q-particles, *wh*-words, COMP heads, disjunctions etc.). The following subsection lays out the basic features I assume are relevant here for the syntactic analysis (see above Chapter 3 for details of the

syntactic framework assumed in this study); this is followed by a detailed examination of the role these features play in the four languages which provide the primary data for this study.

9.3.1 Detailed analysis of the formal syntactic features of Sinhala, Malayalam, Tlingit, and Japanese

Using the system of feature-valuation introduced in Chapter 3.1, I present a detailed analysis of the interaction of the formal syntactic features of Q-particles, complementiser heads, 'junction' (J), and *wh*-words, which provides the all but last piece of the account of the distribution of Q-particles in Sinhala (early and modern), Malayalam (early and modern), and Tlingit.

I assume that complementisers may be sub-divided into interrogative complementisers (CP-INT) and declarative complementisers (CP(decl)), and that C-INT may bear different features from C(decl). I further posit that CP-INT may be (in some languages) further differentiated as necessary as specific to *wh*- or non-*wh*-questions, and the latter category may distinguish between yes/no and alternative questions. In some languages (e.g. Old Malayalam) these may be the only two types of COMP heads; in other languages C-INT heads may be further subdivided into two subtypes, *wh*-associated interrogative complementisers (C-INT(wh)) and non-*wh*-associated interrogative complementisers (C-INT(non-wh)); in other languages (e.g. early Sinhala), the non-*wh*-associated interrogative complementiser may occur as two distinct elements: yes/no-complementiser heads (C-INT(y/n)) and alternative-question complementiser heads (C-INT(alt)).

A question arises regarding the nature of the relevant formal syntactic features. Consider again the distribution of Q-particles in the languages under consideration, Table 9.1, repeated below as Table 9.3.

	Old Sin	Class Sin	Lit Sin	Colloq Sin	Old Mal	Mod Mal	Tlin	Jap
y/n-ques.	(da)	(da)	da	də	-oo	-oo	gé	ka, no, kai, kadooka
wh-ques.	(da)	(da)	da	də	-oo	—	sá	ka, no, ndai
wh-indef.	—	—	hō (aff. & neg.), vat (neg.)	də (aff.), hari (aff.), vat(neg.)	-oo	-oo	sá	ka
decl. disj.	hō, heva(-t)	hō, heva(-t)	hō (aff. & neg.), vat (neg.)	hari (aff. & neg.), vat (neg.),	-oo	-oo	khach'u	ka
interr. disj.	da	da	da	də	-oo	-oo	gé... gwáa	[ka]

Table 9.3: Distribution of Q particles in various stages of Sinhala & Malayalam; Tlingit, Japanese (repeated)

One feature which immediately suggests itself, particularly in the case of Tlingit, is a [wh] feature, since Tlingit *sá* only occurs with *wh*-words, whereas *gé* occurs in other contexts.

However, in all of the languages under consideration, the semantically simple *wh*-words (the *wh*-words of Tlingit are of this type) share a common syntactic property—while there may be no island barriers between the Q-particle itself and the complementiser head of the clause in which the *wh*-word takes scope—there may be (theoretically an infinite number of) island barriers in-between the *wh*-word and the Q-particle.

Consider the following data: CNPs (Complex Noun Phrases) are islands in Sinhala. Again, *wh*-words may be internal to islands, but the Q particle *də* may not, as shown in (13), (14):

(13) a. [[ranjit monəwa gatta kiənə] kaṭəkatāwə] **də** chitra æhuve?
[[ranjit what bought-A that] rumour] **Q** chitra heard-E
'What did Chitra hear the rumour that Ranjit bought?'

b. *[[ranjit monəwa **də** gatta kiənə] kaṭəkatāwə] chitra æhuve?
[[ranjit what **Q** bought-A that] rumour] chitra heard-E

(14) a. [kauru liyəpu potə] **də** ranjit gatte?
[who written book] **Q** ranjit bought-E?
'Who wrote the book Ranjit bought?'

b. *[kau **də** liyəpu potə] ranjit gatte?
[who **Q** written book] ranjit bought-E?

One might suppose in (13a) and (14a) that the *wh*-word has moved covertly to a position within the Spec of the lower CP—an 'escape hatch' position from which it is still visible to syntactic operations of the next phase. However, example (15) shows that this cannot be the case, as the *wh*-word may in fact be inside of an island inside of an island, so long as the Q-particle has no island barriers between it and COMP.

(15) a. [[ranjit [kauru liyəpu potə] gatta kieənə] kaṭəkatāwə] **də** chitra æhuve?
[[ranjit [who written book] bought-A that] rumour] **Q** chitra heard-E
'Who is the person *x* such that Chitra heard the rumour that Ranjit bought the book that *x* wrote?'

b. *[[ranjit [kauru liyəpu potə] **də** gatta kieənə] kaṭəkatāwə] chitra æhuve?
[[ranjit [who written book] **Q** bought-A that] rumour] chitra heard-E

c. *[[ranjit [kau **də** liyəpu potə] gatta kieənə] kaṭəkatāwə] chitra æhuve?
[[ranjit [who **Q** written book] bought-A that] rumour] chitra heard-E

Therefore, in modern Sinhala, the Q particle does not enter into any (syntactic) dependency with the *wh*-word. Thus, Q-particles (at least in Sinhala) cannot be conditioned on the basis of whether or not a *wh*-word is present, since the Q particle apparently does not enter into a dependency with the *wh*-word. Cable (2007) suggests a similar situation for Tlingit.

The contexts do differ with respect to whether a disjunction is present and whether the clause is interrogative or not. Therefore—with the exception of Modern Malayalam, where a Wh[] feature is required—the only relevant features are Q[] (a 'Q-particle feature'), Int[] (an 'interrogative' feature), and Junc[] (a 'junction' feature). Since the exact value of these features is irrelevant for my purposes, I indicate valued features simply as '+' (or as '–' where two distinct interpretable values are necessary, as in the case of the interrogative feature Int[ ]).

The follow subsections examine the specific syntactic feature configuration for Sinhala (four stages), Malayalam (two stages), Tlingit, and Japanese.

9.3.2 Modern Colloquial Sinhala

In modern colloquial Sinhala the Q-particle *də* appears obligatorily in all interrogatives; the Q-particle *hari* appears obligatorily in declarative disjunctions; *wh*-based indefinites appear with either *də* or *hari*.[6]

The following constellation of feature-assignments can account (to a large extent) for the distribution of these Q-particles in modern colloquial Sinhala, as shown in Table 9.4.

Categories for which the feature cell of the table is left empty do not bear any features (which are relevant here).

[6]See Chapter 7 above on the difference between the two types of indefinites.

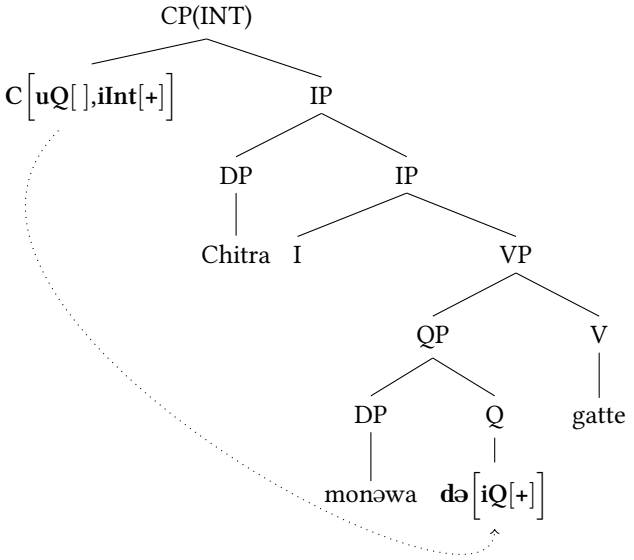
Category	Feature(s)
C-INT	uQ[ ], iInt[+]
***wh*-pronoun**	
də	iQ[+]
hari	
J	

Table 9.4: Modern Colloquial Sinhala feature assignments

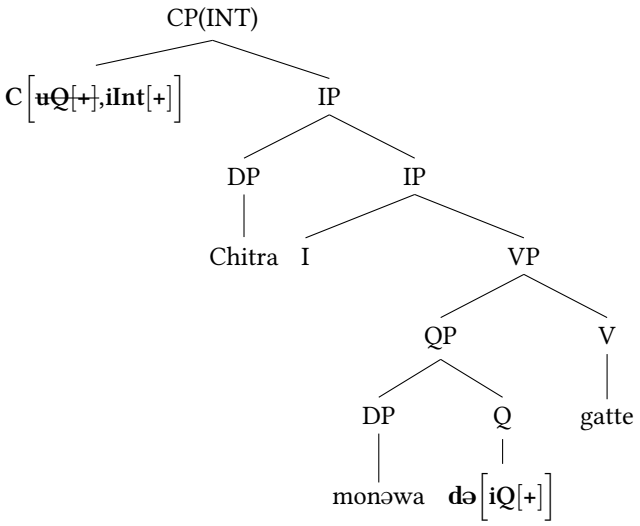
(16) Chitra monəwa də gatte?
Chitra what *də* bought-E
'What did Chitra buy?'

The feature-valuation of (16) is shown below in (17). C, bearing an unvalued uQ[ ] feature, Probes and finds *də* bears a matching iQ[+] feature, allowing an Agree relationship to be established, as shown in (17a). The Agree relationship results in C's uQ[ ] feature receiving a value, as in (17b).

(17) a.



b.

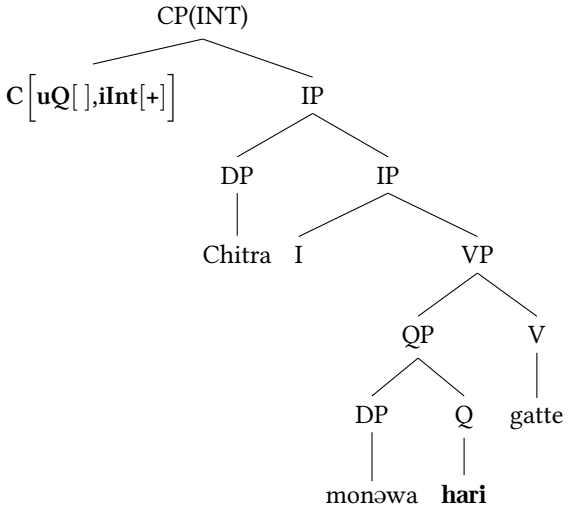


The Q-particle *hari* is disallowed in interrogative contexts, thus a sentence like (18) is ungrammatical.

(18) *Chitra monəwa hari gatte?
Chitra what *hari* bought-E
'What did Chitra buy?'

This ungrammaticality can be derived syntactically as resulting from the failure of the uQ[ ] feature of C to acquire a value. *Hari* does not bear any Q feature and thus when C Probes, it finds no matching feature and remains unvalued, as shown in (19). Unvalued features are uninterpretable at the interface, and thus this lack of a value for C's uQ[ ] feature results in a crash when the CP is sent to Spellout/Transfer.

(19) *

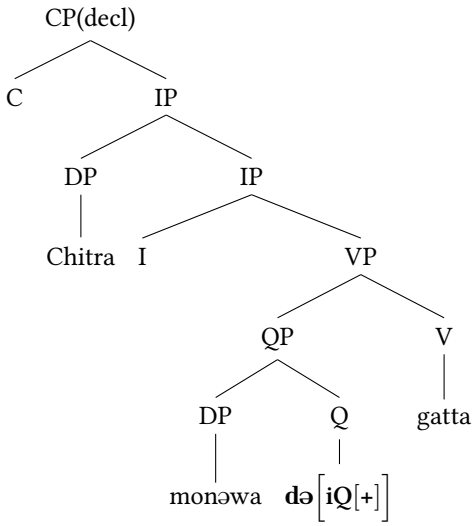


In indefinite contexts, both *də* and *hari* are permissible, see (20), (21), as in fact neither *də* or *hari* nor the declarative C head bear any unvalued features, see (22), (23), respectively.

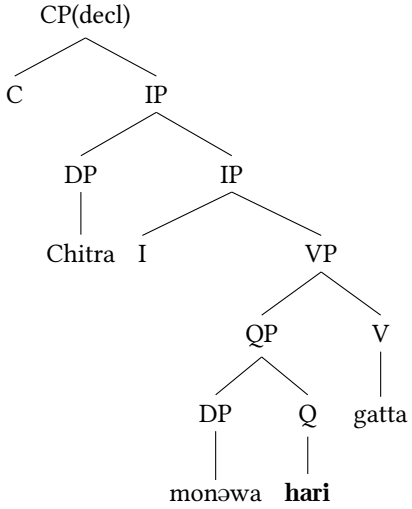
(20) Chitra monəwa də gatta.
Chitra what də bought-A
'Chitra bought something.'

(21) Chitra monəwa hari gatta.
Chitra what hari bought-A
'Chitra bought something.'

(22)



(23)



9.3.3 Modern Literary Sinhala

The relevant feature assignments for Modern Literary Sinhala are shown in Table 9.5.

Category	Feature(s)
C-INT	uQ[ ], iInt[+]
***wh*-pronoun**	
da	iQ[+], uInt[ ]
hō	
J	

Table 9.5: Modern Literary Sinhala feature assignments

With respect to the formal syntactic features of Q-particle structures, Literary Sinhala differs fairly minimally from Colloquial Sinhala. Relevant differences include the use of *hō* rather than *hari*, and the inadmissibility of *da* in

non-interrogative contexts; i.e. *da* cannot be used to form indefinite pronouns. This latter difference is captured by assigning *da* an unvalued uInt[ ] feature, thus requiring the presence of an interrogative C head to value it.

9.3.4 Early Sinhala

In Old and Classical Sinhala, there are no *wh*-based indefinites. Instead, alongside of the possibility of using indefinite NPs (a possibility which still exists in modern Sinhala, see above Section 7.3), we find the monomorphemic elements *kisi* and *yam*.[7]

There is a certain amount of variation in both Old and Classical Sinhala, with respect to the use of *da* in yes/no and *wh*-questions; I posit two grammars to handle this variation. Table 9.6 represents the feature assignments for the grammar in which *wh*- and yes/no questions appear without *da*; the feature assignments for the grammar in which *da* appears in yes/no and *wh*-questions are given in Table 9.7.

Category	Feature(s)
C-INT(wh)	iInt[+]
C-INT(y/n)	iInt[+]
C-INT(alt)	uQ[ ], iInt[+]
***wh*-interrog.**	
indef. pronoun	
da	iQ[+], uInt[ ]
hō	
J	

Table 9.6: Early Sinhala I feature assignments

Category	Feature(s)
C-INT	uQ[ ], iInt[+]
***wh*-interrog.**	
indef. pronoun	
da	iQ[+], uInt[ ]
hō	
J	

Table 9.7: Early Sinhala II feature assignments

As shown by Tables 9.6 and 9.7, this variation is handled by the presence or absence of a uQ[ ] feature on the interrogative C head. I posit that at this stage of Sinhala *da* bears an unvalued uInt[ ] feature, which in effect disallows *da* from appearing in any environment in which there is no element available to value this feature. In the Early Sinhala I grammar, see Table 9.6, the only element available for valuing uInt[ ] is C-INT(alt) and thus *da* only appears in alternative questions in this grammar (and there too obligatorily). In the Early Sinhala II grammar, see Table 9.7,

[7] *Kisi* derives from Sanskrit *kiṁcit* (> Pāli *kiṁci*), which is composed of *kim* 'what' and the particle *cit*—and thus is in fact diachronically a *wh*-based indefinite—but by the time of Sinhala has become a non-analysable monomorphemic element. *Yam* is based on the Old Indo-Aryan *ya*-stem (used to form relatives in relative-correlative structures).

Yam and *kisi* can also co-occur, and thus we find both *kisi-yam* and *yam-kisi*, "some, any". Geiger (1938: §134.2) also cites a 10th-c. instance where *yam* does in fact occur with a *wh*-word (*kavari* "what, which"), but this appears to be a rather marginal construction.

(i) mahaṇ-vannavun atin yam- kavari vatak no gannā isā
yam- what/which hand.INDEF
"To receive nothing whatever from the hand of those entering the order" (10th c. inscription; Zilva Wickremasinghe et al. 1912–1933: I.49[51])

the C-INT category is collapsed (in other words all subcategories of C-INT bear the same set of features) and *da* is therefore allowed and obligatory in all interrogative environments.

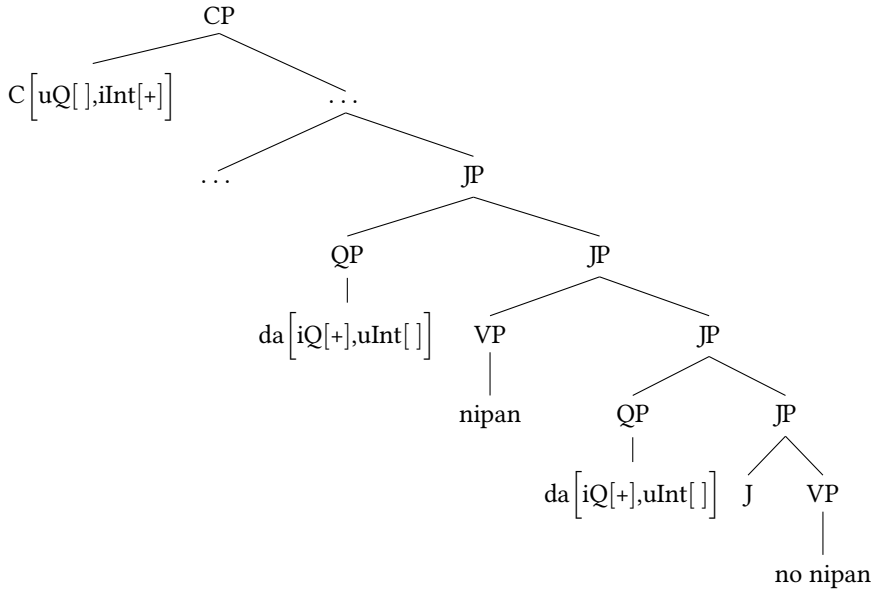
In terms of the syntactic features involved, in regard to the differences between the various stages of Sinhala (with respect to the distribution of Q-particles) note that the changes involved are relatively minimal: in early Sinhala, there is variation in which interrogative heads bear unvalued uQ[ ] features. This variation is settled in favour of all interrogative heads bearing unvalued uQ[ ] features in modern literary Sinhala. The only difference in terms of syntactic features between the system of Q-particles in modern literary Sinhala and modern colloquial Sinhala is that the particle *da*/*də* in modern colloquial Sinhala no longer enters the derivation with an unvalued uInt[ ] feature.

Let us consider the case of alternative questions in Early Sinhala grammar. In this situation the use of the notion of "feature-sharing" (adopted from Pesetsky & Torrego (2007)) is crucial.[8] Consider an alternative question like (24), where we find two Q-particles.

(24) mā ... nuvaṭahu arabhayā kī dǣ nipan **da** no nipan **da**?
my . . . religious mendicant about said things QUOT born ***da*** NEG born ***da***?
"Did my predictions regarding the religious mendicant prove correct or did they not?" (12th century, *Ama.* 178)
(Wijemanne 1984: 75) [CS]

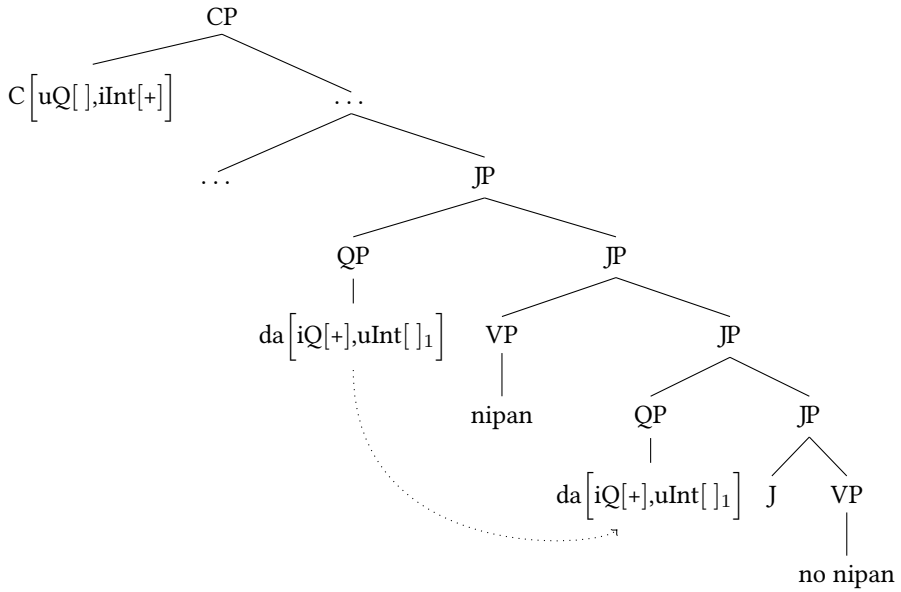
According to the analysis adopted here, each Q-particles bears an unvalued uInt[ ] feature which needs valuing. The process of feature-valuation is shown in (25): first the higher *da* (active due to its unvalued uInt[ ] feature) Probes and Agrees with the lower *da*, establishing a feature-sharing relationship (for uInt[ ]) as shown in (25b); then C Probes and Agrees with the higher *da*, resulting in the valuing of C's uQ feature and the valuation of the higher *da*'s uInt feature—the latter, since it shares its value with that of the lower *da*'s uInt feature, results in the valuation of the lower *da*'s uInt feature as well, as shown in (25c). The valued uninterpreted features then are deleted, as shown in (25d).

(25) a.

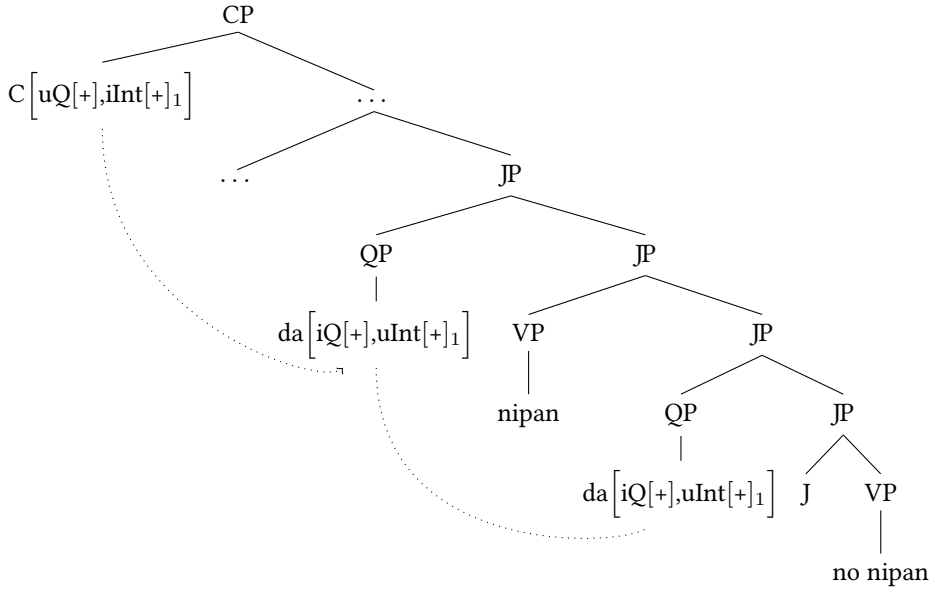


[8]Recall from Chapter 3.1 that this entails conceptualising Agree in terming of feature-sharing, meaning that Agree between two unvalued features is not vacuous: if two unvalued features have established a feature-sharing via Agree, then any subsequent Agree relation which is established between either of the 'shared' unvalued features and a valued instance of that feature will result in both of the unvalued features acquiring a value.

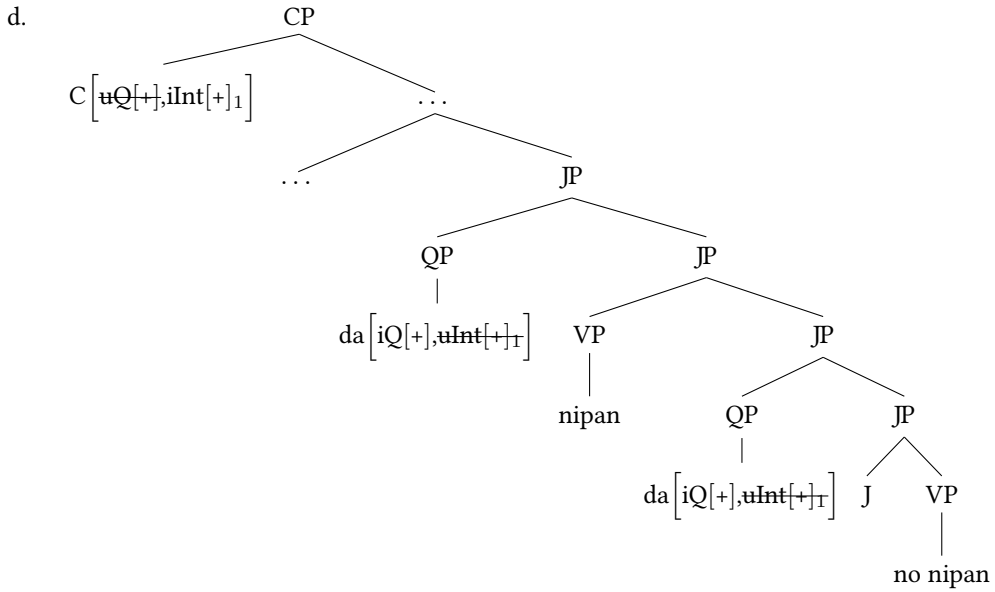
b.


CP
C [uQ[ ],iInt[+]]
. . .
. . .
JP
QP
JP
da [iQ[+],uInt[ ]1]
VP
JP
nipan
QP
JP
da [iQ[+],uInt[ ]1]
J
VP
no nipan

c.


CP
C [uQ[+],iInt[+]1]
. . .
. . .
JP
QP
JP
da [iQ[+],uInt[+]1]
VP
JP
nipan
QP
JP
da [iQ[+],uInt[+]1]
J
VP
no nipan

d.



9.3.5 Old Malayalam

I present the Old Malayalam Q-particle related feature system before the Modern Malayalam system because the former is simpler. In Old Malayalam the particle *-oo* appears in obligatorily yes/no, alternative, and *wh*-questions, in disjunctions, and in the formation of *wh*-indefinites. The feature assignments are shown below in Table 9.8; the Old Malayalam system is thus similar in many respects to the modern colloquial Sinhala system, though it is simpler in that there is only one Q-particle.

Category	Feature(s)
C-INT	uQ[ ], iInt[+]
***wh*-pronoun**	
-oo	iQ[+]
J	

Table 9.8: Old Malayalam feature assignments

9.3.6 Modern Malayalam

The Q-particle valuation system of Modern Malayalam closely resembles that of Old Malayalam, except that in the modern language *-oo* no longer appears in *wh*-interrogatives, which requires treating the interrogative complementiser of *wh*-questions separating from that of other questions, as indicated in Table 9.9, as bearing an additional unvalued feature: a Wh-feature. The only element which can value this feature is the interrogative pronoun, bearing valued iQ[+] and iWh[+] features.[9]

[9]This is accompanied by a lexico-semantic split of *wh*-words from being uniformly semantically-simple elements which could act either as interrogative or indefinite pronouns in Old Malayalam to being either semantically-simple (and serving only as indefinite pronouns) or else being semantically-complex (*wh*-interrogative pronouns). See above Section 9.2.

Category	Feature(s)
C-INT(wh)	uQ[ ], uWh[ ], iInt[+]
C-INT(non-wh)	uQ[ ], iInt[+]
C(decl)	
-oo	iQ[+]
J	
***wh*-words** (sem. simple)	
***wh*-interrog.** (sem. complex)	iQ[+], iWh[+]

Table 9.9: Modern Malayalam features

9.3.7 Tlingit

I posit the following set of feature assignments for Tlingit, shown in Table 9.10

Category	Feature(s)
C-INT	uQ[ ], iInt[+]
***wh*-pronoun**	
sá	iQ[+]
gé	iQ[+], uJunc[ ], uInt[ ]
khach'u	uJunc[ ], iInt[−]
J	iJunc[+], uInt[ ]

Table 9.10: Tlingit feature assignments

The Q-particle *sá* appears obligatorily in *wh*-questions (see example (26)) and is also used to form *wh*-based indefinites (see example (27)), thus similar in some respects to modern colloquial Sinhala *də*.

(26) Daa **sá** aawax̲áa i éesh?
what ***sá*** he.ate.it your father
'What did your father eat?' (Cable 2007: 75)

(27) Kéet ax̲á daa **sá**.
killer.whale he.eats.it what ***sá***
'A killer-whale will eat anything.' (Cable 2007: 66)

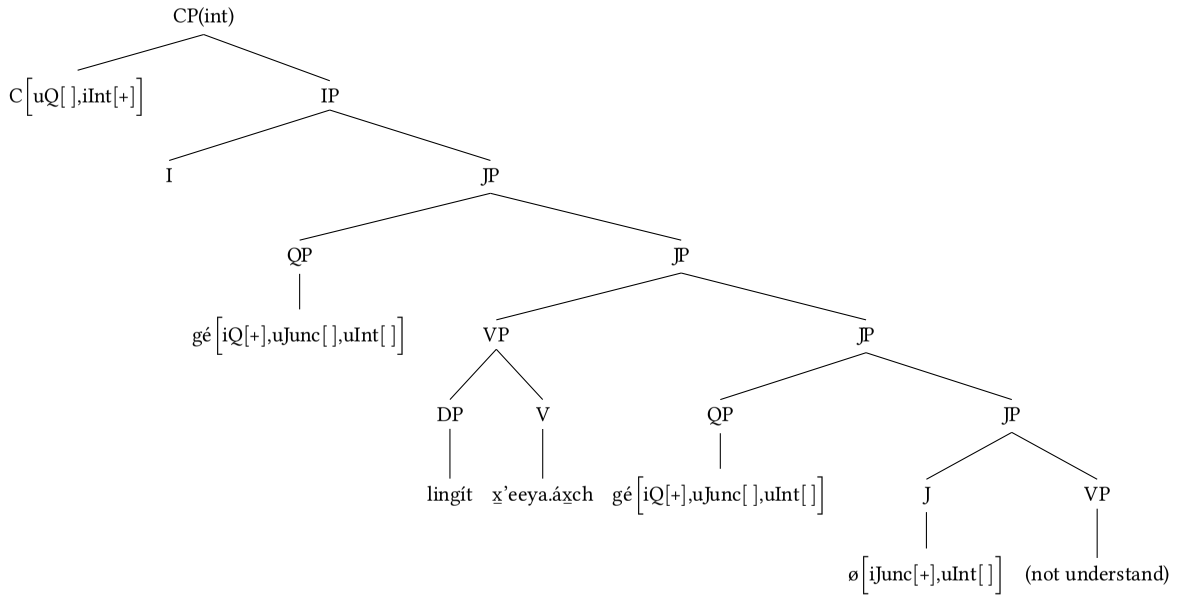
However, a distinct particle, *gé*, is employed in yes/no-questions, as in (28).

(28) Lingít gé x̲'eeya.áx̲ch?
Tlingit *gé* you.understand.it
"Do you speak Tlingit?" (Cable 2007: 74n40)

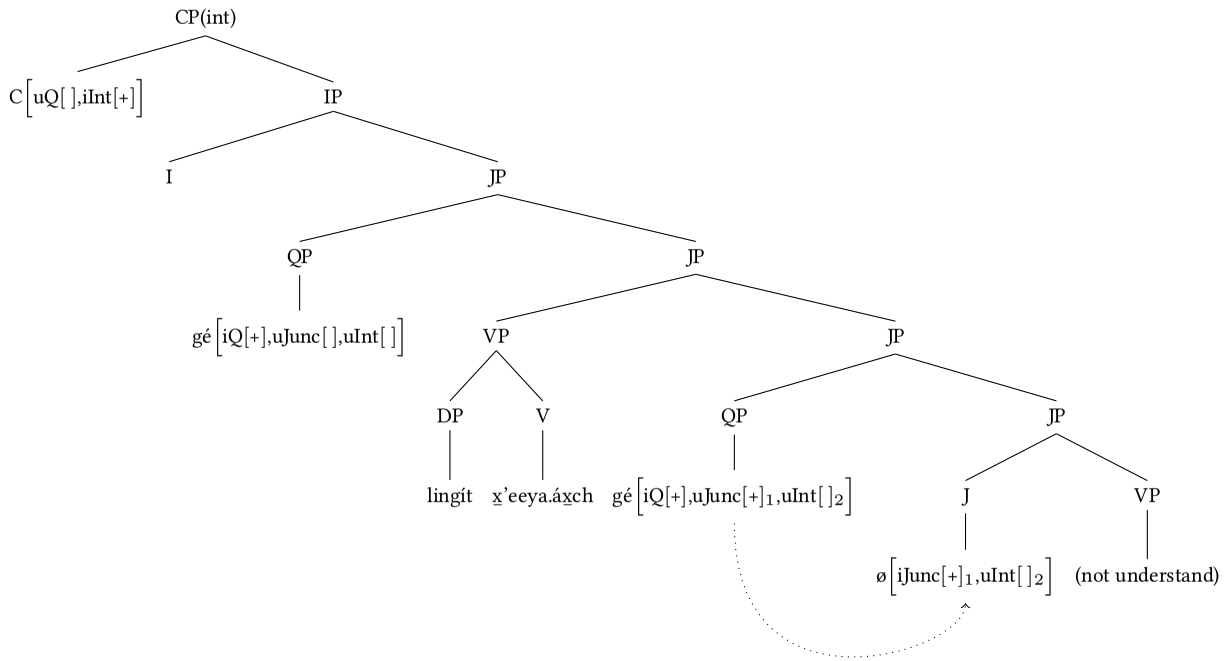
The distinct feature assignments for *gé* and *sá* predict the complementary distribution of these two particles: *sá* bears no unvalued features and is thus possible in *wh*-indefinites (where no Agree operations are necessary) and in *wh*-interrogatives (where C-INT simply requires a valued Q feature, which *sá* provides), while *gé* bears both an unvalued uJunc[ ] and an unvalued uInt[ ] feature, rendering it admissible only in interrogatives contexts containing a disjunction.

Assuming the above feature assignments, the derivation of feature assignments for (28) is shown below in (29)–(33), following a pattern similar to that observed for Early Sinhala alternative questions shown above in (25).

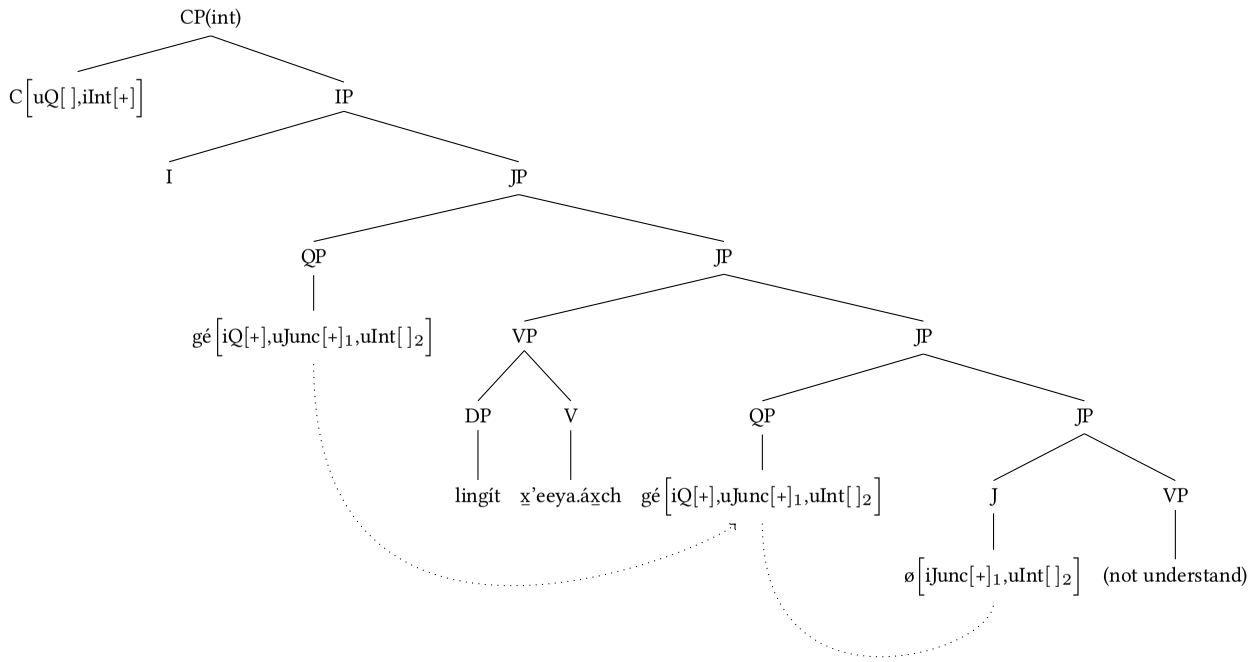
(29)


CP(int)
C [uQ[ ],iInt[+]]
IP
I
JP
QP
JP
gé [iQ[+],uJunc[ ],uInt[ ]]
VP
JP
DP
V
QP
JP
lingít
x̲'eeya.áx̲ch
gé [iQ[+],uJunc[ ],uInt[ ]]
J
VP
ø [iJunc[+],uInt[ ]]
(not understand)

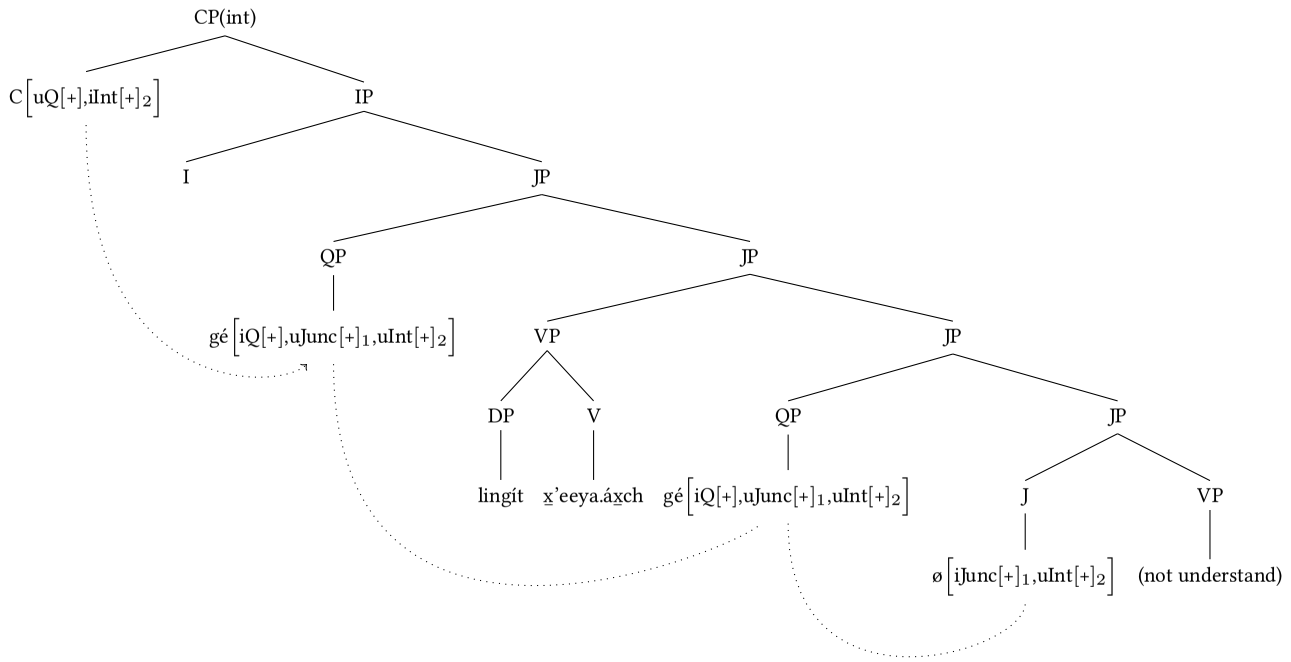
(30)


CP(int)
C [uQ[ ],iInt[+]]
IP
I
JP
QP
JP
gé [iQ[+],uJunc[ ],uInt[ ]]
VP
JP
DP
V
QP
JP
lingít
x̲'eeya.áx̲ch
gé [iQ[+],uJunc[+]1,uInt[ ]2]
J
VP
ø [iJunc[+]1,uInt[ ]2]
(not understand)

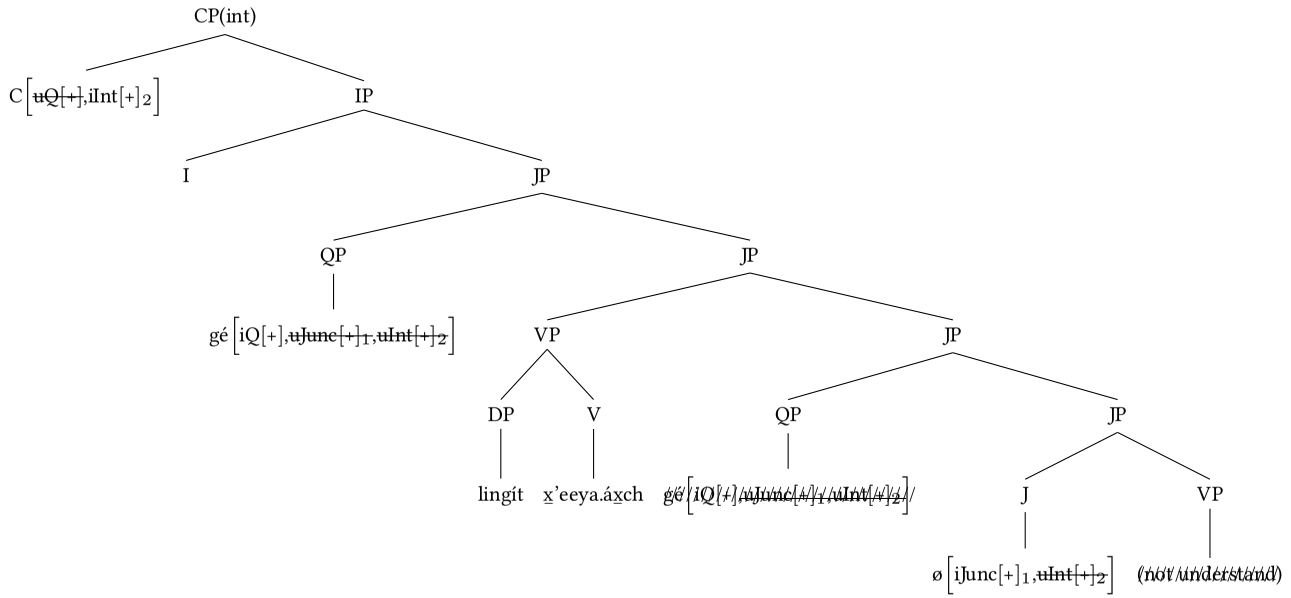
(31)


CP(int)
C[uQ[ ],iInt[+]]
IP
I
JP
QP
JP
gé[iQ[+],uJunc[+]1,uInt[ ]2]
VP
JP
DP
V
QP
JP
lingít
x̲'eeya.áx̲ch
gé[iQ[+],uJunc[+]1,uInt[ ]2]
J
VP
ø[iJunc[+]1,uInt[ ]2]
(not understand)

(32)


CP(int)
C[uQ[+],iInt[+]2]
IP
I
JP
QP
JP
gé[iQ[+],uJunc[+]1,uInt[+]2]
VP
JP
DP
V
QP
JP
lingít
x̲’eeya.áx̲ch
gé[iQ[+],uJunc[+]1,uInt[+]2]
J
VP
ø[iJunc[+]1,uInt[+]2]
(not understand)

(33)


CP(int)
C [~~uQ[+]~~ , iInt[+]2]
IP
I
JP
QP
JP
gé [iQ[+] , ~~uJunc[+]1~~ , ~~uInt[+]2~~]
VP
JP
DP
V
QP
JP
lingít
x̲’eeya.áx̲ch
gé [iQ[+] , ~~uJunc[+]1~~ , ~~uInt[+]2~~]
J
VP
ø [iJunc[+]1 , ~~uInt[+]2~~]
(not understand)

In step (30), the lower *gé* Probes and Agrees with J, valuing its uninterpretable uJunc[ ] feature, additionally a feature-sharing relationship is established between the unvalued uInt[ ] features of *gé* and J. Next, as shown in (31), the higher *gé* Probes and Agrees with the lower *gé*, picking up the shared value for the Junc feature, and extending the feature-sharing relationship of the uInt feature. Finally, in (32), the interrogative C head Probes and Agrees with the higher *gé*, valuing its own uQ[ ] feature and providing a value for the shared uInt[ ] feature of the other elements.

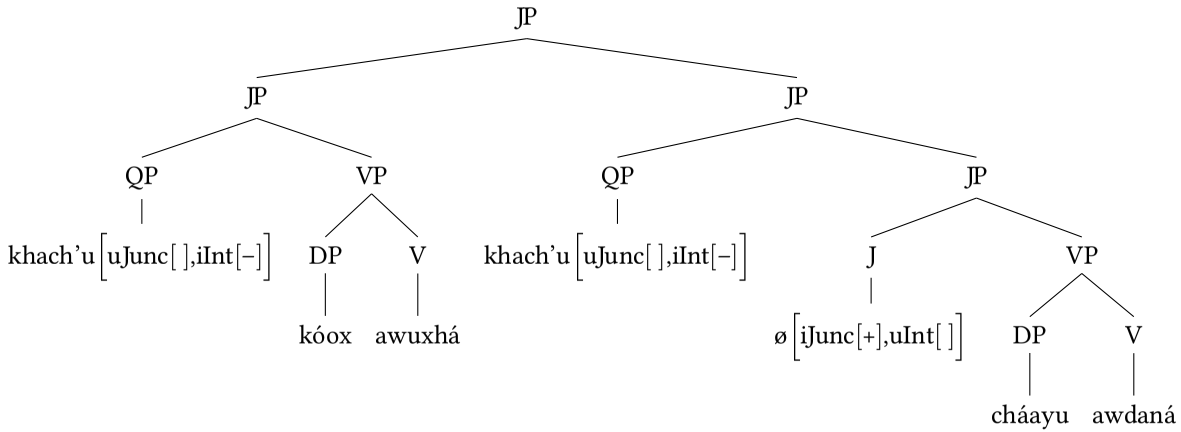
The grammaticality of (28) is therefore correctly predicted (step (33) shows the apparently obligatory elision of the *or not. . .* constituent). The Q-particle *sá* is also correctly predicted to be ungrammatical in yes/no-questions given that it lacks a uJunc[ ] feature, and thus J would remain with an unvalued uInt[ ] feature since no constituent would enter into an Agree relationship with it.

Tlingit shows a further complication in that yet another Q-particle, *khach'u*, appears in declarative disjunctions, as shown in (34)

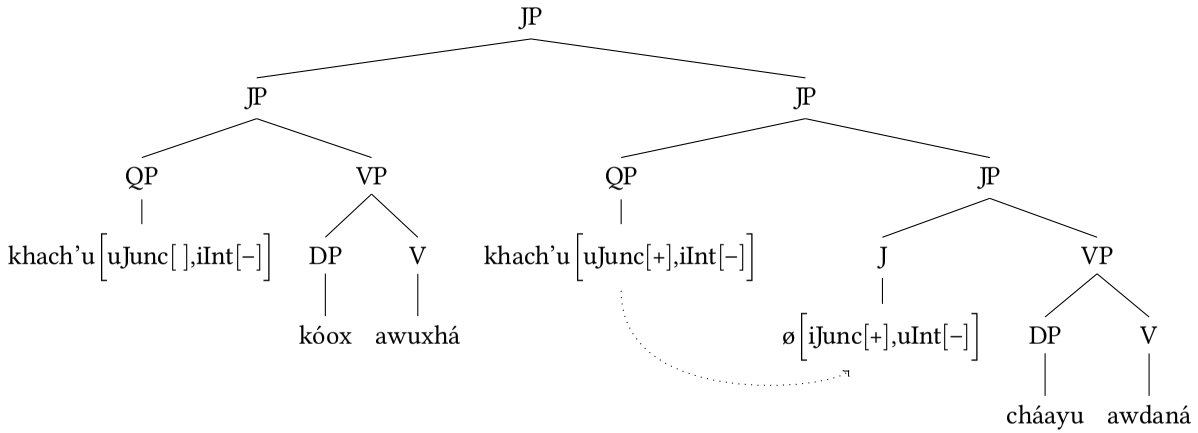
(34) Tlél aadóoch sá kóox awuxhá khach'u cháayu awdaná.
not who.ERG *sá* rice ate *khach'u* tea drank
"Nobody ate rice or drank tea." (Seth Cable, p.c.)

The feature valuation for (34) is shown below in (35)–(38).

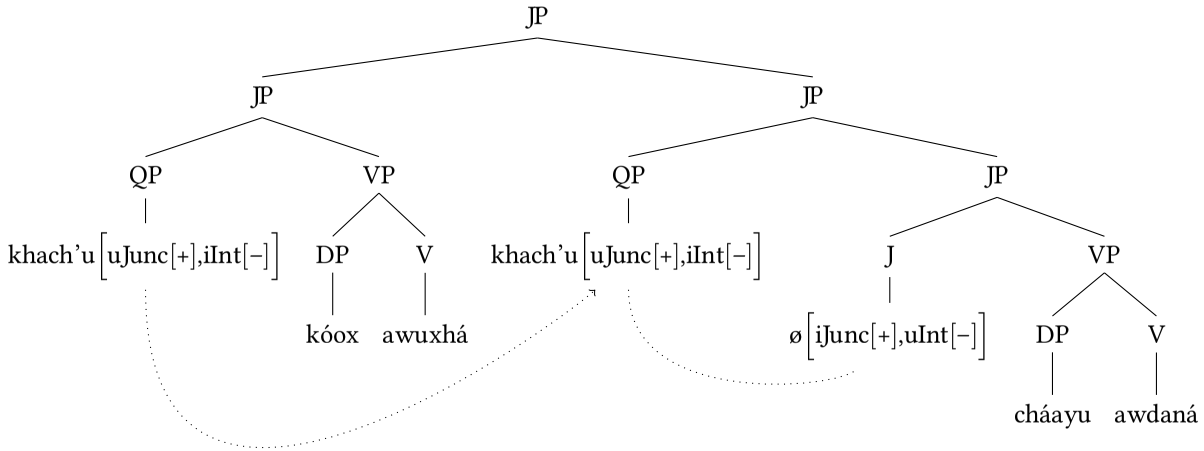
(35)


JP
JP
JP
QP
VP
QP
JP
khach'u [uJunc[ ],iInt[−]]
DP
V
khach'u [uJunc[ ],iInt[−]]
J
VP
kóox
awuxhá
ø [iJunc[+],uInt[ ]]
DP
V
cháayu
awdaná

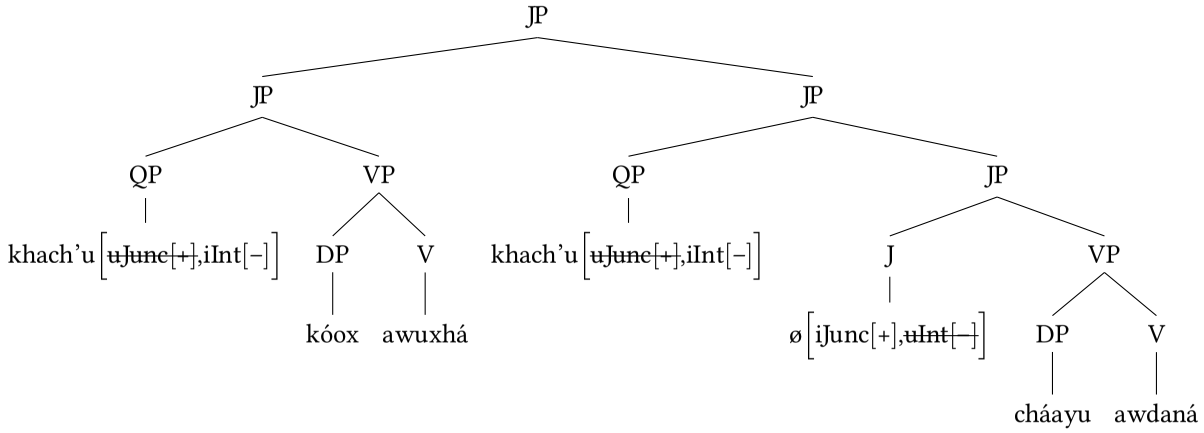
(36)


JP
JP
JP
QP
VP
QP
JP
khach'u [uJunc[ ],iInt[−]]
DP
V
khach'u [uJunc[+],iInt[−]]
J
VP
kóox
awuxhá
ø [iJunc[+],uInt[−]]
DP
V
cháayu
awdaná

(37)


JP
JP
JP
QP
VP
QP
JP
khach'u [uJunc[+],iInt[−]]
DP
V
khach'u [uJunc[+],iInt[−]]
J
VP
kóox
awuxhá
ø [iJunc[+],uInt[−]]
DP
V
cháayu
awdaná

(38)


JP
JP
JP
QP
VP
QP
JP
khach'u [~~uJunc~~[+],iInt[−]]
DP
V
khach'u [~~uJunc~~[+],iInt[−]]
J
VP
kóox
awuxhá
ø [iJunc[+],~~uInt[−]~~]
DP
V
cháayu
awdaná

Here *khach'u* crucially bears a valued iInt[−] feature, thus allowing J's uInt[ ] feature to be valued. Neither *sá* or *gé* are admissible in this syntactic context since either would result in J bearing an unvalued uInt[ ] at the interface.[10]

9.3.8 Japanese

In Japanese only *ka* occurs in the full range of possible Q-particle environments. None of the other Q-particles can occur in declarative disjunctions or in the formation of *wh*-based indefinites. The Q-particle *ndai* is restricted solely to *wh*-questions, and the particles *kai* and *kadooka* are restricted to yes/no-questions.

I propose the following set of features for Japanese:

Category	Feature(s)
C-INT	uQ[ ], uInt[+]
***wh*-pronoun**	iWh[+]
ka	iQ[+]
no	iQ[+], uInt[ ]
ndai	iQ[+], uInt[ ], uWh[ ]
kai*/*kadooka	iQ[+], uInt[ ], uJunc[ ]
J	iJunc[+]

Table 9.11: Japanese feature assignments

The interrogative C-head requires the presence of some Q-particle due to its unvalued uQ[ ] feature. The Q-particle *kai* and *kadooka* can occur only in yes/no-questions due to their unvalued uInt[ ] and uJunc[ ] features; while *ndai* is restricted to *wh*-questions due to its unvalued uInt[ ] and uWh[ ] features. The particle *no* is restricted to interrogatives due to its unvalued uInt[ ] feature. Only *ka* may occur in declarative contexts, due to the fact that bears no uInt[ ] feature which needs valuing.

There are additional restrictions on the distribution of these particles which concern politeness-marking and matrix/embedded clause restrictions which I do not treat here.[11]

Additionally, Yoshida & Yoshida (1996) and Hagstrom (1998) report that in informal speech, *wh*- and yes/no-questions can occur without any Q-particle (see also Ko 2005). This suggests the co-existence of a grammar in which C-INT bears no unvalued Q feature (and thus does not require the presence of a Q-particle).[12]

The use of formal syntactic features can thus serve to constrain the environments in which certain Q-particles may appear in particular languages. In some cases there may be additional pragmatic constraints on Q-particles, including issues of politeness, as in Japanese. Another, rather more complicated pragmatically-based constraint on Q-particle in Sinhala is discussed in the following section.

[10] Tlingit alternative questions involve complexities I do not attempt to explain here. See the example below in (i).

(i) Káxwei gé i tuwáa sigóo, cháau gwáa, héen gwáa?
coffee *gé* you.want, tea *gwáa*, water *gwáa*
"Do you want coffee, or tea, or water?" (Seth Cable, p.c.)

(i) probably is actually three matrix questions: "Is it coffee you want?" "Or else is it tea?" "Or else is it water?"

[11] Specifically, only *ka* may be appear in both matrix and embedded clauses, while *no*, *ndai*, and *kai* are restricted to matrix clauses and *kadooka* to embedded clauses. *Ka* is the most polite form, with *no* being of middling politeness, and *kai* and *ndai* occurring in informal speech. See further Ginsburg (2009).

[12] Such questions, however, can differ semantically from their Q-particle-ful counterparts: see further Hagstrom (1998), esp. chapter 6 and following chapters.

9.4 Defining Q-particle environments IV: pragmatic restrictions

Recall that the inadmissibility of *da/də* in affirmative disjunctions cannot be explained on the basis of formal feature specifications. The pragmatics of *də* required for the explanation of the distribution of *də* and *hari* indefinites (in terms of intensionally- and extensionally-unknown indefinites, discussed above in Chapter 7) actually provides the solution.

Let us first consider the acceptable affirmative declarative, formed with *hari*, as in example (39):

(39) Gunəpālə hari Chitra hari gaməṭə giyā.
Gunapala *hari* Chitra *hari* village.DAT went-A
"Gunapala or Chitra went to the village."

Here the speaker asserts:

(40) $\forall w \in F.\exists f.\text{BASISCH}(f)[f(w)(\{\text{Gunapala, Chitra}\})$ went to the village in $w]$

The speaker also presupposes that:

(41) $\neg\big[\exists f.\text{BASICH}_{\sim\text{ImpPred}}(f)\ \neg[\exists w', w'' \in F$:
$f(\{\text{Gunapala, Chitra}\})(w')$ went to the village in $w'=1$ &
$f(\{\text{Gunapala, Chitra}\})(w'')$ went to the village in $w''=1$ &
$f(\{\text{Gunapala, Chitra}\})(w') \neq f(\{\text{Gunapala, Chitra}\})(w'')]\big]$

This means that the speaker has no means of singling out an individual that satisfies the proposition in all epistemically-accessible worlds. This is an eminently reasonable pragmatic signal for this sort of disjunction.

Now consider the unacceptable affirmative declarative disjunction formed with *də*:

(42) *Gunəpālə də Chitra də gaməṭə giyā.
Gunapala *də* Chitra *də* village.DAT went-A
"Gunapala or Chitra went to the village."

Just as in (39) above, the speaker asserts that:

(43) $\forall w \in F.\exists f.\text{BASISCH}(f)[f(\{\text{Gunapala, Chitra}\})$ went to the village in $w]$

This means that in all epistemically-accessible worlds it is either the case that Gunapala went to the village or Chitra went to the village:

(44) a. $\forall w \in F.\exists f \in G\ [f(\{\text{Gunapala, Chitra}\})$ went to the village in $w] \equiv$
b. $\forall w \in F[[\text{Chitra went to the village in } w] \vee [\text{Gunapala went to the village in } w]]$

However, now the speaker also presupposes that:

(45) $\exists w' \in F \neg\big[\exists f.\text{BASICH}_{\sim\text{ImpPred}}(f)$:
$f(\{\text{Gunapala, Chitra}\})(w')$ went to the village in $w'=1\big]$

This entails that there is some epistemically-accessible world for which the speaker can assert neither that Chitra went to the village or Gunapala went to the village, therefore allowing for the possibility that neither Chitra nor Gunapala

went to the village in that world. In other words (44) (=(43)) asserts $\Box[p \vee q]$, and (45) signals that $\Diamond[\neg p \wedge \neg q]$. Thus (45) contradicts (44).[13]

Why does this contradiction not also arise in the case of WH+*də* indefinites? Consider (46).

(46) kau də gamaṭa giyā.
who *də* village.DATA went-A
"Someone went to the village."

Here, the speaker asserts that:

(47) $\forall w \in F.\exists f.\text{BASISCH}(f).f(\{x \in D_e \mid x \in \text{human}' \text{ in } w\})$ went to the village in w

Let us assume for the sake of exposition that human′={Gunapala, Chitra, Ranjit}. In this case, since the speaker does not actually (necessarily) know the content of the set denoted by *kau* "who"—he knows only that $[\![\text{kau}]\!]^g=\{x \in D_{\langle se \rangle} \mid x \in \text{human}'\}$, but he may not know for which values of x human′(x) is true—it is not the case that he necessarily knows for which values of f f({Gunapala, Chitra, Ranjit}) is true, and thus he asserts only (47).

Sentence (46) presupposes that:

(48) $\exists w' \in F \neg \big[\exists f.\text{BASICH}_{\sim \text{ImpPred}}(f)$:
$f(\{x \in D_{\langle se \rangle} \mid x \in \text{human}'\}(w')$ went to the village in $w'=1\big]$

Here (48) does not contradict (47) since here the speaker only asserts that there are values of f for which f({x ∈ $D_{\langle se \rangle}$ | x ∈ human′}(w′) satisfies the proposition, but does not assert that he knows which values these are.

Thus the pragmatics of *də* explains why it is inadmissible in declarative disjunctions, providing the last piece of the puzzle of how to account for the distribution of Q-particles in Sinhala.

9.5 Summary & Conclusions

This chapter provides an account of the crosslinguistic distribution of Q-particles in Tlingit, Japanese, and various stages of Sinhala and Malayalam. The data analysed here highlight the necessity of considering various components of the grammar (semantics, pragmatics, syntax) for a complete account of complex linguistic phenomena, such as the distribution of Q-particles.

The semantic analysis of Q-particles as crosslinguistically bearing a uniform denotation as variables over choice functions provides a straightforward account of why, crosslinguistically, Q-particles tend to appear not only in interrogatives, but also in (non-interrogative) disjunctions and indefinites, if we treat all of these environments as involving an element with an Hamblin-type set denotation—the type of denotation to which choice-functions, as defined here, apply.

[13] One might wonder why *də* then cannot appear in a disjunction involving two indefinite NPs, e.g. "A boy or a girl went to the village" as in example (i), where the speaker thus may know of no individual concept whose unique extension satisfies the existential claim ("boy" and "girl" are not uniquely identifying individual concepts), a configuration which would seem to involve no contradiction with the pragmatics of *də*.

(i) *pirimi-lamayek də gæhænu-lamayek də gamə̀ṭə giyā.
boy.INDEF *də* girl.INDEF *də* village.DAT go.PAST.A
"A boy or a girl went to the village."

Recall, however, that də indefinites, though intensionally-unknown, are specific indefinites (see above Chapter 7.8), and thus *də* is incompatible with the non-specific reading of the indefinites which would be required if the speaker lacked some uniquely identifying individual concept about who went to the village.

Crosslinguistic variability in the distribution of Q-particles, including in languages employing multiple Q-particles, can be largely accounted for in terms of differences in formal syntactic feature assignments between languages. This includes not only differences in the feature specifications of Q-particles, but also differences in the feature specifications of other elements with which Q-particle interact, such as the heads of interrogative CPs.

In some cases the analysis of the distribution of Q-particles requires the consideration of pragmatics. In the case considered here, namely the question of how to correctly rule out *də* from appearing in non-interrogative disjunctions, we found that the presuppositions assigned to *də* and *hari* with respect to their status in forming epistemic indefinites (see Chapter 7) also correctly predict the inadmissibility of *də* in non-interrogative disjunctions (but allows, correctly, for the appearance of *hari* in this context).

The restriction on Q-particles requiring that the choice-functions which they denote may only be applied to elements of the "//" type (Hamblin-type sets) accounts for the basic pattern of the distribution of Q-particles: they are only found in environments containing Hamblin-type elements (e.g. *wh*-words, disjunctions); see Section 9.1 above. Differences in lexico-semantic properties of *wh*-words—i.e. whether they are semantically-simple (denoting only Hamblin-type sets) or semantically-complex (in essence, incorporating a Q-particle)—can account for the absence of Q-particles from *wh*-interrogatives (and indefinites) in certain languages, even languages which employ Q-particles in other contexts (e.g. modern Malayalam), as shown in Section 9.2.[14] Differences in formal syntactic feature specifications account for the majority of the crosslinguistic and historical differences in the distribution of Q-particles across different contexts, as discussed above in 9.3. Finally, in a limited set of contexts, the pragmatics associated with the presuppositions borne by particular Q-particles may also serve to restrict the distribution of particular Q-particles, as is the case for Sinhala *də* and *hari*, discussed above in Section 9.4.

The historical evidence afforded by Sinhala (see above Section 9.3.4) also supports the analysis advanced in this thesis, namely that crosslinguistically Q-particles can be given a semantically unified analysis as variables over choice functions, while crosslinguistic differences can be accounted for largely in terms of differences of formal syntactic features. The changes in the distribution of *da*/*də* and *hō*/*hari* in Sinhala from the 8th century to the present day can be largely accounted for in terms of fairly minimal changes in formal syntactic features associated with Q-particles and interrogative complementisers. If we treated Sinhala *da* in *wh*-interrogatives as an element separate from (though homophonous with) the *da* which appears in alternative- and yes/no-questions—as suggested by Cable (2007: 74–75n40), who says "given the distinction between *gé* and *sá* in Tlingit, I assume that the use of *da*/*ka* in Sinhala/Japanese polar questions reflects the existence of a separate, homophonous 'yes/no' particle"—then we would have no direct way of connecting the early use of *da* in alternative and yes/no questions with the later appearance of *da* in *wh*-interrogatives, and the still later appearance of *da* in *wh*-indefinites. Here we find that diachronic evidence thus plays an important role in deciding between possible synchronic analyses (see further Chapter 11 for more discussion).

This concludes the basic account of the crosslinguistic distribution of Q-particles. The next chapter returns to the topic of Q-particles in disjunctions, and provides a more in-depth examination of the structure of both the syntax and semantics of disjunction, as well as the possible extension of a Q-particle-based analysis to the treatment of conjunction.

[14] Even English, I argue, employs Q-particles in disjunctive contexts. See Chapter 10.

Chapter 10

The syntax and semantics of (dis)junction

> '. . . When you come closer, you will then define it as an animal, even if you do not yet know if it is a horse or an ass. And finally, when it is still closer, you will be able to say it is a horse even if you do not yet know whether it is Brunellus or Niger. . . '
>
> —William of Baskerville to his pupil Adso of Melk,
> in Umberto Eco's *The name of the rose*

Adopting a semantically-unified analysis of Q-particles, as uniformly denoting variables over choice functions in all contexts, requires a novel analysis of the structure of disjunction. In this chapter I provide an examination of the details of such an analysis.

The semantic treatment of Q-particles as variables over choice-functions can thus account for their appearance in both *wh*-interrogatives and *wh*-indefinites as discussed previously. What of Q-particles in yes/no and alternative questions, as in (1) and (2), respectively?[1]

(1) Gunəpālə gamaṭa giya **də**?
Gunapala village.DAT went.A ***də***
"Did Gunapala go to the village?"

(2) Gunəpālə **də** Chitra **də** gamaṭa giye?
Gunapala ***də*** Chitra ***də*** village.DAT went.E
"Was it Gunapala or Chitra who went to the village?"

I argue that Q-particles like Sinhala *də* play the same role in disjunction that they do in interrogatives and indefinites: they are variables over choice functions. This entails that disjunctive structures, like *wh*-words, involve Hamblin-type sets. The treatment of disjunctions as involving Hamblin-type sets is motivated on independent grounds, as shown by Alonso-Ovalle 2006.[2]

[1] The difference in the marking on the verb in (1) and (2) is the result of the difference between the two sentences in whether or not there is a focussed element in the c-command domain of the verb. Example (1) has no focussed element in the c-command domain of the verb, while (2) does. See above Chapter 3.6 for further discussion.

[2] Specifically, Alonso-Ovalle (2006) shows that:
(1) A standard semantic treatment of disjunction fails to capture the natural interpretation of counterfactual conditionals which involve disjunctive antecedents, predicting that such counterfactuals are evaluated by selecting the closest worlds from the union of the propositions that *or* operates over; whereas the natural interpretation requires the selection of the closest worlds from *each* of the propositions that *or* operates over—an interpretation predicted under a Hamblin-style semantic analysis if conditionals are analysed as correlative constructions.
(2) A standard semantic treatment of disjunctions under the scope of modals incorrectly predicts that a sentence like *John may leave or stay* is true so long as John has at least one of the rights (to leave or to stay); whereas a Hamblin-style treatment of disjunction allows for the correct derivation (as an implicature of domain widening) that such a sentence is true iff John has both rights (to leave or to stay).
(3) A Hamblin-style analysis is better equipped to handle unembedded disjunctions with an exclusive component: where the exclusive component of a disjunction S (with more than two atomic disjuncts) can be derived as an implicature if S competes in the pragmatics with all of the conjunctions that can be formed of the atomic disjuncts—the generation of the pragmatic competitors is difficult under a standard analysis of *or* since the interpretation system does not have access to the atomic disjuncts.

However, that Q-particles act as variables over choice-functions in disjunctions is perhaps not obvious from their surface syntax, e.g. *də* appears *after* each of the disjuncts, as in (3).

(3) gunəpaalə də chitra də gaməṭə giye?
Gunapala *də* Chitra *də* village.DAT went-E
"Was it Gunapala or Chitra who went to the village?"

Examples like (3) (typical not only of Sinhala, but also of other languages like Japanese and Malayalam) raise two, related questions: (i) if *də* acts as a choice function and applies to a set, then how is this set created?; (ii) how can *də* apply to the set, since it appears in the surface syntax in a structurally lower position than the set itself (=the entire disjunction)? To put the question another way: how is it that Q-particles like *də* can act as choice functions in disjunctive structures when it appears to be *də* itself that acts as the disjunction?

I will argue that (i) Q-particles themselves do not act as disjunctions, rather the actual disjunction is an unpronounced element (a 'junction' element, J, which heads its own projection, JP) whose semantic function is to create a Hamblin-type set; (ii) Q-particles in disjunctive structures actually originate in positions which c-command the disjunct, with PF-level rules accounting for the post-disjunct positioning in surface syntax (unsurprising due to their status as enclitics).

10.1 Evidence for a category J(unction)

English itself furnishes evidence for such an analysis. In English *either*. . . *or* constructions, as den Dikken (2006) points out, *either* does not always occur on the edge of the leftmost disjunct; rather it can apparently 'float' to positions structurally lower, see (4), *or* structurally higher, see (5), than the leftmost disjunct's edge.

(4) a. Either [John ate rice] or [he ate beans].
b. [John **either** ate rice] or [he ate beans]. [*either* too low]

(5) a. John ate either [rice] or [beans].
b. John **either** ate [rice] or [beans]. [*either* too high]
c. **Either** John ate [rice] or [beans]. [*either* too high]

Based on this evidence, den Dikken (2006) argues that neither *either* nor *or* (nor *both* or *and*) are themselves the lexicalisation of the (dis/con)junction, but rather are phrasal categories which adjoin directly to their (dis/con)junct (or to a node on the θ-path projected from the contrastive focus). He proposes that the actual (dis/con)junction is an unpronounced head J ('junction').

Den Dikken's syntactic evidence is persuasive, but the question arises: if elements like *either* and *or* are not themselves lexicalisations of the actual disjunction, then what is their semantic function? I suggest that elements like English *or* are in fact Q-particles, and that, like other Q-particles, their semantic function is that of a variable over choice functions, which apply to Hamblin-type sets.[3]

Alonso-Ovalle's (2006) Hamblin-style analysis of disjunction involves a very simple disjunction rule, (6).

(6) Where $[\![B]\!]^g, [\![C]\!]^g \subseteq D_\tau$, $\left[\!\!\left[\begin{array}{ccc} & A & \\ B & or & C \end{array} \right]\!\!\right]^g \subseteq D_\tau = [\![B]\!]^g \cup [\![C]\!]^g$

There are at least two difficulties which the straightforward adoption of (6) would present for a unified semantic analysis of Q-particles like Sinhala *də*: (a) Alonso-Ovalle (2006) takes *or* (∼ Sinhala *də*) itself to be the disjunction operator, and (b) his formulation involves non-binary branching structures.

Taking the 'junction' operator in English (and Sinhala etc.) to be an unpronounced J, as argued by den Dikken (2006), we can reformulate (6) as in (7).

(7) 'Junction rule' (1st version):

Where $[\![B]\!]^g, [\![C]\!]^g \in D_\tau$, $\left[\!\!\left[\begin{array}{ccc} & JP_a & \\ B & & JP_b \\ & J \quad C & \end{array} \right]\!\!\right]^g \subseteq D_\tau = [\![B]\!]^g \cup [\![C]\!]^g$

10.2 Maintaining a choice-functional analysis of Q-particles

However, (7) requires that, like Alonso-Ovalle (2006) we adopt the position that all elements bear Hamblin-type denotations (e.g. $[\![\text{John}]\!]^g = \{\text{John}\}$, $[\![\text{smoke}]\!]^g = \{\lambda x.x \text{ smokes}\}$ etc.), rather that the mixed Montagovian/Hamblin analysis argued for here. If we were to adopt a Hamblin-semantics-all-the-way-down analysis, we would lose the crosslinguistic correlation between a restricted set of constructions/elements (i.e. indeterminate pronouns, disjunction) and the appearance of Q-particles, since, if all elements bear Hamblin-type denotations, then, in theory, a Q-particle could apply to any type of element. However, in fact, a choice-functional analysis of Q-particles is

[3]That *or* is a Q-particle in English is also suggested by the fact that can act as a Q-particle in yes/no questions in earlier English, both direct, (i) and indirect, (ii) (examples cited from the *OED* (Murray et al. 2011):

(i) **Or** not in wrathefulnesse of hym is lettid þe sunne, & oo dai maad as two?
"Was not the sun stopped in his anger, and one day made as two?" [Ecclus. 46.5 (Wycliffe Bible, ca. 1382)]

(ii) He asked the lordes. . . **or** they wolde therfore warre.
"He asked the lords if they would therefore go to war." [Virgilius sig. aiiij^{v}, ca. 1518]

Or may also appear preceding each of the disjuncts in alternative questions in earlier English, again, in both direct, (iii), and indirect, (iv), questions.

(iii) How kenst thou, that he is awoke? **Or** hast thy selfe his slomber broke? **Or** made preuie to the same?
"How do you know that he is awake? Have you woken him yourself, or have you been made privy (to the knowledge that he is awake)?"
[Spenser, *Shepheardes Cal.*, Mar. 29, 1579]

(iv) Tell me where is fancie bred, **Or** in the hart, **or** in the head.
"Tell me where is fancy bred: whether in the heart, or in the head." [Shakespeare, *Merchant of Venice* III. ii. 64, 1600]

See also Jayaseelan (2008), who makes a similar observation.

simply incompatible with a Hamblin-semantics-all-the-way-down analysis, given that the pointwise composition rule Hamblin employs would require the incorrect analysis shown in (8).

(8) $\exists f \in D_{cf}.\{f\}(\{A, B, C\}) = \{f(A), f(B), f(C)\}$

Since choice-functions can apply only to sets, (8) results in an undefined denotation.

Therefore, the choice-functional approach to Q-particles requires the mixed Montague/Hamblin approach discussed above. This in turn entails that part of the denotation of J must involve the transformation of Montagovian elements into Hamblin-type elements. Adopting this revision results in the translation of J as in (9).

(9) 'Junction rule' (2nd version):

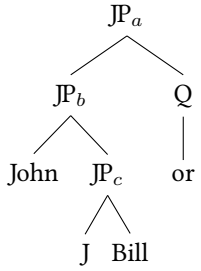
Where $[\![B]\!]^g, [\![C]\!]^g \in D_\tau$, $\left[\!\!\left[\begin{array}{ccc} & JP_a & \\ & \diagup\ \diagdown & \\ B & & JP_b \\ & & \diagup\ \diagdown \\ & & J\ \ C \end{array} \right]\!\!\right]^g \subseteq D_{\langle \tau, t \rangle} = \{[\![B]\!]^g\} \cup \{[\![C]\!]^g\}$

The rule in (9) correctly derives the denotation of *John J Bill* as shown in (10).

(10) $[\![\text{John J Bill}]\!]^g = \{\text{John,Bill}\}$

If we then assume that *or* is a Q-particle, semantically realised as a variable over choice-functions, and that it adjoins to the entire JP, as shown in (11), then we derive the semantic representation in (12).

(11)

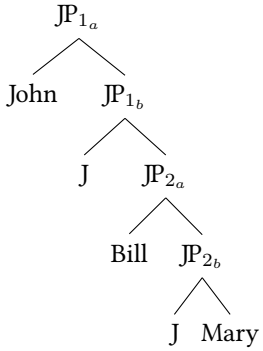


(12) $f(\{\text{John, Bill}\})$

Under this approach the multiplicity of Q-particles which is possible in English (e.g. *John **(or)** Bill **or** Mary*) and obligatory in Sinhala (e.g. *Gunapala* ***(də)** *Chitra* ***(də)** *Ranjit* ***(də)** 'Gunapala or Chitra or Ranjit') remains inexplicable. That is, the intuition is that J creates Hamblin-type sets of alternatives to which *a single* Q should apply. Before discussing the explanation of the "extra" Q-particles, let us first consider in detail what the rule in (9) would predict in the case of more than two disjuncts, as in (13).

(13) a. ⟦John J_1 Bill J_2 Mary⟧^{g}

b.



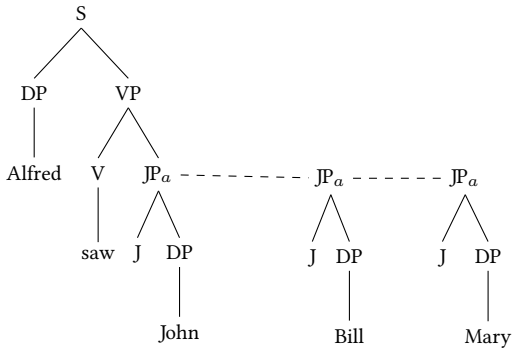
c. {John,{Bill,Mary}}

Using the rule in (9) incorrects predicts the denotation of (13a) to be (13c), since the rule requires recursive embedding of sets.[4] This results in the undesired outcome that a choice-function attempting to apply to (13c) will return either John or the set {Bill,Mary}. We have the additional complication both for (13) and even for the binary (10) that adjunct of *or* to the JP predicts that *or* should appear only once in the syntax, either preceding or following all of the disjuncts, e.g. **or John Bill Mary* or **John Bill Mary or.* An alternative translation of J is therefore required.[5]

[4]E.g. rule (9) predicts that ⟦John J Bill J Mary J Fred⟧^{g} = {John, {Bill, {Mary, Fred}}} and that ⟦John J Bill J Mary J Fred J Kim⟧^{g} = {John, {Bill, {Mary, {Fred, Kim}}}}, etc.

[5]We might adopt the idea that disjunction (and 'junction' more generally) is not binary-branching, but rather 'flat' or 'three dimensional'. Adopting the latter option, we might then suggest that junction requires a third dimension, with the 'juncts' being integrated into the tree in parallel, as shown in (i) for the sentence "Alfred saw John or Bill or Mary."

(i)



Now, if we assume that the Q-particle *or* adjoins to the JP, we might explain the multiple appearance of *or* in the syntax of disjunctions like *John or Bill or Mary* by adopting the position that in the case of adjunction to 'three dimensional' objects like junctions the adjoined element appears, in surface syntax, on each of the members of the disjunction. This approach would require some addition finnessing, at least in languages like English where 'or' cannot appear following the last member of the disjunction, and further has the possibility of appearing only between the final and penultimate disjunct. Presumably at least the second issue could be dealt with in terms of elision.

The three dimensional approach allows for the correct semantic representation of disjunctions involving more than two members. We can reformulate the rule for intepreting J(P) as follows:

(ii) 'Junction rule' (revised):

a. ⟦J⟧^{g}(⟦XP⟧^{g}) = {⟦XP⟧^{g}}

b. ⟦JP_1 - - - JP_2⟧^{g} = ⟦JP_1⟧^{g} ∪ ⟦JP_2⟧^{g}

This allows the JP in (i) to be represented semantically as {John,Bill,Mary}, as desired, as it does not encounter the same recursion problem as (9).

However, adopting either a flat or three dimensional approach to junction is problematic, since there is binding-based evidence that disjunctions and conjunctions have binary-branching internal structure, just like any other XP. Consider the asymmetry between (iiia) and (iiib).

(iii) a. [Every man]$_i$ or his$_{i,j}$ brother was there.

b. His$_i$ brother or [every man]$_{*i,j}$ was there.

The syntax of (dis/con)junction would appear then to be necessarily binary-branching and therefore a different approach is required.

Given that in Sinhala each disjunct is followed by an instance of *də* (or *hari* in declaratives), the choice-functional analysis of Q-particles suggests therefore that disjunctions might thus involve multiple instances of "set-recursion", e.g. {John, {Bill, {Mary}}}, which are "undone" by the application of multiple choice functions (represented by multiple occurrences of Q-particles like *də*). Let us continue using English as our object language for the nonce; I will return to analysis of Sinhala disjunctions anon.

In standard modern colloquial English, generally *or* must precede the ultimate disjunct, and may optionally precede each of other disjuncts, with the exception of the first disjunct (in which position we sometimes find *either*, but see above Section 10.1). In earlier English, and even in modern poetical English, the first disjunct may also be preceded by *or*, as shown by the examples below.[6]

(14) a. **Or** he shal singe *si dedero*, **or** al geineth him noht.
"He shall sing *si dedero*, or all profits him naught." [ca. 1330; T. Wright, *Polit. Songs Eng.* (1839) 324]

b. Loth to leaue vnsought **Or** that, **or** any place.
"Loath to leave unsought either that or any place." [a1616; Shakespeare, *Comedy of Errors* (1623) i. i. 136]

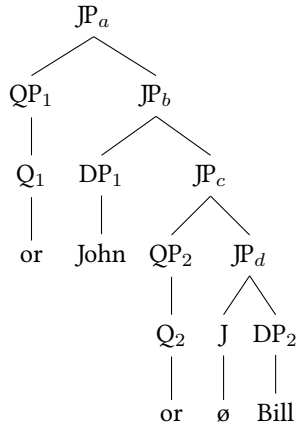
c. Without **or** wave **or** wind. [1800; S. T. Coleridge, *Anc. Mariner* vi, in Wordsworth & S. T. Coleridge, *Lyrical Ballads* (ed. 2) I. 185]

d. Learn that to love is the one way to know **Or** God **or** man. [1867; J. Ingelow, *Story of Doom* vii. 266]

e. His eyes are all glazed, **Or** far **or** near he can see nothing straight. [1957; D. L. Sayers, translation of the *Song of Roland* 128]

In view of these data, let us adopt the following working hypothesis: at some level of representation the structure of English disjunction is similar to that of Sinhala, in that each disjunct is associated with *or*—although in English they precede rather than follow the disjunct. In modern colloquial English the *or* preceding the initial disjunct is obligatorily elided (or else appears in the form *either*[7]), and all other instances of *or*, save the one preceding the final disjunct, are optionally elided—but this elision of course takes place on the PF-side, while on the LF-side each disjunct is still preceded by *or*. Adopting this analysis, and the JP hypothesis discussed above in Section 10.1, suggests the following (underlying) structure for the fragment *(either) John or Bill*:

(15)



Each Q-particle adjoins to the minimal structure containing (i) the head of JP and (ii) the disjunct with which the particle is associated. For QP$_2$ this involves adjoining to the segment of the JP containing [J DP$_2$]; for QP$_1$ this involves adjoining to the JP at the maximal level (JP$_a$).

[6] Cited from Murray et al. (2011).

[7] Although it may be that *either* has additional semantic and/or syntactic properties, as compared with *or*.

Adopting the syntactic structure of (15), we can revise the definition for the "junction" J accordingly.

(16) 'Junction rule' (final version):
$[\![\text{J}]\!]^{g} = \lambda X_{\langle \tau,t\rangle}\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle \tau,t\rangle}.\{Y\} \cup \{Z(\{\lambda P_{\langle \tau,t\rangle}.P\}(X))\}$

Here J takes three arguments: two XPs of the same semantic type (e.g. two DPs) and a choice-function variable (i.e. a Q-particle). The Hamblin-type identity function $\{\lambda P_{\langle \tau,t\rangle}.P\}$ guarantees, by the rules of flexible function application, that the lower disjunct is a Hamblin-type element.[8] The basic function of rule (16) is to perform a union operation over the set containing the higher disjunct and the set containing the result of the choice-function variable applied to the lower disjunct.

The semantic rule in (16) is much more satisfactory than our previous formulations. Up to this point I have been treating the multiple appearance of the Q-particle as a sort of quirk which requires an explanation. Adopting the syntactic structure of (15) and the semantic definition in (16), it becomes apparent that multiplicity of Q-particles is not a glitch, but rather the key to the understanding the syntax and semantics of disjunction. Each disjunct involves one level of recursive set-formation ("embedding"); and each disjunct is associated with a Q-particle which—since it denotes a variable over choice functions—successively transforms one layer of Hamblin set-formation back into ordinary Montagovian semantics. This is demonstrated in (17), which shows the derivation of (15).

(17) $[\![(15)]\!]^{g}$ =

a. $[\![Q_1]\!]^{g}([\![\text{JP}_b]\!]^{g})$ =
b. $[\![Q_1]\!]^{g}(\lambda X_{\langle \tau,t\rangle}\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle \tau,t\rangle}.$
$\{Y\} \cup \{Z(\{\lambda P_{\langle \tau,t\rangle}.P\}(X))\}([\![DP_1]\!]^{g})([\![Q_2]\!]^{g})([\![DP_1]\!]^{g}))$ =
c. $[\![or_1]\!]^{g}(\lambda X_{\langle \tau,t\rangle}\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle \tau,t\rangle}.$
$\{Y\} \cup \{Z(\{\lambda P_{\langle \tau,t\rangle}.P\}(X))\}([\![John]\!]^{g})([\![or_2]\!]^{g})([\![Bill]\!]^{g}))$ =
d. $f_1(\lambda X_{\langle \tau,t\rangle}\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle \tau,t\rangle}.$
$\{Y\} \cup \{Z(\{\lambda P_{\langle \tau,t\rangle}.P\}(X))\}(John)(f_2)(Bill))$ =
e. $f_1(\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle \tau,t\rangle}.$
$\{Y\} \cup \{Z(\{\lambda P_{\langle \tau,t\rangle}.P\}(Bill))\}(John)(f_2))$ =
f. $f_1(\lambda Y_{\langle \tau,t\rangle}.$
$\{Y\} \cup \{f_2(\{\lambda P_{\langle \tau,t\rangle}.P\}(Bill))\}(John))$ =
g. $f_1(\{John\} \cup \{f_2(\{\lambda P_{\langle \tau,t\rangle}.P\}(Bill))\})$ =
h. $f_1(\{John\} \cup \{f_2(\{c \mid \exists x \in \{\lambda P_{\langle \tau,t\rangle}.P\}[x(Bill)]\})\})$ =
i. $f_1(\{John\} \cup \{f_2(\{\lambda P_{\langle \tau,t\rangle}.P(Bill)\})\})$ =
j. $f_1(\{John\} \cup \{f_2(\{Bill\})\})$ =
k. $f_1(\{John, f_2(\{Bill\})\})$

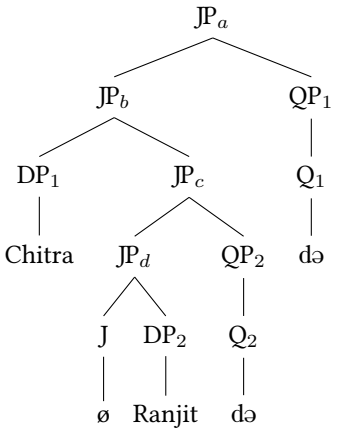
The choice function f_2 applies to the singleton set {Bill} and returns Bill; the choice function f_1 applies to the set containing {John, Bill} and returns one of these members. Thus the structure of disjunction and semantics of J proposed above does indeed provide a logical semantic treatment of disjunction—and furthermore allows us to explain the use of Q-particles in disjunctive structures under a unified analysis of Q-particles as denoting variables over choice functions.

Returning to Sinhala—which has an advantage over English, in terms of ease of analysis, in that the Q-particles overtly accompany each disjunct—note again that Sinhala differs from English in the placement of the Q-particle

[8] Essentially, the identity function transforms the lowest of a sequence of disjuncts into a Hamblin-type element, while leaving all subsequent disjuncts unaltered.

with respect to the disjunct, namely Q-particles follow disjuncts in Sinhala, rather than preceding them as in English. This is more problematic for our analysis than it might seem at first blush. Using the same structure for disjunction proposed above for English in (15), but positing that QPs right-adjoin to the segments of JP, results in the incorrect prediction that—rather than following each disjunct—the Q-particles should "stack up" at the end of the disjunct. Consider the Sinhala fragment *chitra də ranjit də* "(either) Chitra or Ranjit". Using the structure of (15), but with right-adjoined QPs results in:

(18) *



= * Chitra Ranjit də də

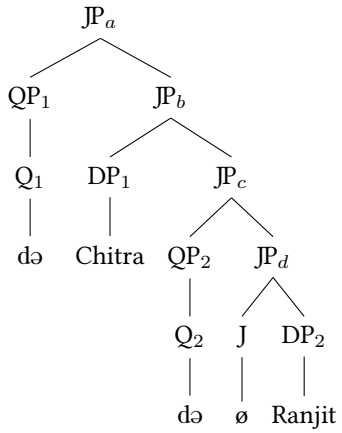
So rather than right-adjoining QPs, it seems preferable to adopt a syntactic analysis of Sinhala disjunction which is fully identical to that of English, but which involves later "PF-level" refashioning resulting in the Q-particles being aligned to the right of their disjuncts. I assume that this refashioning takes place according to the principles of Lowering as described in Embick & Noyer (2001) within the Distributed Morphology framework (on which see also Halle & Marantz 1993, Embick & Noyer 2007, amongst others). The Lowering operation involves a one head being appended to a structurally lower head, as schematised below:

(19) Lowering:
$[_{XP}\ X^0 \ldots [_{YP} \ldots Y^0 \ldots]] \rightarrow [_{XP} \ldots [_{YP} \ldots [_{Y^0}\ Y^0{+}X^0] \ldots]]$ (from Embick & Noyer 2001: 561)

We can justify this Lowering operation along the following lines. The particles *də* and *hari* are enclitics (as is evident from their positioning in other constructions). Syntax dictates that they occupy a particular position; however PF-rules can operate on the output of syntax, and transform this output according to morphophonological principles. In the case of Sinhala Q-particles, PF lowers the Q-particle to the head of the nearest disjunct (in our example, the head of the nearest DP).

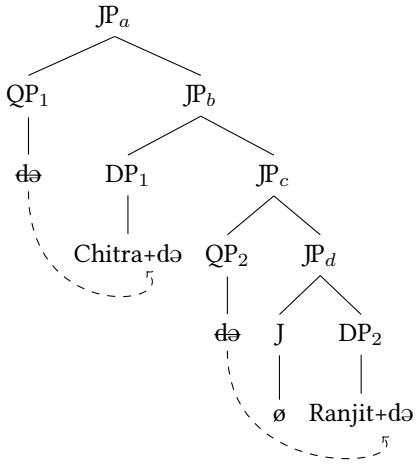
Thus the base-generated structure for *chitra də ranjit də* "(either) Chitra or Ranjit" would be like that of English:

(20) Pre-Spellout/"S-Structure":



At "PF" the Q-particles undergo morphophonological Lowering:

(21) "PF-Structure":

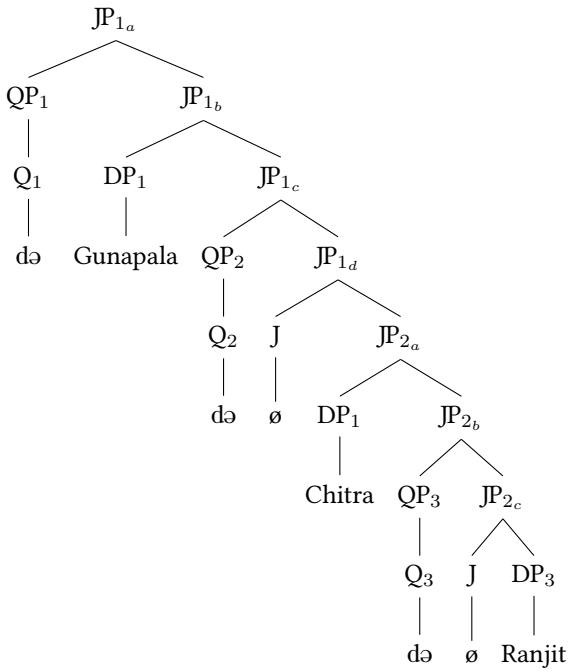


Consider a more complicated example like (22), involving three disjuncts—just to show that the analysis generalises correctly to cases of more than two disjuncts.

(22) gunəpaalə də chitra də ranjit də gamət̪ə giye?
Gunapala *də* Chitra *də* Ranjit *də* village.DAT went-E
"Was it Gunapala or Chitra or Ranjit who went to the village?"

The (pre-PF) syntactic structure of the disjunction of (22) is provided in (23) and the semantic derivation is provided in (24).

(23)



(24) a. $[\![\text{JP}_{2_a}]\!]^g$ =

(i) $\lambda X_{\langle\tau,t\rangle}\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle\tau,t\rangle}.$
$\{Y\} \cup \{Z(\{\lambda P_{\langle\tau,t\rangle}.P\}(X))\}([\![\text{DP}_3]\!]^g)([\![\text{Q}_3]\!]^g)([\![\text{DP}_2]\!]^g)$ =

(ii) $\lambda X_{\langle\tau,t\rangle}\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle\tau,t\rangle}.$
$\{Y\} \cup \{Z(\{\lambda P_{\langle\tau,t\rangle}.P\}(X))\}(\text{Ranjit})(f_3)(\text{Chitra})$ =

(iii) $\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle\tau,t\rangle}.$
$\{Y\} \cup \{Z(\{\lambda P_{\langle\tau,t\rangle}.P\}(\text{Chitra}))\}(\text{Ranjit})(f_3)$ =

(iv) $\lambda Y_{\langle\tau,t\rangle}.$
$\{Y\} \cup \{f_3(\{\lambda P_{\langle\tau,t\rangle}.P\}(\text{Chitra}))\}(\text{Ranjit})$ =

(v) $\{\text{Ranjit}\} \cup \{f_3(\{\lambda P_{\langle\tau,t\rangle}.P\}(\text{Chitra}))\}$ =

(vi) $\{\text{Ranjit}\} \cup \{f_3(\{c \mid \exists x \in \{\lambda P_{\langle\tau,t\rangle}.P\}[x(\text{Chitra})]\})\}$ =

(vii) $\{\text{Ranjit}\} \cup \{f_3(\{\lambda P_{\langle\tau,t\rangle}.P(\text{Chitra})\})\}$ =

(viii) $\{\text{Ranjit}\} \cup \{f_3(\{\text{Chitra}\})\}$ =

(ix) $\{\text{Ranjit}, f_3(\{\text{Chitra}\})\}$

b. $[\![\text{JP}_{1_b}]\!]^g$ =

(i) $\lambda X_{\langle\tau,t\rangle}\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle\tau,t\rangle}$.
$\{Y\} \cup \{Z(\{\lambda P_{\langle\tau,t\rangle}.P\}(X))\}([\![\text{JP}_2]\!]^g)([\![\text{Q}_2]\!]^g)([\![\text{DP}_1]\!]^g)$ =

(ii) $\lambda X_{\langle\tau,t\rangle}\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle\tau,t\rangle}$.
$\{Y\} \cup \{Z(\{\lambda P_{\langle\tau,t\rangle}.P\}(X))\}(\{\text{Ranjit}, f_3(\{\text{Chitra}\})\})(f_2)(\text{Gunapala})$ =

(iii) $\lambda Z_{\langle cf,t\rangle}\lambda Y_{\langle\tau,t\rangle}$.
$\{Y\} \cup \{Z(\{\lambda P_{\langle\tau,t\rangle}.P\}(\{\text{Ranjit}, f_3(\{\text{Chitra}\})\}))\}(f_2)(\text{Gunapala})$ =

(iv) $\lambda Y_{\langle\tau,t\rangle}$.
$\{Y\} \cup \{f_2(\{\lambda P_{\langle\tau,t\rangle}.P\}(\{\text{Ranjit}, f_3(\{\text{Chitra}\})\}))\}(\text{Gunapala})$ =

(v) $\{\text{Gunapala}\} \cup \{f_2(\{\lambda P_{\langle\tau,t\rangle}.P\}(\{\text{Ranjit}, f_3(\{\text{Chitra}\})\}))\}$ =

(vi) $\{\text{Gunapala}\} \cup \{f_2(\{c \mid \exists x \in \{\lambda P_{\langle\tau,t\rangle}.P\}, \exists y \in \{\text{Ranjit}, f_3(\{\text{Chitra}\})\}\ [c=x(y)]\})\}$ =

(vii) $\{\text{Gunapala}\} \cup \{f_2(\{\lambda P_{\langle\tau,t\rangle}.P(\text{Ranjit}), \lambda P_{\langle\tau,t\rangle}.P(f_3(\{\text{Chitra}\}))\})\}$ =

(viii) $\{\text{Gunapala}\} \cup \{f_2(\{\text{Ranjit}, f_3(\{\text{Chitra}\})\})\}$ =

(ix) $\{\text{Gunapala}, f_2(\{\text{Ranjit}, f_3(\{\text{Chitra}\})\})\}$

c. $[\![\text{JP}_{1_a}]\!]^g = f_1(\{\text{Gunapala}, f_2(\{\text{Ranjit}, f_3(\{\text{Chitra}\})\})\})$

Adopting the translation of J as in (16) therefore allows for the correct representation of disjunctive structures as shown by the derivation in (24). Thus, a unified semantic treatment of Q-particles like Sinhala *də*—appearing not only in interrogative and indefinite constructions, but also in disjunctions—is shown to be possible.

10.3 A brief note on conjunction

As discussed above in Chapter 2, the Japanese Q-particle *ka* appears (amongst other contexts) in both disjunctions (25) and in the formation of (existential) *wh*-indefinites (26) (cf. Hagstrom 1998: 17–8).

(25) a. John-**ka** Bill-(**ka**-)ga hon-o katta.
John-***ka*** Bill(***ka***-)NOM book-ACC bought.
"John or Bill bought books."

b. John-ga hon-o katta-**ka** Bill-ga hon-o katta-(***ka*** desu).
John-NOM book-ACC bought-***ka*** Bill-NOM book-ACC bought(-***ka*** is).
"John bought books or Bill bought books." (Kuroda 1965: 85)

(26) a. dare-**ka**-ga hon-o katta.
who-***ka***-NOM book-ACC bought.
"Someone bought books."

b. John-ga nani-**ka**-o katta.
John-NOM what-***ka***-ACC bought.
"John bought something." (Kuroda 1965: 97)

In an intriguingly parallel formation, the Japanese particle *mo* appears both as a "conjunction marker" (27) and, added to *wh*-words, in the construction of universal quantifiers (28) (cf. Hagstrom 1998: 18–9).

(27) a. John-ga hon-**mo** zassi-**mo** katta.
John-NOM book-***mo*** magazine-***mo*** bought.
"John bought both books and magazines."

b. John-ga hon-o kai-***mo***-si, zassi-o kai-***mo*** sita.
John-NOM book-ACC buy-***mo***-do, magazine-ACC buy-***mo*** did.
"John bought books and John bought magazines." (Kuroda 1965: 77–8)

(28) a. dare-**mo**-ga kita.
who-***mo***-NOM came.
"Everyone came." (Kawashima 1994: 147)

b. dare-**mo** hon-o kaw-anakat-ta.
who-***mo*** book-ACC bought-NEG.
"No-one bought books." (Kuroda 1965: 94)

Similar data are found in Sinhala. As discussed *passim*, the Sinhala Q-particles *də* and *hari* both function to form disjunctions as well as indefinite pronouns (when added to a *wh*-word). Sinhala also possesses a clitic *-t* which appears—like Japanese *mo*—both as a "conjunction marker" and, when added to a *wh*-word, forms universal quantifiers (29).

(29) a. chitra kauru-**t** ekkə kataa kəlaa.
Chitra who-***t*** with talk did
"Chitra talked with everyone."

b. ranjit kaurun-ṭə-**t** gæhuwa.
Ranjit who-DAT-***t*** hit.
"Ranjit hit everyone." (Kishimoto 1992: 55)

The clitic *-t* is not, however, the only conjunction element in Sinhala; in fact, it is only used to link clauses (Fairbanks et al. 1968: 228), while an element at least homophonous with the focus marker (*-y(i)*), is used to link non-phrasal conjuncts, as in (30) below.[9] Like Japanese *mo*, Sinhala *-t* means "also" when attached to an NP (Fairbanks et al. 1968: 197).

(30) a. putek(u)-**y** duwek(u)-**y**
son-***y*** daughter-***y***
"a son and a daughter"

b. bas-ekək(u)-**y** ṭæksiy-ək(u)-**y**
bas-INDEF-***y*** taxi-INDEF-***y***
"a bus and a taxi" (Fairbanks et al. 1968: 105)

Sinhala thus exhibits the same basic pattern as Japanese, suggesting a connection between disjunction and existential quantification on the one hand, and between conjunction and universal quantification on the other, see (31), a connection which has not gone unnoticed in previous literature, cf. Reichenbach (1947); Rohrer (1973).

(31) a. $(\exists x)\ f(x) \equiv f(x_1) \vee f(x_2) \ldots \vee f(x_n)$
b. $(\forall x)\ f(x) \equiv f(x_1)\ \&\ f(x_2) \ldots\ \&\ f(x_n)$

The above data suggest a treatment of "conjunctive particles" like Sinhala *-t* and Japanese *mo* which parallels the treatment of Q-particles like Sinhala *də* and Japanese *ka*, i.e. as denoting variables over choice functions. However,

[9] Literary Sinhala also has the element *saha* which can function to mean "also" when appearing after an NP and also may form conjunctions.

while Sinhala *də* and Japanese *ka* have been analysed as denoting simple variables over choice functions which obtain their quantification from outside (i.e. from the denotation of a COMP head or through existential closure), *-t* and *mo* must be analysed as bearing inherent universal quantification, i.e. as in (32).

(32) $[\![$"and" XP$]\!]^g$ (e.g. Japanese *mo*) = $\forall f.\in D_{cf}.f([\![XP]\!]^g)$

Note further that such an approach accords with den Dikken's proposal that the junction phrase JP is involved in both disjunction and conjunction. Under this analysis the role of the junction head in both cases is simply to form a (Hamblin-type) set, to which may be applied either an existentially-quantified choice function (resulting in disjunction) or a universally-quantified choice function (resulting in conjunction).

However, as pointed out by Haspelmath (1997: 157–158), in many languages the combination of *wh*-words with "conjunction markers" does not in fact form universally-quantified pronouns. Such is the case in Malayalam, where the addition of *-um* "and" to a *wh*-word like *aar* "who" results in *aarum* "anybody" (Jayaseelan 2001). Haspelmath (1997: 157–158) provides further examples of NPI and/or free-choice pronouns formed from the combination of a *wh*-word and an element translatable as "and, even, also" in Serbo-Croatian, Indonesian, Tagalog, Kannada, and Ancash Quechua. Further, in some cases the addition of an element meaning "and, also, even" to a *wh*-word results in an existential indefinite pronoun, as in Hittite (*kuiš* "who", *-ki* "and, also", *kuiš-ki* "someone"). Since Japanese is the only language included in Haspelmath's study where WH+"and" results in a universally-quantified element, the Japanese pattern appears to him to be the exception and not the rule, and he therefore suggests that the connection in (31b) may be spurious.

However, as noted above, Sinhala also possesses WH+"and" universal pronouns, and similar data are found in other languages. Gothic, while it does not form *wh*-indefinites like Sinhala or Japanese, possesses an element *-(u)h*—a clitic particle meaning "and"—which, when suffixed to *wh*-words forms universal quantifiers. Thus Gothic *ƕas* "who, what", *ƕarjis* "who, which", *ƕaþar* "which of two" : *ƕaz-**uh*** "every, each", *ƕarjiz-**uh*** "every, each", *ƕaþar-**uh*** "each of two", also *áinƕarjiz-**uh*** "everyone" (Wright 1910: §275).[10]

Latin too displays a similar phenomenon: *quis* "who, which", *-que* "and, also", *quis-que* "every" (Haspelmath 1997: 156).[11] Thus the Japanese pattern of WH+"and"=universal-quantifier is not as exceptional as Haspelmath's study suggests.

But even the combination of *wh*-words with "disjunction markers", i.e. the WH+Q-particle indefinites explored extensively in this study, does not always result in existential indefinites, crosslinguistically-speaking. Haspelmath (1997: 166) provides examples of NPI and/or free-choice indefinites formed from the combination of a *wh*-word and an element translatable as "or" in Korean, Russian, Hungarian, Basque, Latvian, Romanian, Ossetic, Hausa, and West Greenlandic.

Though further investigation of the questions raised by this crosslinguistic variability in the meaning of WH+"or" and WH+"and" falls outside of the scope of this study, I suggest that the Japanese/Sinhala pattern is the one which is synchronically-motivated (i.e. as in (31)), and some of the other patterns observed by Haspelmath (1997) may be the result of historical change.[12]

[10]The more usual element used for "and" in Gothic is *jah*, which also can mean "also, even", while *-(u)h* only rarely possesses these latter two values (Klein & Condon 1993: 2–3).

[11]Sinhala *-t*, Latin *-que*, and Gothic *-(u)h* do ultimately derive (at least in part) from the same PIE source: $*k^we$ "and". Gothic *-(u)h* appears to stem from PIE $*u + *k^we$ (see Brugmann 1904: 62–63, Klein & Condon 1993: 2). Sinhala *-t* from Sanskrit *ca* (< PIE $*k^we$) > *ja* > *da* > *d* > *t*. However, the use of elements deriving from PIE $*k^we$ in these three languages appears to be an independent development in each case.

[12]In fact, in many of the languages the "conjunctive markers" which participate in the formation of various sorts of *wh*-based pronouns actually can mean not only "and", but also "even" or "also". Given that "and", "even", and "also" have different semantics/pragmatics, it may be possible to derive some of alternatives to universal quantifiers noted by Haspelmath (i.e. NPI or free-choice pronouns) as deriving from the senses of "even" or "also" rather than from "and" (which I suggest here denotes a universally quantified choice function variable).

In any event, the connection between disjunction and existential quantification on the one hand, and between conjunction and universal quantification on the other, as exhibited by the Japanese and Sinhala data examined above suggests that an analysis of conjunction involving universally-quantified choice functions is a promising approach to the semantics of conjunction, especially as it accords with the proposal of den Dikken (2006) that a covert junction phrase JP is involved in both disjunction and conjunction, allowing for a uniform treatment of semantic role of the junction head J as creating Hamblin-type sets.

10.4 Conclusion and implications for generative grammar

This chapter provides the details of an analysis for disjunction which makes possible a unified semantic analysis of Q-particles like Sinhala *hari*, Japanese *ka*, as denoting variables over choice functions. Such an analysis requires that disjunction involves some element which creates Hamblin-type sets to which choice functions may apply. I adopt den Dikken's (2006) proposal that disjunctions involve a covert junction element J, heading a projection JP, with the disjuncts themselves appearing in J's complement and specifier positions. The visible markers of disjunction like Sinhala *hari*, Japanese *ka*, English *either*, *or* are phrasal categories adjoining to the disjuncts. This syntactic analysis allows for the division of labour which the unified semantic treatment of Q-particles requires: namely, J is responsible for creating Hamblin-type sets, to which choice functions (the denotation of Q-particles) may apply. These choice function variables are bound by existential closure, thus resulting in the desired denotation for disjunction. As noted, Alonso-Ovalle (2006) provides a number of independent arguments for a Hamblin analysis of disjunction, so the analysis of disjunction advocated here has a number of advantages over previous treatments.

If this analysis is correct, then it suggests that—while Q-particles are found in interrogatives, indefinites, and relative clauses only in certain languages—in disjunctions, Q-particles may be ubiquitous. That is, elements generally classified as disjunction operators like English *or* are in fact Q-particles.

Independent support for such a notion can be found in earlier stages of English, where *or* may appear not only in the formation of disjunctions, as in modern English, but also seems to sometimes function as a marker of yes/no-questions, as in examples (33) and (34).

(33) **Or** not in wrathefulnesse of hym is lettid þe sunne, & oo dai maad as two?
"Was not the sun stopped in his anger, and one day made as two?" [Ecclus. 46.5 (Wycliffe Bible, ca. 1382)]

(34) He asked the lordes. . . **or** they wolde therfore warre.
"He asked the lords if they would therefore go to war." [Virgilius sig. aiiij^{v}, ca. 1518]

Jayaseelan (2008) presents additional evidence for the existence of a connection between "disjunction markers" (like English *or*) and "yes/no question markers". Noting first that in English *if* functions as a yes/no question marker in embedded questions (e.g. "he asked the lords if they would go to war"), Jayaseelan points out that English *if* is cognate with Dutch *of* (Murray et al. 2011), the latter being the "disjunction marker" of Dutch (e.g. *of Jan of Marie* "John or Mary"). Thus what is a yes/no question marker in one language has a cognate in the other which serves as a disjunction marker. Further Jayaseelan remarks that in "colloquial substandard Dutch" *of* also appears as a question marker introducing not only embedded yes/no-questions (see example (35)) but also *wh*-questions (see examples (36a), (36b)).[13]

[13]The following data are from Hoekstra (1993) via Bayer (2004), who notes that similar data are found in Frisian, West Flemish, and certain varieties of Swiss German.

(35) Ik vraag me af [**of** [dat [Ajax de volgende ronde halt]]]
I ask me PART [**if** [that [Ajax the next round reaches]]]
"I wonder whether Ajax will make it to the next round."

(36) a. Hij weet [**hoe** [**of** [je dat moet doen]]]
he knows [**how** [**if** [you this must do.INF]]]
"He knows how you must do this."

b. Ze weet [**wie** [**of** [dat [hij had willen opbellen]]]]
she knows [**who** [**if** [that [he had wanted call.INF]]]]
"She knows who he wanted to call."

These data are particularly interesting since they show the presence of an overt Q-particle in Germanic which appears uniformly in all types of embedded questions, including *wh*-questions, that also appears in the formation of disjunctions (e.g. *of Jan of Marie* "John or Mary").

This strengthens the argument that so-called "disjunction markers" are universally Q-particles (elements denoting variables over choice functions), suggesting that the basic analysis of disjunction proposed here is not confined to languages like Sinhala and Japanese, but is appropriate for the analysis of disjunction more generally, including in languages like modern English which otherwise do not appear to possess overt Q-particles.

In section 10.3 above, I suggested that conjunction can be handled in a similar fashion, except that the choice functions involved are universally- rather than existentially-bound. A choice-functional analysis thus offers a promising new approach to the syntax and semantics of both disjunction and conjunction, though further research is warranted, particularly in the case of conjunction.

The following chapter turns to the examination of the diachronic development of Q-particle and focus constructions in the history of Sinhala, and discusses the ramifications of these philological data for synchronic formal analysis.

Chapter 11

Historical developments in Sinhala Q-particle and focus constructions

> '. . . And only when you are at the proper distance will you see that it is Brunellus (or, rather that horse and not another, however you decide to call it). And that will be full knowledge, the learning of the singular. So an hour ago I could expect all horses, but not because of the vastness of my intellect, but because of the paucity of my deduction. . . '
>
> —William of Baskerville to his pupil Adso of Melk,
> in Umberto Eco's *The name of the rose*

In this chapter I investigate two sets of changes in the history of Sinhala–namely the development of the focus construction and the evolution of Q-particles—and show that an understanding of these diachronic developments leads to new insights into the proper formal synchronic analysis of these structures in modern colloquial Sinhala.

The relevant stages of Sinhala I examine are: (1) Modern Colloquial Sinhala [MCS], the usual form of the present-day spoken language; (2) Modern Literary Sinhala [LS], the formal written variety of the modern language, which preserves earlier features of the language; (3) Classical Sinhala [CS], ca. 12th–15th centuries A.D.; (4) Old Sinhala [OS], represented by the Sigiri graffiti texts of the 8th–10th centuries A.D. Recall that modern literary Sinhala preserves a number of archaisms such as overt subject-verb agreement morphology, which makes it reasonable to treat it as representing a stage historically earlier to that of the modern colloquial language.

The development of the Sinhala focus construction is examined in Section 11.1 and the development of Q-particle constructions in Section 11.2. Both sections involve a certain amount of etymological discussion; I deem this philological investigation relevant to the overall study because it is necessary in order to determine the actual sources of (i) the *-e* verbal form used in focussing constructions and (ii) the Q-particles *də* and *hari*. An understanding of the sources of these elements helps us to understand their synchronic properties, and—in the case of the etymological history of *də* and *hari*—provides us with a detailed picture of possible sources of Q-particles.

11.1 The development of the Sinhala focus construction

In Chapters 4 and 5, I presented formal syntactic and semantic analyses of modern colloquial Sinhala focus constructions. In this section, I investigate the historical antecedents of the modern focus construction in Sinhala, and show that its earliest extant antecedents in Old Sinhala have a radically different structure. The initial origin of this construction appears to be an internal development, not due to contact with Tamil—though later developments appear to reflect Dravidian influence. The changes which take place in Classical Sinhala bring the structure more into line with the Dravidian construction, but the subsequent development in modern Sinhala again appears to represent an internal development, resulting in divergence from the Tamil focus construction.

11.1.1 "Focus" constructions in Old Sinhala (8th–10th c.)

The present participle of Sinhala comes from Old Indo-Aryan primary derivatives in *-ana*,[1] e.g. Skt. *kr̥-* 'to do' > Skt. *káraṇa* 'deed', Skt. *dā* 'to give' > Skt. *dā́na* 'giving' etc. (see Paranavitana 1956). Sinhala past participles derive from Old Indo-Aryan *ta*-participles in *-ta, -na* or *-ita*.[2] Thus Old Sinhala *giya* < Old Indo-Aryan *gataḥ*.

The Old/Classical Sinhala present participle has two basic forms, one of which takes the form *-nu* or *-nā*, and is used as an abstract action-noun, the Sinhalese grammarians' *hav-kiriya* (Skt. *bhāva kriyā*),[3] e.g. *balanu* 'seeing (noun)', *basnā* 'sinking (noun)'. Old/Classical Sinhala present participles are declined as inanimates, and thus retain their stem forms in nominative and accusative cases, and can be inflected in the instrumental, dative and locative cases. See example (1).

(1) ***balanu*** *tabay* 'without **looking**' (*S.G.* 375) [OS]

The other form of the present participle ends in *-na* or *-nā*.[4] Present participles ending in *-na*/*-nā* can occur as adjectives, in which event they do not inflect; otherwise, they can be declined as animate nouns (as either masculines or feminines). The past participle, typically ending either in *-u* or *-i*, can also appear as an uninflected adjective, or be inflected as a masculine or feminine noun. See Table 11.1.

Class	Endings	Declension	Function	Examples
Present	*-nu, -nā*	inanimate	action-noun	*balanu, basnā*
Participle	*-na, -nā*	non-declined	adjective	*balana, basnā*
	-nu, -nā	animate (masc. & fem.)	participant-noun	*balannē (balanuyē), basnē*
	-nu, -nā	animate (masc./neu.)	impersonal	*balannē (balanuyē), basnē*
Past	*-u, -i*	non-declined	adjective	*bælu, bæsi*
Participle	*-u, -i*	animate (masc. & fem.)	participant-noun	*bæluve, bæssē (bæsiyē)*
	-u, -i	animate (masc./neu.)	impersonal	*bæluve, bæssē (bæsiyē)*

Table 11.1: Participles in early Sinhala

It is the adjectival/nominal forms of the participles which will concern us here, particularly the 'impersonal' use of the adjective/nominal participle, which ultimately forms the basis for the later Sinhala E-verbal forms.

However, it is important to note that, in both Old and Classical Sinhala, these adjectival/nominal forms of the participle are used not only with the 'impersonal' function (examined in detail below), but also as modifying adjectives or nominals meaning "the one who Xs". See (2) for an example of this form used as a adjectival participle modifying a noun; and (3)–(5) for examples of this form used as a nominal (examples (2) and (3) show present-tense forms of the participle; (4) and (5) show past-tense forms of the participle).

(2) *gala* ***yana*** *minisun* 'people [*minisun*] **going** to the rock' (*S.G.* 495) [OS]

(3) *maga-****yanno*** *yati* 'path-**goers** go' (*S.G.* 51) [OS]

(4) ***hindi*** *se* '**having-stayed** manner', i.e. 'as if having remained' (*S.G.* 189) [OS]

(5) *beyada* ***giyanaṭa*** '**to those who have gone** to the mountainside' (*S.G.* 111) [OS]

[1] See Whitney (1879/1889: §1150).

[2] This participle is sometimes called past passive participles—see Whitney 1879/1889: §952–8, but more properly treated as ergative than passive—see Hock 1986.

[3] See De Alwis (1852). The distinction between *-nu* and *-nā* forms later largely disappears in favour of the former. The latter appear in Old and Classical texts only with verbs of Geiger's conjugation II (Geiger 1938: §140, 142), i.e. verbs which later take an *-i* stem vowel. See further Reynolds (1964: 130n7,9).

[4] Again the form chosen depends on the root's conjugation membership; the former occurs in Geiger's conjugation I, the latter in conjugation II.

We also find examples where such nominal uses are patient-oriented, rather than agent-oriented, as in example (6) below.

(6) pere tā no **laddiya** dæyek no **lad** veye
formerly you.GEN not **obtain**.PAST.PTCP.NOM.FEM thing.INDEF.NOM not **obtain**.PAST.PTCP.NEU.NOM be.PRES.3SG
ranaṭ
gold.DAT
'**She who had not been obtained** by you earlier is a thing not **obtained** for gold.' (*S.G.* 609) [OS]

These nominal uses of the participle can, like other nouns, appear in the full range of possible case forms, e.g. in the dative case as above in example (5) or in the accusative as in example (7) below.

(7) pansiya agnan saga **yannā** keḷe pasu
five-hundred.NOM woman.GEN.PL heaven **go**.PRES.PTCP.ACC make.PAST.PTCP.NOM behind/afterwards
'The five hundred women[5] delayed **him who was going** to heaven.' (*S.G.* 44) [OS]

In contrast, the ancestor of the modern Sinhala 'focus' E-forms of the verb descends from a masculine/neuter nominative singular case-form (whose suffix takes a variety of forms depending on the conjugation class, but prototypically appears as *-e*), which is used 'impersonally', without any syntactic subject, in the sense "there is Xing". The logical subject, if present, is expressed by the use of a noun or pronoun in genitive case, as in example (8).

(8) **vayane** ma^{F} sihigiriye taṭa
play-music.PRES.PTCP.NOM I.GEN^{F} Sihigiri.LOC you.DAT
'It is by me that music is played, to you at Sigiri.'
(Lit., 'There is playing of music by me to you at Sihigiri.') (*S.G.* 526) [OS]

Example (8) resembles later Sinhala focus constructions further in that the element which appears to be focussed appears immediately following the participial verb (with *-e* suffix).

At this stage, this 'impersonal' use of the participle does not appear to be specialised as a focussing construction. It is a marked construction, its primary function seeming to be to either foreground or background some constituent. Thus in example (9), despite the use of two 'impersonal' participial constructions, there are no elements which appear to be focussed; rather, the subjects of both clauses appear to be backgrounded/de-emphasised.

(9) [[no bæṇæ]$^{(F)?!}$ **visi** mā gala]$^{(F)?!}$ **kī** ta
[[not speak.GER]$^{(F)?!}$ **remain**.PAST.PTCP.NOM I.GEN rock.LOC]$^{(F)?!}$ **say**.PAST.PTCP.NOM you.GEN
"'I remained on the rock without speaking"–thus you said.'
(Lit."'There was remaining of me on the rock without speaking." There was saying by you (of this).') (*S.G.* 508) [OS]

However, I have discovered one graffito, shown below in (10), in which the impersonal participial construction is obviously used with a focussed element, a fact made clear by the appearance of the emphasiser *nu*. Here the focussed constituent is dislocated immediately to the right of the verb, the typical position of focussed elements in later Sinhala.

[5]Many of the Sigiri graffiti concern and/or are addressed to the female figures painted on the Mirror Wall.

(10) ma sova niva æ meseyi duḍuḷa-sela-aḍadarihi **visi** ya
my.GEN sorrow extinguish.ABS she.GEN this-manner fortress-rock-edge.LOC **dwell**.PAST.PTCP.NOM PRES.3SG
yaha-asaraṭa^{F} **nu**
happy-companionship.DAT^{F} **indeed**
'It is indeed for the sake of happy companionship that she, having extinguished my sorrow, dwelt in this manner at the edge of the fortress rock.'
(Lit. 'Having extinguished my sorrow, there is her having dwelt in this manner at the edge of the fortress rock indeed for happy companionship.') (*S.G.* 32) [OS]

Unlike later Classical and modern literary Sinhala, displacement of focussed elements is not, however, obligatory, even where the "impersonal" participial is employed, see examples (11), (12).

(11) kavar dese āma [maṁg me]^{F} **mæ** pænini
which(ever) direction.LOC come.COND [path this]^{F} **EMPH** appear.PAST.PTCP.NOM
"In whatever direction (one) came, **this very path** appeared." (*S.G.* 35) [OS]

(12) [æsi-piyev eyun]^{F} **layu** nætte
[eyelid-movement their.GEN.PL]^{F} **EMPH** is-not.PRES.PTCP.NOM
"There is not **even** movement of their eyelids. (*S.G.* 17) [OS]

Further, unlike later Classical and modern literary Sinhala, the appearance of a copula or an agreement-clitic like *ya* is not obligatory in Old Sinhala; note the absence of either copula or *ya* in example (8) above. When a copula or *ya* does appear in impersonal participial constructions, it always appears immediately to the right of the verbal participle (see example (10) above for *ya* and example (13) below for the copula *ve*), not to the right of the focussed constituent—while the latter is the invariant position of *ya*/copula in both Classical and modern literary Sinhala, and of optionally-appearing focus-marking *yi, tamaa, tamay* in modern colloquial Sinhala.

(13) yannaṭ^{F} **ā** ve sihigiri
go.PRES.PTCP.DAT^{F} **come**.PAST.PTCP.NOM be.PRES.3SG Sihigiri
'It was to leave Sigiri that I came.'
(Lit. 'It was to go (away) that there was coming (of me) to Sihigiri.') (*S.G.* 168) [OS]

Furthermore, note that there is no fixed position in which focussed elements occur in Old Sinhala. In some cases they appear immediately to the right of the impersonal participle, as in examples (8), (10) above, sometimes they appear on the left edge of the sentence, as in example (13) above and example (14) below.

(14) [man buṇa kalak mata dækæ sihigiri]^{F} **ā**
[mind break.PAST.PTCP.ACC woman.ACC.INDEF summit.LOC see.ABS Sihigiri.GEN]^{F} **come**.PAST.PTCP.NOM
ve mā
be.PRES.3SG I.GEN
'It is because I saw a damsel whose pride had been shattered on the summit of Sigiri that I came back.'
(Lit. 'There was coming back of me, having seen on the summit of Sihigiri a damsel whose pride has been shattered.') (*S.G.* 510) [OS]

Some of the clearest examples of impersonal participles in focussing constructions occur in negative yes/no questions.[6] In these examples, the negation *no* occurs immediately following the focussed element, see examples (15)–(17).[7]

[6]See Han & Romero (2001); Romero & Han (2004) for a formal treatment of negative yes/no questions.
[7]The negative particle *no* appears to also signal that the clause is an interrogative.

(15) un [rahasaṭa]^{F} **no** beyada ature siṭi
she.GEN.PL [secrecy.DAT]^{F} **NEG** mountainside.LOC midst.LOC remain.PAST.PTCP.NOM
Lit. "Was there not remaining by them (fem.) in the midst of the mountainside for (the sake of) secrecy?" (apparently to be understood as "Is it not for the sake of secrecy that they (fem.) [*scil.* the women whose images are painted on the Mirror Wall] remained in the midst of the mountainside?") (*S.G.* 65) [OS]

(16) [gīyak]^{F} **no** kī
[verse.ACC.INDEF]^{F} **NEG** say.PAST.PART.NOM
"Is it not a verse said (by me)?" (*S.G.* 194) [OS]

(17) [sahaja-piyanaṭ]^{F} **no** meyhi saran andva tubu
[born-lovers.DAT]^{F} **NEG** here.LOC *apasaras*(=nymphs) paint.PAST.PART.NOM
"Is it not for the sake of born lovers that nymphs have been painted here?" (*S.G.* 175) [OS]

In any event, at this stage of Sinhala, there is no direct correlation between the presence of a focussed element and the use of the impersonal participial construction, as there is in later Classical Sinhala (or between the presence of a focussed element and the use of E-marked verbs as in modern Sinhala). Rather, focus seems to have been optionally realised through the use of emphatic particles like *nu*, *mæ*, *layu* etc. (see (10), (11), (12) above) and presumably through intonation as in modern Sinhala. The co-occurrence of these impersonal participial constructions with focussed elements is due to the fact that impersonal participial constructions are compatible with the presence of a focussed element since they are a marked construction which appears to function to either foreground (≈ focus) or background elements.

11.1.2 Focus Constructions in Classical Sinhala (12th–15th c.)

It is in Classical Sinhala that we find a focus construction which more closely resembles the Dravidian focus construction, and which thus is more likely to reflect convergence with Tamil. Note, however, that internal developments led to the appearance of the 'impersonal' participial construction of Old Sinhala (described above), which was itself already compatible with the presence of focussed elements. This compatibility with focus, along with the syntactic equivalence of the Sinhala nominal participle with the Tamil nominalised verb, contributed to make the impersonal construction an obvious potential candidate for 'calqueing' the Tamil focus construction.

Recall from Chapter 4 that in Tamil the nominalised verb is employed both in focus constructions (18), (19), (20), as well in clausal nominalisations, as in (21).

(18) naan poonatu yaaLppaaNattukku^{F}
I.NOM went.PAST.VN.NOM Jaffna.DAT^{F}
'It was to Jaffna that I went.' (cited from Gair 1986[1998]b: 156) [SLT]

(19) yāṉ oṟu-pp-atu numar-ai^{F}
I.NOM punish-NONPAST.VN.NOM relation-ACC^{F}
'It is the relations whom I punish.' (Old Tamil, *kali* 58.20; cited from Lehmann 1998: 97) [OT]

(20) a. nāṉ maturai.y-il piṟa-nt-ēṉ
I.NOM Madurai-LOC be_born-PAST-1SG
'I was born in Madurai.'

b. [$_{S}$ [$_{NP_1}$ nāṉ piṟa-nt-atu] [$_{NP_2}$ maturai(-y.il)]^{F}]
[$_{S}$ [$_{NP_1}$ I.NOM be_born-PAST-VN-NOM] [$_{NP_2}$ Madurai-NOM(-LOC)]^{F}]
'Madurai is where I was born.' (cited from Annamalai & Steever 1998: 123) [TNT]

(21) mani pooRatu
Mani.NOM go.PRES.VN.NOM
'Mani's going' (cited from Gair 1986[1998]b: 156) [TNT]

As discussed, the Sinhala adjectival/nominal participles appeared not only in the 'impersonal' constructions discussed above, but also as clausal nominalisations (see examples (2)–(7) above). In Classical Sinhala this function as a clausal nominalisation remains a prominent function of the adjectival/nominal participle, as shown by the examples below.

(22) a. ayināḍan **gatte** diḷiňdu veyi
not-given-thing **take**.PAST.PTCP.SG.NOM poor become.PRES.3SG
'He who takes ungiven things becomes poor.' (*Ama.* 68) [CS]
(cited from Wijemanne 1984: 157)

b. vaṭālā śilpa ugannaṭa **enno** no veti
arts learn.PRES.PTCP.DAT **come**-PRES.PTCP.3PL.NOM not become-PRES.3PL
'There will not be people who come to learn arts again.', lit. 'To-learn-arts-coming (ones) will not be again' (*Ama.* 131) (cited from *Ibid.*) [CS]

c. ætam budusavu kenek ... taman **maranuvan** piriyesti
disciple.NOM.PL Buddha.GEN some.NOM ... self.PL.OBLQ **kill**.PRES.PTCP.PL.OBLQ seek.PRES.3PL
'Some disciples of the Buddha ... search for **those who kill** them.' (*Ama.* 191) [CS]

d. manah-karma mahatæ yi **kiyanne** nokalakireṇu nam ve
mind-action important QUOT **say.PRES.PTCP.SG.NOM** not_backslide.ADJ EMPH be.PRES.3SG
'The person who says that the action of the mind is the important thing will not backslide.' (*Ama.* 53)
(cited from Gair 1986[1998]b: 167) [CS]

e. ovun no sanhinduva badu karavuvara **ganne** daham
them.ACC NEG calm.PRES.PTCP taxes royal-revenue **take.PRES.PTCP.SG.NOM** proper-doctrine(*dharma*)
noweyi
NEG.be.PRES.3SG
'Taking the royal revenue without calming them is not proper.' (*Ama.* 103) [CS]
(cited from *Ibid.*)

The Old Sinhala "impersonal" constructions have undergone change and appear to now be specialised for focus. I note several differences in the Classical Sinhala "focussing" construction from the earlier Old Sinhala "impersonal" construction. The first is these is that *ya*(/*yæ*) or an (overt) copula has become obligatory in these constructions. Further, the copula/*ya* now must immediately follow the focussed constituent—not the participial itself as in Old Sinhala (see above Section 11.1.1). Thus it is clear that this construction is no longer treated as impersonal, but rather as a copular construction, with *ya*(/*yæ*) acting either as a copula linking the nominalised clause and the focussed constitutent, or else as a predicator which transforms the focussed element into a predicate which can apply to the nominalised clause. However, like Old Sinhala and unlike later modern literary Sinhala, the focussed element can appear either to the right of the verb, as in (23a), or else on the left-edge of the clause, as in (23b).

(23) a. taṭa **karuṇe** [mahat lābha]^{F} yæ
you.DAT **do.PAST.PTCP.NOM** [great fortune]^{F} 3SG
'What has been done for you is a great fortune.' (*Ama.* 245) (cited from Paolillo 1994: 161) [CS]

b. [ovun saňdahā]^{F} yæ mā dan **denne**
[they.ACC for]^{F} 3SG I.ACC alms **give.PRES.PTCP.NOM**
'It is for them that I am giving alms.' (*Ama.* 107) (cited from *Ibid.*) [CS]

The obligatory presence of *ya* or a copula is not itself very reminiscent of either the Old Tamil or modern Tamil focus construction, in which the copula is non-overt. The similarity between the above Classical Sinhala focus constructions and the Tamil constructions is that they are both treated as copular constructions (where in Tamil the copula is simply non-overt).[8]

A further development can be seen in the use of accusative case-marking for the subject of the nominalised verb, see example (23b) above. The analysis of *mā* as accusative rather than genitive requires a few remarks about the history of Sinhala nominal morphology. Already in Old Sinhala, the old case system was starting to break down, in the sense that the distinction between accusative and genitive was being eroded, with the extension of genitive endings to accusative case nouns and ponouns.

By the period of Classical Sinhala, the distinction between accusative and genitive case-forms began to be renewed by the addition of *-ge* to the now ambiguous genitive/accusative forms of animate nouns and pronouns, with the old simplex form becoming restricted more and more to accusative-case only. Etymologically, *ge* derives from the locative form of (vulgar) Old Indo-Aryan *gehá* 'house' (loc. *gehe*) < older *gṛ̥há* 'house'. Thus *ge* originates as a locative used with reference to physical location. Even in Old Sinhala *ge* is found, though almost always in proper names, e.g. (24).

(24) diyavāṇa sivala-malun **ge** siri-devu gī
Diyavana.GEN Sivala Mala.GEN **of the house of** Siridevu.GEN song
'The song of Siridevu **of the house** of Sivala Mala of Diyavana.' (*S.G.* 263)

Ge appears only once in Old Sinhala with a pronoun, foreshadowing its later usurption of the genitive function, in *atak hæ**ge*** 'a hand of hers' (*S.G.* 104). Even in Classical Sinhala *ge* is not completely obligatory as both nominal and pronominal genitives are found with *ge* (*baṁbahu ge tepul* 'the brahmā's word' (*Ama.* 297), *mahaṇa goyumhu ge savu* 'a follower of the monk Goyum' (*Ama.* 90), *nahannahu ge jaṭāyehi* 'in the hair of him who was bathing' (*Ama.* 85), *mā ge kāryayaṭa* 'for my business' (*Ama.* 287), *topa ge mahaṇa* 'your monk' (*Ama.*154)) and without *ge* (*sorahu atæ* 'at the hand of the robber' (*Ama.* 133), *ætam budusavu* 'disciples of the Buddha' (*Ama.* 191), *ma put'haṭa* 'to my son' (Ama 82)). But its use obviously can serve to disambiguate genitives from accusatives—and by the modern period (both Literary and Colloquial) *-ge* appears to have taken on the status of an affix, and is the invariable ending of animate genitives.[9]

This use of *ge* parallels the use in Tamil of *il* (etymologically also 'house'), at least in its earliest function as a locative marker (see above example (15)), suggesting that by the end of the 10th century and certainly by the 13th century there is good evidence for structural change in Sinhala induced by contact with Tamil; though reanalysis/grammaticalisation of words meaning "house" as generalised locative markers is not unknown elsewhere, cp. the French preposition *chez* (e.g. *chez moi* "at/to my place", *chez le boucher* "at/to the butcher's" etc.) < Latin *casa* "house", with more recent developments in some varieties of *chez* X to mean "in the possession of X"—a development which neatly parallels that of Sinhala *ge* ("house" > locative-type (pre/post)position > genitive-type (pre/post)position).[10] As Paranavitana (1956: §362) remarks, '[t]he change in meaning from the loc. to the gen. is a natural one: what is one's house is one's own, so the postposition *ge* came to denote ownership'. However, in the

[8]Note that the similarity between the Classical Sinhala construction and the Dravidian focus construction is more obvious in the case of Malayalam, where the copula *āṇū* obligatorily follows the focussed element—see fn.6 above and example (i) below:

(i) innale bālan^{F} āṇū vannatu?
yesterday Balan^{F} be came.NOM
'Was it Balan who came yesterday?' (cited from Lindholm 1972: 306) [MM]

[9]Apparently with further phonological development in the colloquial language in the first person plural pronominal series, where the *g* has disappeared.

[10]See Harrison & Ashby (2003) for extensive discussion of various developments of French *chez*. The French varieties in question where *chez* can indicate possession are spoken in Lebanon and Sub-Saharan Africa; Harrison & Ashby (2003: 393) provide the example *"Où est ton crayon?" "**Chez** Marie. Regarde! Elle dessine avec."* ("Where is your pencil?" "*Chez* Marie. Look, she's drawing with (it).")

graffiti texts (which contain hundreds of genitives) it is found only once with pronoun (see above). In other words, in Old Sinhala *ge* has a very limited distribution, occuring almost solely in proper names, and the old simplex form of the genitive is almost always the form used elsewhere, despite the frequent homophony between accusative and genitive forms. By the modern period, the genitive in *ge* has become the standard form for both nominals and pronominals, with the old homophonous form being restricted to the accusative, thus resolving any potential ambiguity between accusative and genitive forms.

However, in focus constructions in Classical Sinhala, overt subjects of the nominalised clause always appear with the marking of the old simplex genitive (which by modern Sinhala clearly becomes the marking of the accusative), suggesting that the structure has been reanalysed as requiring accusative case-marking on the subject (if present).

As Paolillo (1994) notes, the pattern exhibited by the examples in (23) is found only in focussed sentences where the focussed constituent is a non-subject. Compare the subject-focussed examples, (25) below, with (23) above, in which the focussed element is a non-subject.

(25) a. [mama]^{F} yæ maṅḍavā **pīyemi**
[I.NOM.SG]^{F} 3SG **trample**.PAST.PTCP.MSC.NOM.1SG
'It is I who trampled it.' (*Ama.* 152) (Wijemanne 1984: 171) [CS]

b. [buduhu]^{F} yæ maṭa saṅgraha saṅdahā meyaṭa **æḍiyaha**
[Buddha.NOM.PL]^{F} 3SG I.DAT favour for this.DAT **come**.PAST.PTCP.3P.MSC.PL.NOM
'It is the Buddha who has come here to favour me.' (*Ama.* 135) (Wijemanne 1984: 172) [CS]

c. [paviṭ diṭ'ha]^{F} yæ taṭa **upanne**
[sinful thought.NOM.SG]^{F} 3SG you.DAT **arise**.PAST.PTCP.NOM
'It is an evil thought that has arisen upon you.' (*Ama.* 175) (*Ibid.*) [CS]

In the examples above in (25), where the subject is the focussed element, we do not find the invariant third-person, singular form of the participle. Rather the participles shows inflection for number and person just as a verb of a non-focussing construction does.[11] This represents another divergence of the Classical Sinhala focus construction from the earlier impersonal participial construction of Old Sinhala, where in the latter even focussed subjects fail to control number/person marking on the participle (see above example (8)).

Paolillo (1994: 168n16) suggests that the *yæ* which appears in the examples in (25) is not the same *yæ* of (23), but rather a homophonous element deriving perhaps from a form of *ye-* "to speak". This seems an unsatisfactory analysis; there is no reason to assume that the *yæ* of the examples in (25) is not the same *yæ* of examples like (23).

The proper synchronic analysis of examples like (25a) above appears to require a biclausal structure, as in (26). Note that in (26) each clause has its own finite verb, which strongly suggests a biclausal analysis.

(26) [mama yæ] [maṅḍavā **pīyemi**]
[I.NOM.SG COP.3SG] [**trample**.PAST.PTCP.MSC.NOM.1SG]
Lit. '(It) is I. (I) trampled it.' (*Ama.* 152) [CS]

Excepting examples of focussed subjects (as in (25)), focus constructions in Classical Sinhala thus appear to be similar to those found in Tamil in certain respects. Most importantly they represent clausal nominalisations put into a copular relationship with the focussed constituent, just as in Dravidian.

[11] Note that the analysis of (25c) is ambiguous, since the subject of this sentence is a third-person singular the participle's appearing with no further agreement morphology could be considered an agreeing form (since this is the expected agreement morphology for a third-person singular subject) or a non-agreeing form (since the third-person singular morphology is the default form).

11.1.3 Focus Constructions in Modern Literary Sinhala

The structure of modern literary Sinhala focus constructions is very similar to what we found for Classical Sinhala. As in Classical Sinhala, the presence of the default third-person singular 'agreement clitic' *ya* or a form of the copula is obligatory.

(27) a. mā kiyevuvē [ema pota]^{F} ya
I.ACC read.PAST.E [that book]^{F} COP.3SG
'It was that book I read.' [LS]

b. mā kiyevuvē [ema pota]^{F} veyi
I.ACC read.PAST.E [that book]^{F} be.3SG
'It was that book I read.' [LS]

c. *mā kiyevuvē [ema pota]^{F}
I.ACC read.PAST.E [that book]^{F}
'It was that book I read.' [LS]

As in Classical Sinhala, the subject of the clause of the E-verbal form must appear in the accusative case (and at this stage the distinction between accusative and genitive is clearly demarked).

(28) *mama kiyevuvē [ema pota]^{F}
I.NOM read.PAST.E [that book]^{F}
[LS]

Also like Classical Sinhala, the focussed element cannot remain in situ in its base-generated position, but must be dislocated. However, unlike Classical Sinhala, dislocation to the left edge of the clause is no longer an option; the focussed element must appear to the right of the verb.

(29) *[ema pota]^{F} ya mā kiyevuvē
[that book]^{F} COP.3SG I.ACC read.PAST.E
[LS]

Another difference from Classical Sinhala is that even focussed subjects appear in the accusative case and do not control number/person agreement on the predicate, see (30)—cp. with (25).

(30) ema pota kiyevuvē mā^{F} ya
that book read.PAST.E I.ACC^{F} COP.3SG
'It was I who read that book.' [LS]

The appearance of accusative-case subjects of focussed clauses in literary Sinhala seems to be part of more general bidirectional entailment condition on nominative case. Gair (1995[1998]: 253) states this in the form of a two-part rule:[12]

> Nominative Case is assigned only by AGR, represented by verbal agreement. It surfaces as nominative case inflection.
>
> Accusative (oblique) Case inflection occurs on nominals that are licensed in a syntactic construction but are not assigned Nominative Case (or some lexically assigned case).

[12] cp. Gair 1968[1998], 1992, 1995[1998].

Since E-verbal forms do not inflect for person/number, they show no verbal agreement (and in the terms of Gair 1995[1998] lack an AGR projection), which is the only possible assigner of Nominative Case in literary Sinhala.

By this stage of the language, the direct connection between E-verbal forms and other nominalisations has been lost. As discussed above at (19a), neither literary nor colloquial modern Sinhala utilise E-verbal forms in the construction of nominalised clauses (unlike both Old and Classical Sinhala, discussed above in Sections 11.1.1, 11.1.2), instead employing either special 'gerund' forms of the verb or adjectival participles.

So are literary Sinhala focus constructions still to be treated as nominalisations ('clefts')? Despite the loss of the E-verbal form as a possible nominalisation in other constructions, other facts suggest that literary Sinhala may still treat focus constructions as involving nominalisation of a clause, which is put into a copular relationship with the focussed element. Both the lack of agreement and the use of accusative case for the subjects of literary Sinhala focus clauses are properties which surface elsewhere as associated with nonverbal predicates (see discussion above, as well as Gair & Paolillo 1989[1998]).

11.1.4 Development of the Modern Colloquial Sinhala Focus Construction

In literary Sinhala we find the absence of subject-controlled person/number morphology on verbs in focus constructions, along with the assignment of accusative case to subjects of focussed clauses—in contrast to the assignment of nominative case to subjects and presence of agreement morphology on the verb in non-focussing clauses. This argues that though in modern literary Sinhala focus constructions may remain true 'clefts' as in earlier Classical Sinhala (and as in Dravidian), the situation is rather different in modern colloquial Sinhala. As discussed above in Section 4.1, there are two other crucial differences between focus constructions in literary and colloquial Sinhala: (1) dislocation of the focussed element (to the right of the verb) is no longer obligatory in colloquial Sinhala (compare (31b) and (31d) against (31a)), and (2) the element *y(i)* (< the earlier 'agreement clitic' *ya/yæ*) is (a) no longer obligatory (compare (31b) and (31d) against (31a) and (31c)), and (b) no longer substitutes for the copula, but instead may be replaced by other focus-marking particles like *tamaa, tamay* (see (31b) and (31d) below).

(31) a. mamə gamə̣ə̣^{F} yanne
I village^{F}.DAT go.PRES.E
'It is to the village I go.' [MCS]

b. mamə gamə̣ə̣^{F}-y(i)/tamaa/tamay yanne
I village^{F}.DAT-EMPH go.PRES.E
'It is to the village I go.' [MCS]

c. mamə yanne gamə̣ə̣^{F}
I go.PRES.E village^{F}.DAT
'It is to the village I go.' [MCS]

d. mamə yanne gamə̣ə̣^{F}-y(i)/tamaa/tamay
I go.PRES.E village^{F}.DAT-EMPH
'It is to the village I go.' [MCS]

Further, up to the stage of literary Sinhala, *yi/ya* is not an isolated element, but rather is part of a full paradigm of 'agreement clitics' (see Gair 1968[1998]: 216): *yi/ya*=third-person singular, *ti*=third-person plural, *mi*=first-person singular, *mu*=first-person plural etc. But with the disappearance of verbal agreement morphology, *yi* becomes an isolated element which appears almost only in focus constructions.[13]

This leaves *yi* ripe for reanalysis as a focus-marking particle rather than an agreement-clitic, especially since colloquial Sinhala lacks verbal agreement morphology.

It seems likely that these developments represent an acquisition-based change, as from the point of view of (child) language-learners, it makes sense to analyse the E-verbal form of focussing clauses as finite (parallel to the 'neutral' A-verbal form) and *yi* as a focus-marking particle rather than an agreement-clitic. Since colloquial Sinhala lacks (overt) agreement morphology on ALL predicates, this means that the E-verbal forms of focussing clauses are no longer exceptional in their lack of subject-predicate agreement morphology on the verb—and thus there is less motivation to analyse them as nominalised verbal predicates rather than simply as finite forms of the verb. Further, the loss of verbal agreement morphology in all contexts breaks the connection between subject-predicate agreement and assignment of nominative case, with the apparent result that subjects of both 'neutral' and 'focussing' clauses are assigned nominative case in colloquial Sinhala. Between the lack of agreement morphology in all contexts and the uniform case assignment of subjects of both neutral and focussing clauses, there is no longer sufficient motivation for language-learners to treat E-verbal forms as nominalised verbs nor *yi* as agreement morphology. This leads to the reanalysis of E-verbal forms as finite, and *yi* as a focus-marking particle, with the consequence that dislocation of the focussed element and the appearance of *yi* both become optional.

Thus, modern colloquial Sinhala focus constructions, unlike those of Classical and modern literary Sinhala, involve neither obligatory movement of the focussed element nor the appearance of a copular element. Semantically, however, as discussed above in Chapter 5, they do still behave as clefts, in the sense that the focus is associated with an existence presupposition.

11.1.5 Excursus: The origin of the modern Sinhala neutral *-a* verbal form

The "focussing" *-e* verbal forms of modern Sinhala, as discussed above, comes from the masculine/neuter NOMINATIVE singular case-form (whose suffix takes a variety of forms depending on the conjugation class, but prototypically appears as *-e*) of Old Sinhala. The "neutral" *-a* verbal forms are of a more opaque origin; however, they appear to descend from an OBLIQUE case-form of the participial.[14]

[13] The exception to this is that vowel-final adjectives in predicate position appear with final *-y*, compare (i) against (ii), (iii) below (cited from Gair & Paolillo (1989[1998]: 92–93), who refer to *y(i)* as a 'assertion marker').

(i) mee potə alut
this book new
'This book is new.' [MCS]

(ii) mee potə hoňda-y
this book good-*yi*
'This book is good.' [MCS]

(iii) mee kææmə rasa-y
this food tasty-*yi*
'This food is tasty.' [MCS]

[14] Certainly by the time of Classical Sinhala and even in Old Sinhala, distinguishing between accusative and genitive case-forms can be difficult; we find syncretism of the accusative and the genitive (adopting the morphology of the genitive), resulting in later "renewal" of the distinction by the use of the accusative/genitive case-form followed by the postposition *ge* to mark the genitive. Therefore, especially in the discussion of Classical Sinhala examples I often use the gloss OBLQ "oblique" to indicate the use of the syncretic form used by both accusative and genitive case nominals. See further Section 11.1.2 above.

One context in which oblique forms of participles are found in early Sinhala is preceding finite forms of the root *ve-* "be, become"—as apparently was first noted by Reynolds (1964: 145), who cites examples such as (32). I suggest that it is this sort of construction which may have been the ultimate origin for the "neutral" *-a* verbal forms of modern Sinhala.

(32) buduhu ma ohu **dæmuvā** vet
Buddha.NOM.PL EMP he.SG.OBLQ **tame**.PAST.PTCP.SG.OBLQ became.PRES.3PL
'The Buddha himself **tamed** him'. (*Ama.* 246.33) [CS]

Wijemanne (1984: 170-1) also notes this usage, remarking that it is found with past but not with present participles, (33).

(33) a. me se kiyanuye ... 'nivanudu næti' yi **kīvā**
this manner say.PRES.PTCP.SG.NOM ... extinction.NOM.SG is-not.PRES.3SG QUOT **say**.PAST.PTCP.SG.OBLQ
ve
be.PRES.3SG
'The person who says thus **has said** "there is no extinction."' (*Ama.* 296) [CS]

b. (idin to baṁbahu ge tepul ikmiṇihi nam,) geṭa siri vadut mohol **gattā**
house.DAT Śrī enter.COND club **take**.PAST.PTCP.SG.OBLQ
veyi
be.PRES.3SG
'(If you go beyond the brahma's word,) [it is like] one **took up** a club when the goddess of prosperity enters the house.' (*Ama.* 297) [CS]

In Old Sinhala, the variety of different forms often makes it difficult to determine the case of a participle: in addition to the clearly masculine nominative ending *-e*, participles often appear in stem-form, or stem+*a*. These, however, appear to be accusative-forms; and, in the following two examples, (34) and (35), we find the same verb appearing multiple times with different endings. These differences appear to reflect whether the participle is used as participant-nominal as subject (nominative case) or as a simple predicate (accusative case).

(34) ataṭ ho a men surat-atni **gat** sapu-mal
hand.DAT she.SG.FEM.NOM come.PAST.PTCP.ACC? like rosy-hand.INSTR **took**.PAST.PART.ACC? *sapu*-flower
känen naham gata he ja mana **gattī**
cluster.INSTR not take.INF able.PRES.PTCP.NOM.SG though mind.ACC **take**.PAST.PTCP.NOM.SG.FEM
muḷullen
whole.INSTR
'She **took**, in (her) rosy hand, a *sapu*[15] flower like (it) came (by itself) to (her) hand. Though (she) (is) one who is unable to take (it) from the cluster, (she) (is) **one who has taken** (my) mind in its entirety.' (*S.G.* 313) [OS]

[15]Skt. *campaka*, Michelia Champaca.

(35) pere tā no **laddiya** dæyek no **lad** veye
formerly you.GEN not **obtain**.PAST.PTCP.NOM.FEM thing.INDEF.NOM not **obtain**.PAST.PTCP.ACC? be.PRES.3SG
ranaṭ / jo kā **lada** ayut sihineka ahas
gold.DAT / Q who.NOM?/ACC? **obtain**.PAST.PTCP.ACC? come.COND dream.INDEF.NOM? sky
lada vey
obtain.PAST.PTCP.ACC? be.PRES.3SG
'**She who had not been obtained** by you earlier is a thing not **obtained** for gold. Who, having come, **obtained** (her)? A dream **obtained** the sky.'[16] (*S.G.* 609) [OS]

Here it appears that we have the same pattern as in the case of Classical Sinhala: the oblique (accusative) form appears when used as a simple predicate. However, in Old Sinhala, the oblique participle is not always followed by an overt form of *ve-*. In (35) the oblique forms are followed by *ve-* in two of the instances, but not in the third or in the case of the oblique participle in (34).

Though this account remains speculative at this point, the path of development seems to be that the oblique forms of the past participle—which apparently act as simple predicates when combined with a form of *ve-* "to be"—thus appear to be the origin for the modern "neutral" *-a* forms. In the colloquial language this use of the oblique form appears to have been generalised, as it is also found in the present tense (Classical Sinhala retains the old finite present, with subject/verb agreement morphology intact, in "neutral" contexts).

11.1.6 Summary and Conclusions

Modern colloquial Sinhala focus constructions, as discussed above in Chapter 4, show formal divergence from the superficially similar constructions of Dravidian and earlier Sinhala. Table 11.2 summarises the properties of these constructions in various stages of Sinhala and in Tamil and Malayalam. Note in particular the differences between modern colloquial Sinhala and Dravidian with respect to whether the focussed constituent is obligatorily displaced when the participial/focussing form of the verb occurs.

[16] Paranavitana (1956) translates the second line rather differently: "'...Having come (here), who has been obtained by you?' 'The sky has been obtained in a dream."' However, *kā* could be either nominative or accusative, and *lada* lacks the second person agreement *-hi* we would expect if there was an implicit subject *you*. Secondly, *sihineka*, Paranavitana treats as an indefinite locative, presumably taking the *-e-* of *-eka* to be the locative ending which appears with inanimates, e.g. *gire* 'on the mountain' (cp. nom/acc *giri*). However, case endings seem to appear following the indefinite suffix—which appears as *-ek(a)* or *-ak(a)* with inanimate nominative/accusatives—as in *katak'hi* 'in the beloved' (*kat(a)* + *-ak(a)*-feminine oblique indefinite marking + *-hi* locative case marking). So *sihineka* appears to be a nominative/accusative form *sihina* 'dream' with the nominative/accusative inanimate indefinite marker *-ek(a)*, and thus seems to be a subject rather than a locative adjunct. Further, the subject of the sentence is unclear otherwise.

	Old Sinhala	**Classical Sinhala**	**Mod. Lit. Sinhala**	**Mod. Colloq. Sinhala**	**Tamil**	**Malayalam**
Use of participial forces displacement of focussed constituent	no	yes	yes	no	yes	yes
Participial used in "focussing"/"clefting" constructions also is used for general nominalisation	yes	yes	no	no	yes	yes
Participial/"focussing" verbs appear with copular element	optional	obligatory	obligatory	none	non-overt	obligatory
Position of copular element	after participle	after focus	after focus	n/a	n/a	after focus
Subject of participial/"focussing" verb takes	genitive	accusative/ old genitive	accusative	normal case	normal case	normal case

Table 11.2: Properties of Sinhala and Dravidian "clefting"/"focussing" constructions

In summary, the Sinhala focus construction has its origins in the "impersonal" construction we find in Old Sinhala. The *-e* marking that appears on the verb in focussing contexts in modern Sinhala derives from the nominative singular case-form of the masculine/neuter participle of Old Sinhala, a nominal form which up until modern Sinhala also functioned as a general means of nominalising verbs. The Old Sinhala "impersonal" construction involves the use of a participle (with or without a following copular element), with the logical subject of the nominalised verb appearing optionally in the genitive case (which becomes the accusative case of later Sinhala). This construction seems to have been compatible with the presence of focussed elements, but signalling the presence of a focussed element does not appear to have been its primary function.

In the next stage of the language, Classical Sinhala, the "impersonal" construction has been reanalysed as directly tied to the presence of a focussed element. Here a copular element is obligatory, and appears immediately following the focussed element, which itself occurs either to the right of the participial verb or elsewhere on the left edge of the clause. Logical subjects of the participial clause, if present, occur with accusative case-marking (in form, the old simplex genitive case-marking). Modern literary Sinhala presents a similar picture, with the exception that the *-e* form becomes morphologically-isolated, as the old participials are no longer used for verbal nominalisations in general. The formal structure of focus constructions in Classical and modern literary Sinhala share a number of properties with focus constructions in Dravidian, and may well represent in part the result of language contact.

In modern colloquial Sinhala, however, we find that the focus construction acquires properties rather different from those of Dravidian. The focus constructions of Classical and modern literary Sinhala do appear to be "clefts" in the sense that they involve a clause put into a copular relationship with another element (the focussed element). Modern colloquial Sinhala focus constructions do not involve any copular element, nor is displacement of the focussed element obligatory.

Thus, the "impersonal" constructions of Old Sinhala—the predecessors of the later focus constructions—are monoclausal and bear little real resemblance to Dravidian focussing constructions. In Classical Sinhala and modern literary Sinhala we find that the "impersonal" construction of earlier Sinhala has been reanalysed as a biclausal "cleft" tightly associated with the presence of a focussed element, a development which quite likely reflects Dravidian influence. Modern colloquial Sinhala focus constructions appear to be the result of language-internal developments, as they no longer are analysable as biclausal.

The co-existence of changes in Sinhala resulting in convergence with and divergence from Dravidian is consistent with the findings of Gair (1976[1998], 1985[1998], 1986[1998]a). While Gair (1985[1998]) suggests that—even though Sinhala has undergone a number of phonological changes setting it off from other Indo-Aryan languages—the majority of these changes cannot be attributed to Dravidian influence,[17] he notes (Gair 1976[1998], 1986[1998]b) a number of changes in Sinhala syntax which appear to reflect convergence with Dravidian. Such similarities include the use of conjunct affixes such as the conditional and concessive and the employment of adverbial clauses formed from verbal adjectives followed by certain particles (Gair 1976[1998]: 208), as well as the transformation of Sinhala into a more consistently left-branching language, like Tamil and other Dravidian languages, and unlike its Indo-Aryan relatives, which generally present more mixture of left- and right-branching structures (Gair 1985[1998]: 199).[18] In the case of the development of Sinhala focus constructions, the mixture of convergence with and divergence from Dravidian would appear to reflect the scenario suggested by Gair (1985[1998]: 196) in which there were historical periods of strong Dravidian influence on Sinhala alternating with periods of weak or absent influence.

The "neutral" *-a* verbal forms of modern Sinhala appear to derive from oblique case-forms of the participle—in

[17]One example of a change which results in greater phonological similarity between Sinhala and Tamil (and other Dravidian languages) is the development of length contrast for *e* and *o* (which likewise makes it unlike other Indo-Aryan languages).

[18]For further discussion of syntactic similarities between Sinhala and Dravidian (not shared by other Indo-Aryan languages), see also Elizarenkova (1972: 135) and Ratanajoti (1975: 117–126).

early Sinhala we find oblique participial case-forms employed preceding forms of "to be".

In the next section I turn to the examination of the historical development of the elements which have been the primary concern of this study: Q-particles.

11.2 The development of the Sinhala Q-particles

In this section I investigate the sources of the Sinhala Q-particles *də* and *hari*, and examine their development from Old Sinhala into modern colloquial Sinhala, including their spread into new environments and the development of the "ignorance component" presuppositions associated with *də* and *hari* indefinites.

11.2.1 Etymology of Sinhala *də*

The Sinhala Q particle *də* derives ultimately from Old Indo-Aryan *utā́ho* (Turner 1962–1966: #1701), which is made up of two particles, *ā́ho* and *utá*.

Utá is a particle used in Vedic with the sense 'and' (Klein 1978), with both X *utá* Y (36) and X Y *utá* (37) orders.

(36) mitró dādhāra pr̥thivī́m **utá** dyā́m
contract.MASC.SG.NOM hold.PRES.3SG earth.SG.ACC ***uta*** heaven.SG.ACC
'Contract holds (together) earth and heaven.' [RV 3.59,1b]

(37) divás pr̥thivyā́ **utá** carkirāma
heaven.SG.GEN earth.SG.GEN ***uta*** praise.SUBJ.1PL
'Heaven and earth we shall praise.' [RV 4.39,1b]

In Classical Sanskrit, this conjunctive sense of *uta* is obsolete (Speijer 1886: §424), though *uta*—like *api* 'too, also, moreover, and; even; though'—may stand at the beginning of a clause as in (39), functioning as a Q particle like Sanskrit *kim*, see (38).[19]

(38) kim śaknoṣi? / śaknoṣi kim?
Q be-able-PRES.2SG / be-able-PRES.2SG Q
Can you? (Speijer 1886: §412)

(39) **uta** daṇḍaḥ patiṣyati?
uta stick-NOM.SG. fall-FUT.3SG
'Will the stick fall?' (Kāśikā on Pāṇini 3,3,152)(Speijer 1886: §412)

[19]Sanskrit *kim* is the nominative/accusative case-form of the neuter interrogative *wh*-pronoun which also is used as a *wh*-interrogative pronoun in the sense "what". Modern Hindi displays a similar phenomenon: the inanimate *wh*-interrogative pronoun *kyā* "what" can also optionally stand at the beginning or end of a yes/no-question:

(i) a. Rām yahā̃ hai kyā?
Ram here be.PRES.3SG Q
"Is Ram here?"
b. Kyā Rām yahā̃ hai?
Q Ram here be.PRES.3SG
"Is Ram here?"

The particle *áho* first appears in late Vedic Sanskrit, positioned at the front of the second clause in a disjoint 'either. . . or' construction, see (40). Note here that *utá* appears at the front of the first clause of the disjunction.[20]

(40) **utá** _avidvā́n amúṁ lokáṁ prétya kaścaná gacchatī3 / **áho**
uta one who does not know.NOM.SG yonder.ACC.SG worldACC.SG depart.GER anyone go.PRES.3SG / ***āho***
vidvā́n amúṁ lokáṁ prétya kaścit sámaśnutā3i /
one who knows.NOM.SG yonder.ACC.SG world.ACC.SG depart.GER anyone reach.PRES.3SG /
'Does anyone who does not know, having died, go to yonder world, or does anyone who knows, having died, attain yonder world?' (Taittirīya Upaniṣad 2.6, cited from Böhtlingk & Roth 1855–1875)[21]

It is from this *utá* A B C . . . *áho* XYZ construction that the form *utā́ho* seems to derive. Since *utá* may appear at the front of an interrogative clause (functioning as a Q particle), it could also immediately precede the *áho* standing at the front of the second part of a disjunction, with the application of sandhi becoming *utā́ho*. And it is in this form (*utāho*), that it appears in Classical Sanskrit, as in (41) below (note that here in addition to *utāho*, we find the normal disjunction *vā*).[22]

(41) kiṁ mama vadhopāyakramaḥ kubjasya vā_**utāho** anyasya vā kasyacit
Q me-GEN murder-plot.NOM.SG hunchback-GEN or_***utāho*** other-GEN or someone-GEN
'Is it I, against whom the murder-plot is laid, or is it the hunchback or somebody else?' (Panc. 332) (Speijer 1886: §415)

In Classical Sanskrit *utāho* may also occur without *kim* (or *vā*), e.g. (42).

(42) buddhi-bhedaḥ para-kr̥̊ta **utāho** te svato 'bhavat
intellect-pollution-NOM.SG enemy-do-PAST.PTCP.NOM.SG ***utāho*** you self-GEN become-PAST.3SG
'Has the pollution of your intellect been brought about by enemies or by you yourself?' (BhP 7.5.10)

---

[20]Disjunctive constructions involving *utá* exist in earlier Vedic verse (pre-dating the Vedic prose discussed above) where we also find *utá* combining with *vā* "or" in the sequence *utá vā*, as in (i); however, it is not clear if or how this differs semantically from *vā* alone.

(i) yā́ḥ ā́paḥ divyā́ḥ **utá vā** srávanti khanítrimāḥ **utá vā** yā́ḥ
REL-PRON.PL water.NOM.PL divine.ADJ.NOM.PL ***utá* or** flow.PRES.3PL produced-by-digging.NOM.PL ***utá* or** REL-PRON.NOM.PL
svayaṁjā́ḥ / samudrā́rthāḥ yā́ḥ śúcayaḥ pāvakā́s tā́ḥ ā́paḥ devī́ḥ
self-born.NOM.PL / ocean-goal.NOM.PL REL-PRON.NOM.PL shining.ADJ.NOM.PL clear.ADJ.NOM.PL those.NOM.PL water.NOM.PL goddess.NOM.PL
ihá mā́m avantu //
here me.ACC.SG watch-over.IMPV.PL //
"Those waters which are divine, or those which flow dug up from the earth, or those which are self-produced, with the ocean as their goal, shining and clear, may those water goddesses watch over me here." [RV 7.49,2]

Graßmann (1873) notes one example where *utá* co-occurs with *vā* but is separated from it:

(ii) idám āpaḥ prá vahata yát kím ca duritám
here(/this side/this place) water.VOC.PL away drive.IMPV.2PL REL-PRON.NEU.NOM/ACC.SG what.NEU.NOM/ACC.SG and/moreover fault
máyi / yát vā _ahám abhidudróha yát **vā** śepé **utá**
me.LOC.SG / REL-PRON.NEU.NOM/ACC.SG or I.NOM.SG hurt.PERF.1SG REL-PRON.NEU.NOM/ACC.SG **or** curse/swear.PREF.1SG ***utá***
ánr̥tam //
untruth/unlawfulness/chaos.NEU.NOM/ACC.SG //
"O Waters, here drive away whatever fault is in me: if I hurt (someone) or if I swore untruthfully." [RV 1.23,22]

[Geldner (1951–1957: 23) appears to prefer to treat *utá* as separate from *vā* and renders it as "and": "Ihr Gewässer, führet all das fort, was von Fehle an mir ist, sei es daß ich treulos war, oder daß ich geflucht habe und jede Unwahrheit!" ("You Waters, drive off all this, what fault is in me, be it that I was disloyal or that I cursed, and every untruth!")]

However, in these examples *utá* does not signal a question, and so there appears to be little connection between the early Vedic constructions with *utá vā* and the later Vedic prose construction *utá . . . áho* (which later becomes *utā́ho*, the ancestor of Sinhala *da*/*də*).

[21]3, e.g. in *gacchatī3*, marks *pluti*, the 'overlong vowel'.

[22]Speijer (1886: §414) remarks that instead of *utāho* or *āho*, we may also find *utasvit* or *āhosvit*. However, it is specifically from *utāho* that Sinhala *də* develops.

And *āho* by itself may appear in place of *utāho*, e.g. (43).

(43) teṣāṁ niṣṭhā ... kā ... sattvam **āho** rajas tamaḥ?
they-GEN state-NOM.FEM ... what-NOM.FEM ... purity ***āho*** passion darkness/ignorance
'What is their condition? Purity, passion or ignorance?' (BhG 17.1)

In Pāli, the distribution of *udāho* (< *utāho*) appears to be similar, cp. the Pāli example in (44) with the Sanskrit example in (41).

(44) kiṇ amhehi saddhiṇ āgamissasi **udāhu** pacchā?
Q us with come-FUT.2SG ***udāhu*** later
'Will you come with us or later?' (DhA ii.96)

Pāli *udāhu* also appears with *uda* (cognate with Sanskrit *uta*, see above) as the question particle:

(45) atthaṁgato so **uda** vā so natthi ... **udāhu** ve sassatiyā arogo
he who has reached the end **Q** or he not exist ... **udāhu** he for eternity free from disease
'He who has reached the end: does he not exist, or is he for eternity without disease?' [BJT p. 330]

As in Sanskrit, Pāli *udāhu* also occurs without the question particle *kiṇ*, as in the examples below.[23]

(46) saccāni sutāni bahuni nānā **udāhu** te takkamanussaranti
truths heard many various **udāhu** they follow conjecture
'Have they learned many various truths or do they follow conjecture?' [BJT p. 280]

(47) ettāvata _aggaṁ no vadanti ha _eke yakkhassa suddhiṁ idha paṇḍitāse **udāhu**
to that extent highest NEG say.PRES.3PL EMPH ones *yakkha*.GEN purification in this place pundits **udāhu**
aṁñampi vadanti etto
other say.PRES.3PL from this
'Do not some of the wise in this world tell us that the purity of the yakkha is the highest to this extent, or do they say something different from this?' [PTS p. 171]

Note that in early Sinhala, in addition to frequently appearing in yes/no-questions, *da* also is used in disjunctive interrogatives, (48).

(48) mā ... nuvaṭahu arabhayā kī dǣ nipan **da** no nipan **da**?
my ... religious mendicant about said things QUOT born ***da*** NEG born ***da***?
"Did my predictions regarding the religious mendicant prove correct or did they not?" (12th century, *Ama.* 178)
(Wijemanne 1984: 75)

[23]Occasionally the question particle *uda* itself seems to function as a disjunction, at least in interrogative contexts, as shown below:

(i) nirāsaso so **udi** āsasāno paṁñāṇavā so **uda** paṁñakapapi
without desire he **uda**? desiring discerning he **uda** still acquiring discernment
'Is he without desire, or desiring? Discerning or still acquiring discernment?' [BJT p. 336]

This use of *uda* could derive from 'ambiguous' uses (where *uda* could be rendered as either 'or' or 'and') of *uda* in examples like:

(ii) na brāhmaṇo no 'mhi na rājaputto na vessāyano **uda** koci no 'mhi
not *brāhmin* not is.PRES.1P not king's son not *vaishya*'s son **uda** any not is.PRES.1P
'I am not a *brāhmin* (=priestly caste), nor a king's son, nor any *vaishya*'s (=agricultural caste) son.' (Lit., 'I am not a *brāhmin*, nor a king's son. And I am not any *vaishya*'s son.' Or perhaps, treating *uda* as 'and': 'I am not a brahmin, not a king's son, not a vaishya's son. And I am not anybody.') [BJT 457]

In additional to its extension to other interrogative contexts, *da* exhibits other syntactic changes from its Sanskrit and Pāli precursors. Specifically, note that in Sanskrit and Pāli *utắho*/*udāhu* PRECEDE the disjunct, while in Sinhala *da* follows its disjunct. Thus the particle has undergone change from being a proclitic-type element to an enclitic-type element. The motivation for this change is not entirely clear, though one suspects that Dravidian influence may have played a role, given that Dravidian Q-particles are enclitics. The status of Q-particles as proclitics vs. enclitics is not a major concern of the present study, so I leave further investigation of this matter to future research.

Turning back to the developments in the distribution of this particle, the general path of change appears to be that the ancestor of *də* (cognate with Sanskrit *utāho*, Pāli *udāhu*) was reanalysed in Classical Sanskrit and Pāli as a disjunction used specifically in interrogative contexts. In prehistoric Sinhala, it was reanalysed as being a Q-particle; the distribution would have originally been restricted to alternative and yes/no-questions (I assume here that yes/no-questions are a special form of alternative question, see Chapter 6.4.2/6.4.3)—representing a widening with respect to its Sanskrit and Pāli precursors. Later, *da* was extended to the construction of relative clauses (existing as an alternative to the "conditional" particle *nam*), as in the modern literary Sinhala example in (49).[24]

(49) [**yam** kumariyak ohu duṭuvā] **da** [oo ohu kerehi piḷiňda sit ætikara
[**REL-PRON** princess.INDEF him see.PAST.3SG.FEM] ***da*** [she him towards connected mind developed
gattāya]
get.PAST.3SG.FEM]
"**Whatever princess** saw him fell in love with him." (cited from Gair & Karunatilaka 1974: 295) [LS]

And later still (in modern colloquial Sinhala) *da* (> *də*) also began to be used in the construction of certain types of indefinites, on which see Chapter 11.2.3 below, as well as Chapter 7 above.

It is of some interest that *utá* itself seems to have earlier been bimorphemic, where the initial element, *ŭ́*, originally functioned as a focussing particle.[25] Though Klein (1974) calls *ŭ́* 'anaphoric' (referring to an element already present in the discourse, rather than deitic), his discussion makes it clear that he analyses its original function as being that of focus: "we can render the *u* [in an English translation–BMS] by 'that very one' or simply by stress" (Klein 1974: 165).[26] Examples of *ŭ́* in this focussing function follow below:

(50) yáṁ tvám agne samádahas / tám **u**
REL-PRON.MASC/NEU.SG.ACC you.NOM.SG Agni.VOC burn-down.PAST.2SG / pron.MASC/NEU.SG.ACC ***u***
nír vāpayā púnaḥ
extinguish.IMPV.2SG again
'The one which you, O Agni, have burnt down, **that very one** extinguish again.' [RV 10.16,13ab]

(51) tám **u** ṣṭuhi _índraṁ yó ha sátvā /
pron.MASC.SG.ACC ***u*** praise.IMPV.2SG Indra.MASC.SG.ACC REL-PRON.MASC.SG.NOM PART warrior.SG.NOM /
yáḥ śū́ro maghávā yó ratheṣṭhắḥ
REL-PRON.MASC.SG.NOM hero.SG.NOM generous one.SG.NOM REL-PRON.MASC.SG.NOM chariot-standing.SG.NOM
'Praise **that very one**, Indra, who is a warrior, who is a hero, generous, who stands in a chariot.' [RV 1.173,5ab]

[24]Note that the extension to relative clauses appears to reflect Dravidian influence, cp. the Tamil example in (i).

(i) [yār aṅkē mutalil vantu ceru-v-ār]$_{RC}$-**oo** [avar ṭikeṭṭu vaṅkalām]$_{CC}$
[who.NOM there first come.CONV arrive.FUT.3PL]$_{RC}$-***oo*** [they.NOM.PL ticket.NOM buy.PERMISSIVE]$_{CC}$
"Let whoever reaches there first buy the tickets." (cited from Annamalai & Steever 1998) [Modern Tamil]

The use of *da* (and *nam*) in Sinhala relative clauses is discussed in more detail in Chapter 8, as is their relation to Dravidian relative clauses.

[25]Re-emphasising the semantic connection between focus and interrogation.

[26]Klein's notion of "anaphoricity" of *u* may be compared to Rooth's (Rooth 1985, 1992, 1996, 1999) idea that focus semantic values make reference to entities which are somehow already in the background of the discourse (the "alternatives" to a focussed element).

(52) tám **u** ṣṭavāma yáṁ gíra / índram ukthā́ni
pron.masc.sg.acc ***u*** praise.1pl rel-pron.masc.sg.acc song.pl.nom / Indra.masc.sg.acc hymn.pl.nom
vāvr̥dhúḥ
increase.past-perf.3pl
'**That very one** shall we praise: Indra, whom the songs (and) hymns have increased.' [RV 8.95,6ab]

(53) agníṁ vaḥ pūrvyáṁ huve / hótāraṁ carṣaṇīnā́m // tám
Agni.masc.sg.acc you.pl.dat first invoke.1sg / priest.masc.sg.acc person.pl.gen // pron.masc.sg.acc
ayā́ vācā́ gr̥ṇe tám **u** va stuṣe
this.sg.inst speech.sg.inst sing.1sg pron.masc.sg.acc ***u*** you.pl.dat praise.1sg
'I invoke Agni, the priest of the folk, for you. I sing (of) him with this speech. **That very one** [=Agni] I praise for you.' [RV 8.23,7]

(54) tám **v** abhí prā́rcata / índram sómasya pītáye
masc.sg.acc ***u*** over praise-forth.impv.2sg / Indra.masc.acc.sg soma.sg.gen drink.inf.sg.dat
'Unto **that very one**, Indra, sing forth the praises for the drinking of the *soma*.' [RV 8.92,5ab]

Later *ŭ* was reanalysed as a conjunction; examples with this function are given below:

(55) sáṁ te páyāṁsi sám **u** yantu vā́jāḥ
together you.sg.gen milk.nom.pl together ***u*** go.impv.3sg loot.nom.pl
'May the milks and the loots unite for you.' [RV 1.91,18]

(56) svàr yád áśmann adhipā́ **u** ándhaḥ
sunlight.nom.sg when rock.sg.loc master.sg.nom ***u*** darkness.sg.nom
'When the sunlight is in the rock, and darkness is master.' [RV 7.88,2c]

The particle *utá* seems to have been composed of this focussing particle *u* combined with **té* (the clitic form of the demonstrative pronoun), originally meaning *"that (very one) there" with later weakening to "and that" and finally to simply "and" (see Brugmann (1904: 100ff.) and Klein (1974: 179–180)).

11.2.2 Etymology of Sinhala *hari*, *hō*

Prior to modern colloquial Sinhala, the normal particle appearing in non-interrogative disjunctions is *hō* rather than *hari*, as in (57).

(57) yuvaraja-væ siṭiyavun **hō** . . . rāja-kumāra-varun **hō** . . . bisōvarun **hō**
heir-apparent ***hō*** . . . princes ***hō*** . . . queen ***hō***
"Either the heir-apparent . . . or the princes . . . or the queen" (12th c. inscription; Zilva Wickremasinghe et al. 1912–1933: ii.161B[5-7])

Geiger (1938: §163.2) suggests that *hō* derives from an earlier construction involving a form of "to be", comparing the Pāli disjunctive construction *hoti. . . hoti*.

The particle *hari* is proper only to the modern colloquial language, and is of uncertain origin. However, as, in the colloquial language, it takes over the same range of functions the particle *hō* fulfils in the modern literary language, it is semantically equivalent.

It is perhaps tempting to relate *hari* directly to *hō*, given that they both contain an initial /h/, i.e. to suppose that *hari* might represent a form of *hō* to which some extension or suffix has been added. However, no such extension/suffix suggests itself.

Another possibility is to relate *hari* to the word which appears in Classical and modern literary Sinhala as *sari*, of which *hari* may plausibly be a colloquial reflex.[27] *Sari* in fact means "similar" (cognate with Prakrit *sarisa*, Sanskrit *sadṛśa*, cp. Geiger 1941), and it is not immediately obvious to me why a word meaning "similar" would come to take on a disjunctive sense.

However, the development from "similar" to disjunction has another possible parallel in Sinhala, in the form of the disjunction *ekko*. Disjunctions in non-interrogative contexts may be formed in colloquial Sinhala either with *hari* following each disjunct, as in examples *passim* and (58) below,

(58) Gunəpālə **hari** Chitra **hari** gaməṭə giya.
Gunapala ***hari*** Chitra ***hari*** village.DAT went-A
"Gunapala or Chitra went to the village."

or else may be formed using *ekko* preceding each of the disjuncts (cf. Reynolds 1980: 69), as in (59):

(59) **Ekko** Gunəpālə **ekko** Chitra gaməṭə giya.
ekko Gunapala ***ekko*** Chitra village.DAT went-A
"Gunapala or Chitra went to the village."

The form *ekko* appears to be cognate with Classical/literary Sinhala *ekvan* "like, similar"(< Skt. *eka-varṇin*), thus there may be a case to be made for *hari* as a descendant of earlier *sari* "similar".

11.2.3 Spread of *də* and *hari* to new syntactic environments

In this section I examine the development and spread of the Sinhala Q-particles *də* and *hari* to new syntactic environments, as summarised in Table 11.3.

	Old Sin	Class Sin	Lit Sin	Colloq Sin
y/n-ques.	(da)	(da)	da	də
wh-ques.	(da)	(da)	da	də
wh-indef.	—	—	hō (aff. & neg.), vat (neg.)	də (aff.), hari (aff.), vat(neg.)
decl. disj.	hō, heva(-t)	hō, heva(-t)	hō (aff. & neg.), vat (neg.)	hari (aff. & neg.), vat (neg.),
interr. disj.	da	da	da	də

Table 11.3: Historical development of Q-particles in Sinhala

The particle *hō*, later in the form *hari*, originally was limited to the formation of non-interrogative disjunctions. This is true up until modern literary Sinhala, when *hō* begins to appear also in the formation of *wh*-based indefinites. This does not involve any change in formal syntactic features.[28]

The particle *da*, as suggested above, in pre-Sinhala would have been restricted to alternative disjunctions (as its Sanskrit and Pāli cognates were). The spread from alternative questions to yes/no-questions is natural, given our assumption that yes/no-questions are actually a special subtype of alternative question (see above, Chapter 6.4.2/6.4.3). From that point *da* seems to have been generalised as a "question particle" in the strict sense, i.e. as

[27] There were two separate changes of /s/ > /h/ in the history of Sinhala, on which see Gair (1985[1998]: 191–2) for more details. However, these changes left a great deal of residue, in the form of: (1) morphophonological alternations, e.g. modern colloquial Sinhala *mini**h**a* "man" but *mini**ss**u* "men"; *ga**h**a* "tree" but *ga**s*** "trees"; (2) variation between registers, e.g. literary Sinhala ***s**ata* "seven" vs. colloquial ***h**atə*; (3) variation within registers, e.g. literary Sinhala *raha* ~ *rasa* "tasty", colloquial Sinhala ***h**oodənəva* ~ ***s**oodənəva* "wash".

[28] I posit that at no stage does *hō*/*hari* bear any relevant formal syntactic features. (At least no features that need valuing or participate in the valuing of the features of other elements associated with disjunctions or indefinites etc.)

a particle employed in interrogatives, and thus is extended to *wh*-interrogatives. Given that by the earliest texts considered (the Old Sinhala of the Sigiri graffiti) *da* appears sporadically in both yes/no and *wh*-questions, this path of development must be considered hypothetical. It is important to note, however, that the extension of *da* to *wh*-interrogatives requires a semantic change in the definition of interrogative pronouns. As discussed in Chapter 9, in the Early Sinhala grammar I[29] interrogative pronouns are semantically-complex, in essence incorporating a Q-particle. Thus, the extension of *da* to *wh*-interrogatives requires the reanalysis of interrogative pronouns as semantically simple, i.e. as denoting simply Hamblin-type sets.

In formal syntactic terms, the change of environment from the earlier (grammar I) to the later (grammar II)[30] simply involves the extension of the features Q[ ],Int[+] from only the CP-heads of alternative questions to the CP-heads of all interrogatives, see the tables below.[31]

Category	Feature(s)
C-INT(wh)	
C-INT(y/n)	
C-INT(alt)	Q[ ], Int[+]
***wh*-interrog.**	
indef. pronoun	
da	Q[+], Int[ ]
hō	
J	

Table 11.4: Early Sinhala I feature assignments (repeated)

Category	Feature(s)
C-INT	Q[ ], Int[+]
***wh*-interrog.**	
indef. pronoun	
da	Q[+], Int[ ]
hō	
J	

Table 11.5: Early Sinhala II feature assignments (repeated)

The extension of *da*/*dɔ* to the formation of *wh*-based indefinites does not occur until modern colloquial Sinhala. This extension involves a more major change in the formal syntactic feature assignments: *da*/*dɔ*, which earlier has the feature assignment Q[+],Int[ ] (see Tables 11.4, 11.5) loses its unvalued Int[ ] feature (see Table 11.6)—a change which allows it to occur in non-interrogative contexts.

[29] Recall that both Old and Classical Sinhala show variation in whether *da* appears in yes/no and *wh*-questions; this variation is accounted for by positing two competing grammars (on which see e.g. Kroch 1989, Lightfoot 1999).

[30] Again, these two Early Sinhala grammars co-exist for some time.

[31] If this change happened in two steps, as suggested above, then the extension would have first affected the CP-heads of yes/no-questions, and then the CP-heads of *wh*-questions.

Category	Feature(s)
C-INT	Q[ ], Int[+]
***wh*-pronoun**	
də	Q[+]
hari	
J	

Table 11.6: Modern Colloquial Sinhala feature assignments (repeated)

Note that *hō*/*hari* appears in the formation of *wh*-indefinites before we find *də* used in *wh*-indefinites. This is not unexpected, given the formal feature analysis proposed here. The extension of *hō*/*hari* to indefinites requires no change in its formal syntactic features,[32] while the extension of *da*/*də* to indefinites requires a major change in its formal syntactic features. It therefore makes sense that *hō*/*hari* appears in the formation of indefinites earlier than does *da*/*də*.[33]

There remains the question of what motivates the extension of *hō*/*hari* to the formation of indefinites in modern literary Sinhala. As noted, such an extension requires no reanalysis of the formal syntactic features of *hari* or any other element. Further, once semantically-simple (i.e. Hamblin-type set denoting) *wh*-pronouns are available in the language, it is a reanalysis which likewise requires no semantic reinterpretation. So, a partial answer to the question of why the extension of *hō*/*hari* to the formation of indefinites occurs in modern literary Sinhala is simply that it was possible without any syntactic or semantic reanalysis.

This raises a related question, however: why did it not occur sooner, say in Old or Classical Sinhala? Recall that in both of these stages semantically-simple *wh*-pronouns co-exist[34] with semantically-complex *wh*-pronouns, the latter of which are incompatible with the co-occurrence of a Q-particle, since they do not denote Hamblin-type sets. By the modern literary Sinhala period, semantically-complex *wh*-pronouns are absent. It may be that the extension of *hō*/*hari* to the formation of *wh*-based indefinites becomes more likely when the only existing *wh*-pronouns are semantically-simple, Hamblin-type set denoting elements—which does not occur until modern literary Sinhala.

Another factor which may have played a role in blocking the extension of *hō*/*hari* to the formation of indefinites in early Sinhala is the fact that in Old and Classical Sinhala we find indefinite pronouns (*kisi*, *yam*) which are morphologically-distinct from the interrogative pronouns. Such forms are absent from modern colloquial Sinhala, but co-exist with the WH+*hō* indefinite in the modern literary language. It is only the *wh*-based indefinites which function as epistemic indefinites, and it is to the development of the pragmatics of these indefinite that I turn in the next section.

11.2.4 The origin of epistemic *wh*-based indefinites in Sinhala

I begin with *hari*(/*hō*), which, as previously discussed, was originally confined solely to non-interrogative disjunctions. As we have already seen in Chapter 7, the presupposition carried by modern colloquial Sinhala WH+*hari* indefinites involves the lack of an Important Predicate approximating basic intensional choice function which selects an individual who both satisfies the proposition and has the same extension in all epistemically-accessible worlds. The formal definition for *hari* is repeated below in (60).

[32] It does require that there exist semantically-simple *wh*-words; but that semantic change had already taken place once *da* was extended to *wh*-interrogatives.

[33] One might suspect, in fact, that it is the extension of *hō*/*hari* to indefinites which might have worked to enable the later extension of *da* to the same context.

[34] As above, I assume that such variation represents the co-existence of competing grammars.

(60) *hari*$_i$(g)(w)(F)(G)(Q)(S), where g is an assignment function, w is a world, F is a modal base supplied by context, G is a set of epistemically-accessible Important Predicates, Q, S are predicates

a. denotes: g(i)(w).BASICH(g(i))

b. "signals": $\neg\left[\exists$f.BASICH$_{\sim \text{ImpPred}}$(f) $\neg[\exists w', w'' \in F$: S(w′)(f(Q)(w′))=1 & S(w″)(f(Q)(w″))=1 & f(Q)(w′) ≠ f(Q)(w″)$]\right]$

Consider a non-interrogative disjunction like that in example (61) below, setting aside the pragmatic analysis proposed for *hari*—let us here allow (61) to stand in for its Classical Sinhala equivalent (with *hō* rather than *hari*), at which stage *hari*/*hō* carries no presupposition.

(61) Gunəpālə hari Chitra hari gaməṭə giya.
Gunapala *hari* Chitra *hari* village.DAT went-A
"Gunapala or Chitra went to the village."

A speaker who utters (61), asserts:[35]

(62) a. $\forall w \in F.\exists$f.BASISCH(f)[f(w)({Gunapala, Chitra}) went to the village in w] ≡

b. $\forall w \in F$.[Gunapala went to the village in w] ∨ [Chitra went to the village in w]

Now by Grice's (1975) Co-operative Principle (specifically the Maxim of Quantity), if the speaker of (61) knew that it was Gunapala who went to the village, or Chitra who went to the village, or that both Gunapala and Chitra went to the village, he would say so, rather than uttering (61).

Thus (61) carries with it a conversational implicature that, in some of the speaker's epistemically-accessible worlds, Gunapala went to the village, and in others, Chitra went to the village (allowing for worlds in which both are true, but disallowing worlds where neither is true). Given that such implicatures will consistently obtain in non-interrogative disjunctive contexts, hearers could re-analyse this pragmatical signal as being associated specifically with the lexeme *hari*(/*hō*), and thus it would become a presupposition carried by that lexeme. Since the presupposition is associated with *hari*(/*hō*), a Q-particle, it makes sense that it would be treated in terms of choice functions, as in the definition in (60).

When the domain of *hari*/*hō* is extended to include the formation of *wh*-based indefinites, the presupposition is carried into the new environment and thus are born the "extensionally-unknown" indefinites. In other words, the pragmatics associated with "extensionally-unknown" WH+*hari* indefinites ultimately derive from the particle's original function in the formation of disjunction.

In the case of *də*, recall that it carries a presupposition that there is at least one epistemically-accessible world w′ for which there is no Important Predicate approximating basic intensional choice function f such that the extension of the individual concept f selects satisfies the speaker's existential claim; the formal definition is repeated in (63).

(63) *də*$_i$(g)(w)(F)(G)(Q)(S), where g is an assignment function, w is a world, F is a modal base supplied by context, G is a set of epistemically-accessible Important Predicates, Q, S are predicates

a. denotes: g(i)(w).BASICH(g(i))

b. "signals": $\exists w' \in F \neg\left[\exists$f.BASICH$_{\sim \text{ImpPred}}$(f): S(w′)(f(Q)(w′))=1$]\right]$

The most natural point of origin for the "intensionally-unknown" indefinite is the use of WH+*də* in *wh*-interrogatives (a reanalysis possibly also motivated by morphological parallelism to the earlier development of WH+*hari* indefinites

[35]Where F is, as above, the set of epistemically-accessible possible worlds.

discussed above). Consider a basic *wh*-question like (64).

(64) īye kau **də** ohē.ta hambavunā
yesterday who ***də*** you.DAT meet.PAST.A
"Who did you meet yesterday?"

According to Gricean maxims of conversation, a cooperative speaker would utter a sentence like (64) in the context that he actually lacks knowledge of the identity of the person that the addressee met. Treating the denotation of *kau* 'who' as being the set of individuals who are human, e.g. $\{x \in D_e \mid x \in \text{human}'\}$, it is crucial to note that the asker of a question like (64) does not necessarily know the full range of possible values of x such that $x \in \text{human}'$.

In other words, in most contexts, the asker of a *wh*-question like (64) expects that a true answer to his question could involve an individual who is unknown to him, which allows for epistemically-accessible worlds for which the speaker knows of no individual concept x such that extension of x in that world would satisfy the proposition that the addressee met that person.

Assuming that the prototypical *wh*-question carried this sort of conversational implicature, it is unsurprising that, here too, as in the case of the development of WH+*hari* indefinites, when *də* is extended to *wh*-indefinites, the conversational implicature associated with the interrogative construction was reanalysed as a presupposition associated with the lexical item *də* itself.[36]

11.2.5 Summary

The Sinhala Q-particle *da*/*də* derives ultimately from an Indo-Aryan form cognate with Sanskrit *utā́ho*, Pāli *udāhu*, which are specialised elements used for forming alternative questions. The predecessor of Sinhala *da* undoubtedly fulfilled the same function, but from the earliest texts we find that *da* also appears in yes/no and *wh*-question—though not consistently. The use of *də* to form *wh*-based indefinite pronouns is a recent development, limited to modern colloquial Sinhala. The Q-particle which appears in modern colloquial Sinhala as *hari* seems to be a replacement of earlier Sinhala *hō*, though the nature of the exact etymological relation between the two forms is uncertain (as is the etymology of *hari* itself). *Hō* was originally confined to the formation of non-interrogative disjunctions, but in both modern literary and modern colloquial Sinhala it can also combine with *wh*-pronouns to form indefinites. These developments are summarised in Table 11.7.[37]

Da/*də* seems to have been slower to be extended to the formation of indefinites due to the fact that this involved a reanalysis of its formal syntactic features.

In summary, I suggest the following path of development for the Sinhala Q-particles:

In the precursor to Sinhala, the ancestor of *da* (corresponding to Sanskrit *utā́ho*, Pāli *udāhu*) occurs only in the formation of alternative questions; whether or not it is formally a Q-particle (in the sense of denoting a variable over choice functions) at this stage is unclear. By the Old Sinhala period, *da* is extended (in some grammars) to the formation of yes/no questions; this is a natural enough extension, if we treat yes/no questions as a special subtype of alternative questions. In this period we also find the Q-particle *hō*, used exclusively for the formation of declarative disjunctions. Later in the early Sinhala period, in some grammars, *da* becomes generalised as a particle appropriate to interrogatives, and thus is extended to the formation of *wh*-questions; this extension requires a reanalysis of *wh*-pronouns as semantically-simple, Hamblin-type set denoting elements.

[36] In both the development of *hari* and *də*, the conversational implicature may have been reanalysed as a generalised conventional implicature (associated with the lexical item) before becoming a presupposition. See further Geis & Zwicky (1971), Cole (1975), Traugott & Dasher (2005).

[37] Pre-Sinh = hypothetical Pre-Sinhala (stages 1 and 2); Old Sinh = Old Sinhala; Class Sin = Classical Sinhala; M Lit Sinh = Modern Literary Sinhala; M Collq Sinh = Modern Colloquial Sinhala. Round brackets () indicate variability/optionality in the use of the Q-particle in that context.

	Pre-Sinh 1	Pre-Sinh 2	Old Sinh	Class Sin	M Lit Sinh	M Collq Sinh
Da/Də						
alt-ques.	✓	✓	✓	✓	✓	✓
y/n-ques.		(✓)	(✓)	(✓)	✓	✓
wh-ques.			(✓)	(✓)	✓	✓
indef.						✓
Monomorphemic indefs. (*yam*, *kisi*)	✓	✓	✓	✓	✓	
Hō/Hari						
indef.					✓	✓
decl. disj.	✓	✓	✓	✓	✓	✓

Table 11.7: Spread of particles *da*/*də*, *hō*/*hari* in Sinhala

By the time of modern literary Sinhala, the earlier grammars in which *da* is restricted to alternative questions is gone, leaving only grammars in which *da* occurs in all types of interrogatives. This entails that the semantic variation between semantically-simple and semantically-complex *wh*-words is also gone, leaving only semantically-simple, Hamblin-type set denoting *wh*-words. This makes more likely the extension of *hō* to the formation of *wh*-based indefinites—a change which was already likely given that such an extension requires no change in the formal syntactic features of *hō*.

In the modern colloquial period, *da*/*də* is extended to the formation of *wh*-based indefinites. The extension of *də* to indefinites presumably takes place later than the extension of *hō*/*hari* due to the fact that the former does require change in the formal syntactic features of *də*, while the latter does not.

The presuppositions of *hari* and *də*, evidenced by the distributions of WH+*hari* and WH+*də* indefinites, are likely to represent the reanalysis of conversational implicatures; the presuppositions of the two particles differ due to the difference in their earlier environments: the presupposition of *hari* derives from *hari*'s earlier use in non-interrogative disjunctions, while the presupposition of *də* derives from *də*'s earlier use in *wh*-interrogatives.

As we found in the examination of the development of Sinhala focus constructions in Chapter 11.1, the development of Q-particles in Sinhala reveals both apparent convergence with as well as divergence from Dravidian. The use of Q-particles in *wh*-questions,[38] (declarative) disjunctions, and (in earlier Sinhala) in relative clauses appears likely to reflect convergence with Dravidian. However, the use of three distinct Q-particles (*da*/*də*, *hō*/*hari*, *vat*), with different syntactic distributions and different pragmatics, finds no parallel in Dravidian. So here too one must conclude that though Dravidian influence at certain periods triggered some convergent developments in Sinhala, much of the development of Q-particles in Sinhala appears to be the result of language internal changes.

11.3 The importance of historical data for synchronic analysis

The analysis of the historical development and spread of the Q-particles *da*/*də* and *hō*/*hari* crucially depends on the synchronic treatment of *da*/*də* as uniformly denoting a variable over choice functions in ALL of the contexts in which it appears, rather than treating the *da* which appears in yes/no- and alternative-questions as a separate

[38]The use of some sort of "question-marker" (potentially a Q-particle) in Indo-Aryan in YES/NO interrogatives is not uncommon, examples include the (optional) use of *kim* (formally the neuter nominative/accusative singular interrogative pronoun "what") in yes/no questions in Sanskrit, either preceding or following the interrogative clause (see above example (38)); and Hindi *kyā* (also literally "what"), which also may optionally appear in yes/no interrogatives, either preceding or following the clause—see above footnote 19. However, such "question-markers" are not used in *wh*-interrogatives in either language.

element from the *da* which appears in *wh*-questions (as per Cable 2007). If the *da* of yes/no questions were analysed as distinct from the *da* of *wh*-interrogatives, there could be no coherent account of the diachronic development of the Sinhala particles discussed here. As shown above in Chapter 9, if we adopt the idea that Q-particles (in both yes/no and *wh*-environments) have a uniform denotation (as variables over choice functions) then the diachronic data can be accounted for in a straightforward manner by positing minimal changes in formal syntactic features. Here we observe the importance of applying both formal synchronic analysis and philological analysis to linguistic data, as it allows for a delimiting of the sets of possible accounts.

Not only is the historical account informed by formal synchronic analysis, but the argument cuts the other way as well: the synchronic analysis is crucially informed by the historical analysis. Consider Cable's (2007) suggestion that the particles we find in yes/no and alternative questions, and in disjunctions more generally are fundamentally different from the particles we find in *wh*-interrogatives and *wh*-indefinites—despite the fact that these are form-identical in Sinhala, Japanese, Malayalam and presumably many other languages as well.

There are two possible lines of argumentation that a Cable-type analysis might pursue. The first is that this "homophony" of particles is entirely accidental. This seems extremely unlikely given the crosslinguistic occurrence of form-identity of these two purported sets of particles. The alternative approach would be to suggest that there is a diachronic relationship between the two sets. This second line of approach fares better with respect to the crosslinguistic data, as the "homophony" then need not be accidental. Let us refer to these two hypothetical sets of particles as type A particles, referring to those used in alternative and yes/no questions and in disjunctions, and type B particles, referring particle used only in *wh*-environments; see Table 11.8. In Sinhala, Malayalam, and Japanese the particles in set A and set B are homophonous, though in Tlingit set A particles are realised as *gé* and set B particles as *sá*. However, whether or not they are form-identical, the semantic denotations of type A and type B particles would be non-identical on a Cable-type account.[39]

Hypothetic Type Label	Environment	Sinhala	Malayalam	Japanese	Tlingit
Type A					
	Alt/Y-N Ques.	də	-oo	ka no kai kadooka	gé
	Decl. Disj.	hari	-oo	ka	khach'u
Type B					
	***Wh*-Ques.**	də	–	ka no ndai	sá
	***Wh*-Indef.**	də hari	-oo	ka	sá

Table 11.8: Hypothetic particle types and examples

The Sinhala data examined in this chapter demonstrates that type A particles can be extended into type B contexts (i.e. the particle *də* was originally restricted to alternative questions [A] and was later extended to *wh*-interrogatives [B] and then to *wh*-indefinites [B]). Given a Cable-type approach the motivation for this extension is opaque. Since the denotations of type A and type B particles are non-identical, then extension of type A particles to type B environments would require semantic re-analysis. But if the two types bear different semantic denotations, then there is no apparent reason for this type of extension to occur in the first place. In other words, assuming this

[39]Presumably, if we were to assume a Tlingit-centric view of universal grammar, declarative disjunctions should really constitute a third type of environment [C], as evidenced by the use of a particle with yet another phonological form (*khach'u*) in such environments in Tlingit.

distinction between type A and type B particles, the extension of type A particles in type B environments constitutes an unmotivated historical change.

On the account proposed in this study, under which particles like Sinhala *də* uniformly denote variables over choice-functions, the motivation for the extension is clear: both type A and type B environments involve elements denoting Hamblin-type sets, and thus constitute the right sort of environment for choice-functions. Changes between stages involve simply changes in (i) the formal syntactic feature specifications (which can restrict particular particles to particular syntactic contexts) and, in some cases, (ii) changes in the lexico-semantics of *wh*-words (from semantically-complex to semantically-simple, or vice-versa, as discussed in Chapter 9.2).

Additionally, Cable's analysis suggests a special connection between *wh*-indefinites and *wh*-interrogatives (both type B contexts). In addition to the fact that—as we have seen in the case of modern Malayalam (Chapter 9.3.6)—some languages may employ Q-particles in *wh*-indefinites but not in *wh*-interrogatives, the examination of the historical data tells against a special relationship between particles in *wh*-interrogative and *wh*-indefinite environments. As we have seen in Section 11.2.5, in Sinhala the particle *hō*/*hari* was extended to the formation of *wh*-indefinites prior to the extension of *də* to *wh*-indefinites. Assuming a division of particles into type A and type B would predict that type B particles, since they are appropriate to *wh*-environments, should be more likely than type A particles to be extended to *wh*-indefinites. As above, the motivation for the extension of type A particles to type B contexts is unclear in any event, but certainly such an extension should FOLLOW the extension of type B particles from one type B environment (*wh*-interrogative) to another (*wh*-indefinite). But the diachronic data from Sinhala reveal exactly the opposite pattern. It is *hō*/*hari*, a type A particle, originally used in non-interrogative disjunctions, which is the first to be extended to the formation of *wh*-indefinites, which would constitute a type B context. Only later is *də*, which occurs in *wh*-interrogatives (type B) extended to form *wh*-indefinites (type B).

Again, under the analysis proposed here, the fact that *hō*/*hari* extends to the formation of *wh*-indefinites prior to the extension of *də* to *wh*-indefinites can be naturally accounted for. *Də* is originally restricted, by dint of its formal feature specification, to interrogative environments. *Hari* has no such restriction. Thus the extension of *hari* to *wh*-indefinites involves no semantic OR syntactic change, whereas the extension of *də* involves the loss of the uninterpretable uInt[ ] syntactic feature from *də*.

The analysis proposed here, in which particles like Sinhala *də* and Japanese *ka* bear a uniform denotation across all syntactic environments is to be preferred to the analysis proposed by Cable (2007), which treats particles in type A environments as distinct from particle in type B environments, since the former allows for a natural account of the observed historical developments in Sinhala. This combined synchronic-formal/historical-philological approach to linguistic data of course can be usefully extended to other linguistic phenomena.

More generally, often we may find many competing potential formal synchronic analyses of a particular stage of a language. Even the use of crosslinguistic data only rules out a certain number of competing analyses, since the languages considered are generally unrelated—and even where related languages are considered, the relationship between them is that of sisters (or cousins etc.) rather than mother-daughter—and so too their grammars bear no direct relation to each other. When different stages of the same language are considered, then the question arises of how a language could transition from grammar I to grammar II. Competing synchronic accounts of these grammars entail different competing accounts of historical transitions. These competing accounts of historical transitions can be evaluated in terms of whether the changes between grammar I and grammar II can be plausibly motivated or not. Synchronic analyses of particular grammars which allow for motivated accounts of historical change are to be preferred over those which do not. Thus the use of philology should crucially inform synchronic formal analysis.

Chapter 12

Conclusion

> '. . . And so the ideas, which I was using earlier to imagine a horse I had not yet seen, were pure signs, as the hoofprints in the snow were signs of the idea of "horse"; and signs and the signs of signs are used only when we are lacking things.'
>
> —William of Baskerville to his pupil Adso of Melk,
> in Umberto Eco's *The name of the rose*

In this study I have examined the properties of Q-particles and focus in Sinhala and other languages, from both formal synchronic and diachronic perspectives. Crosslinguistically, Q-particles tend to occur in a set of environments not limited only to questions (*wh*-, yes/no-, and alternative), but also in the formation of certain indefinite pronouns and in non-interrogative disjunctions. I propose a unified semantic analysis for Q-particles which accounts for their appearance in just this set of environments. It then remains to account for language-specific differences in the distribution of Q-particles, which are handled largely in terms of language-particular differences in formal syntactic features.

The hypothesis that Q-particles can be assigned a single denotation which, crosslinguistically, accounts for their distribution, has certain consequences for our understanding of the grammar of human language more generally. For example, indefinites formed with Q-particles tend to be epistemic indefinites—indefinites which explicitly signal a lack of knowledge concerning who or what satisfies an existential claim, which suggests that Q-particles may carry presuppositions. The nature of these presuppositions is important in determining the possible distributions of Q-particles more generally.

Sinhala was chosen in part because it presents an ideal opportunity to observe the diachronic development and evolution of Q-particles and focus constructions, due to the language's long literary tradition. In addition to examining the sources of Q-particles in Sanskrit and Pāli (languages closely related to the ancestor of Sinhala), I also investigated Q-particles and focus constructions in Old Sinhala (8^{th}–10^{th} c. A.D.) and Classical Sinhala (12^{th}–15^{th} c. A.D.).

The decision to examine focus alongside of Q-particles was non-arbitrary. There are both syntactic and semantic connections between focus and the formation of questions in Sinhala. On the syntactic side, both focus and *wh*- and alternative-questions (as well as focussed yes/no questions) require the presence of the *-e* form of the verb. Focussed constituents, including interrogative pronouns, (optionally in colloquial Sinhala, obligatorily in literary Sinhala) appear to the right of the *-e* marked verb. Semantically, assuming a Hamblin-style analysis of *wh*-words, and a Roothian analysis of focus, both involve evaluation over special sets of "alternatives". Alternative and yes/no-questions, as well as disjunctions more generally, also involve Hamblin-type sets in the analysis proposed here.

Below I summarise the major findings and claims of this study. In Section 12.1, I review the findings concerning the synchronic analysis of focus constructions in Sinhala; Section 12.2 reviews the semantic analysis of Q-particles as

bearing a uniform denotation as variables over choice functions, and how this can account for the crosslinguistic patterns observed. Section 12.3 reviews the novel analysis proposed for disjunction, motivated by the overall semantic analysis of Q-particles. In Section 12.4 I outline the account of crosslinguistic differences in Q-particle distribution, which relies largely on language-specific differences in formal syntactic feature assignments to lexical elements. Section 12.5 reviews the analysis of epistemic indefinites formed with Q-particles, and the role played by pragmatics in constraining the distribution of Q-particles; Section 12.6 rehearses the basic diachronic findings concerning Sinhala focus and Q-particle constructions, and the relation of these data for the formal synchronic analysis. Section 12.7 provides a brief discussion of the greater implications of this study, and suggestions for directions for future research.

12.1 Focus in Sinhala

Both modern literary and modern colloquial Sinhala employ special marking on the matrix verb of a clause containing a focussed element, see example (1a).[1] The same marking also appears on the verb of the (highest) clause over which a *wh*-interrogative pronoun takes scope, as in example (1b); it likewise appears in alternative questions (1c) and in "focussed" yes/no-questions (1d).

(1) a. eyaa gaməṭə^{F} (y/tamay) yanne
he.NOM village^{F}-DAT (EMPH) go.PRES.E
'It is to the village he goes.'

b. Ranjit [[[kau də] aawa] kiyəla] danne?
Ranjit [[[who *də*] came.PAST.A] that] knowPRES.E
'Who does Ranjit know came?'

c. Gunəpālə **də** Chitra **də** gaməṭə giyē?
Gunapala ***də*** Chitra **də** village.DAT go.PAST.E
'Was it Gunapala or Chitra who went to the village?'

d. Chitra ee potə də kieuwe?
Chitra that book *də* read.PAST.E
'Was it that book which Chitra read?'

The appearance of the focussing *-e* verbal marking may optionally be accompanied by movement of the focussed element to the position following the *-e* marked verb, an operation sometimes referred to in the literature as "clefting", see example (2) below.

(2) eyaa yanne gaməṭə^{F} (y/tamay)
he go.PRES.E village^{F}-DAT (EMPH)
'It is to the village he goes.'

Such structures bear a surface similarity to certain constructions in Dravidian languages, as in the (Sri Lankan) Tamil example in (3).

(3) naan poonatu yaaLppaaNattukku^{F}
I.NOM went.PAST.VN.VOM Jaffna.DAT^{F}
'It was to Jaffna that I went.' (cited from Gair 1986[1998]b: 156)

[1]Here I concentrate the discussion on the colloquial language; see Chapters 4 and 11.1 for further discussion of the focus construction in literary Sinhala.

However, though the *-e* forms in Sinhala historically derive from nominalised verbs, they are no longer treated as such in modern Sinhala, in contrast to the case of the Dravidian formations (see further below, Section 12.6.1).

The Sinhala focus construction does exhibit certain properties which distinguish it from, say, focus in English. In addition to the special *-e* marking on the verb,[2] such structures involve the presupposition of the existence of at least one element of the same semantic type as the focussed element which satisfies the proposition—a property found in English cleft constructions of the type "It was John who went to the village" (which presupposes that someone went to the village). Additionally, only one element per clause may be "focussed", in the sense that only one element may bear a focus marker (like *tamaa/tamay/-y(i)*) and/or undergo displacement to focus position on the right of the verb—similar to the restriction on the emphasiser *hī* in Hindi—though additional elements may be semantically and prosodically focussed. I argue that the existence presupposition is borne by the *-e* element itself, which is generated in the head of FocusP. See above Chapters 4 and 5 for a more detailed discussion.

12.2 Accounting for the crosslinguistic similarities of Q-particles and their environments

In a number of languages, including Sinhala, Japanese, Malayalam, and Tlingit (see Table 12.1), we find Q-particles not only in *wh*-, alternative-, and yes/no-questions, but also in non-interrogative disjunctions and in the formation of *wh*-based indefinite pronouns. The appearance of Q-particles in this set of syntactically-heterogeneous environments is accounted for by positing an underlyingly unified semantics for Q-particles.[3]

	Old Sin	Class Sin	Lit Sin	Colloq Sin	Old Mal	Mod Mal	Tlin	Jap
y/n-ques.	(da)	(da)	da	də	-oo	-oo	gé	ka
wh-ques.	(da)	(da)	da	də	-oo	—	sá	ka
wh-indef.	—	—	hō (aff. & neg.), vat (neg.)	də (aff.), hari (aff.), vat(neg.)	-oo	-oo	sá	ka
decl. disj.	hō, heva(-t)	hō, heva(-t)	hō (aff. & neg.), vat (neg.)	hari (aff. & neg.), vat (neg.),	-oo	-oo	khach'u	ka
interr. disj.	da	da	da	də	-oo	-oo	gé... gwáa	[ka]

Table 12.1: Distribution of Q particles in various stages of Sinhala & Malayalam; Tlingit, Japanese (repeated)

Specifically, this study proposes an analysis of Q-particles as denoting variables over choice functions (see Chapter 6)—where a choice function is a function which when applied to a (non-empty) set, returns a single member of that set. I further assume that Q-particles are restricted in that they may only apply to Hamblin-type sets (formally notated as $\langle \alpha // \beta \rangle$ to distinguish them from other sets of the same semantic type, i.e. $\langle \alpha, \beta \rangle$) as discussed in Chapter 9.

Wh-words, including both interrogative and indefinite pronouns, may be treated, as in Hamblin's original formulation, as Hamblin-type sets, e.g. Sinhala $[\![\textit{kau}\text{ "who"}]\!]^g = \{x \in D_e \mid x \in \text{human}'\}$. The existence of a connection between focus and interrogatives, as noted above, led Beck (2006) (followed by Cable (2007)) to propose that the two concepts can be unified by analysing *wh*-words as bearing only focus semantic values, and that this move can provide an explanation of intervention effects which occur if a focus-sensitive operator attempts to apply to a *wh*-word before the Q-particle does. Q-particles, in this analysis, are special in that they make reference only to their complement's focus semantic value—and thus can "convert" a *wh*-word back into a semantically ordinary element. Focus-sensitive operators, on the other hand, make reference to both the ordinary and focus semantic values of their complements,

[2]Note that in literary Sinhala, clauses where the verb takes the *-e* marking differ from ordinary clauses in that there is no subject-verb agreement marking on the verb; in colloquial Sinhala subject-verb agreement marking on verbs is lacking in all cases.

[3]In Table 12.1 round brackets () around Q-particle forms indicate optionality, square brackets [] indicate some further complications.

thus triggering a semantic crash if they attempt to apply to a *wh*-word, which is posited to bear only a focus semantic value.

As intuitively attractive as this analysis is, I have shown in Chapter 6.3 that it suffers from a number of fatal problems. Most obviously, positing that *wh*-words bear only focus semantic values incorrectly predicts that all *wh*-words should bear the SAME denotation—since focus semantic values do not include any semantic restriction aside from a requirement that they be of the same semantic type as the element's ordinary semantic value. Thus both *who* and *what* are predicted to denote $\{x \in D_e\}$. Additionally, *wh*-words do in fact show a distinction between ordinary and semantically-focussed meanings, as evidenced by examples such as (4).

(4) I didn't ask WHAT John saw (, I asked WHO he saw).

Thus, despite the similarity of Hamblin-type denotations (as borne by *wh*-words) and focus semantic values, the two cannot be reduced to focus semantics as proposed by Beck (2006). Rather, the ordinary semantic value of a *wh*-word is a Hamblin-type set.

In questions, choice function variables are bound via the denotation of the interrogative COMP head of CP:

(5) $[\![[\text{COMP}^{\text{INT}}_i \text{ XP}]]\!]^g = \lambda p\left[\exists f \in D_{cf}.p = [\![\text{XP}']\!]^{g^{[f/i]}}\right]$

Thus a question like that in (6) will have the denotation shown in (7).[4]

(6) Kau də aawe?
who *də* come.PAST.E
"Who came?"

(7) $[\![(6)]\!]^g$ =

a. $[\![\text{CP}_{\text{INT}}$ [kau də aawe]$]\!]^g$ =
b. $\lambda p[\exists f \in D_{cf}.p = [\![[\text{kau də aawe}]]\!]^{g^{[f/i]}}$ =
c. $\lambda p[\exists f \in D_{cf}.p = [\![\lambda w.\lambda x.x \text{ came}(g(i)(\{x \in D_e \mid x \in \text{human}'\}))]\!]^{g^{[f/i]}}$ in w] =
d. $\lambda p[\exists f \in D_{cf}.p = \lambda w.\lambda x.x \text{ came}(f(\{x \in D_e \mid x \in \text{human}'\}))$ in w] =
e. $\lambda p[\exists f \in D_{cf}.p = \lambda w.f(\{x \in D_e \mid x \in \text{human}'\})$ came in w]

In non-interrogative contexts, I assume that choice functions are bound by existential closure, so that a statement like (8) will have the denotation shown in (9).

(8) Kau də aawa.
who *də* come.PAST.A
"Someone came."

(9) $[\![(8)]\!]^g = \exists f \in D_{cf}.\lambda w.f(\{x \in D_e \mid x \in \text{human}'\})$ came in w]

Treating *wh*-words—including both interrogative and indefinite pronouns in languages like Sinhala and Japanese—as Hamblin-type sets thus allows for a natural analysis of the Q-particles which appear in these constructions as denoting variables over choice functions.

Disjunctive contexts—which I take to include both alternative and yes/no-questions, as well as non-interrogative disjunctions—like *wh*-words also intuitively involve alternatives, which I propose are formally representable as Hamblin-type sets. The formalisation of this intuition, however, requires a new analysis of both the syntax and

[4] Setting aside the existence presupposition triggered by the *-e* marking on the verb.

semantics of disjunction, as discussed in the following section.

12.3 A new analysis of disjunction

One important consequence of a semantically-unified treatment of Q-particles is that this analysis motivates a new syntactic and semantic analysis of disjunction, building on certain aspects of Alonso-Ovalle (2006) and den Dikken (2006).

Alonso-Ovalle (2006) provides a Hamblin-type analysis of disjunction, which he suggests is superior to the standard semantic treatment of disjunction ($\vee$) for handling (i) counterfactual conditionals involving disjunctive antecedents, (ii) disjunctions under the scope of modals, and (iii) unembedded disjunctions with an exclusion component. However, under Alonso-Ovalle's analysis it is the (overt) disjunction operators themselves (e.g. English *or*) which are responsible for the formation of Hamblin-type sets. This analysis is unavailable if we are to treat Q-particles in all contexts as denoting variables over choice functions.

Based on syntactic evidence, den Dikken (2006) proposes that apparent lexicalisations of disjunction like English *or* are in fact phrasal categories which adjoin directly to their disjunct, and that disjunction itself is in fact accomplished by an unpronounced head J ("junction").[5]

Adopting den Dikken's analysis of disjunction allows for the work of creating Hamblin-type sets to be done by J, which in turn allows us to maintain the semantic uniformity of Q-particles as variables over choice functions. That is, the unpronounced J head creates Hamblin-type sets, to which choice functions (lexicalised as Q-particles) may apply.

As discussed in Chapter 10, this analysis also explains the fact that in many languages a Q-particle accompanies each disjunct either obligatorily (as in Sinhala) or optionally (as in Japanese). Every J creates a set which contains two elements: (i) the higher disjunct, and (ii) a set containing the lower disjunct: this means that each J in a sense creates two layers of "set-recursion". Since the role of a choice function is to convert a set into a member from that set, each disjunct must be associated with a Q-particle. Thus a sentence like (10) will have the denotation shown in (11).

(10) gunəpaalə də chitra də ranjit də gamə̣tə giye?
Gunapala *də* Chitra *də* Ranjit *də* village.DAT went-E
"Was it Gunapala or Chitra or Ranjit who went to the village?"

(11) $[\![(10)]\!]^g = \lambda p[\exists f_1,f_2,f_3 \in D_{cf}.p=\lambda w.f_1(\{\text{Gunapala}, f_2(\{\text{Ranjit}, f_3(\{\text{Chitra}\})\})\})$ went to the village in w]

Yes/no-questions are treated as a special subtype of alternative questions, involving an elided . . . *or not*, and thus receive similar interpretations.

Adopting such an analysis—where disjunction involves (i) the creation of a Hamblin-type set by J, and (ii) the application of choice function(s) to the output of J—thus allows us to maintain a semantically unified denotation for Q-particles.

In this way, the range of environments in which Q-particles appear crosslinguistically can be given a unified analysis by way of the semantics of Q-particles. That is, all of the environments in which Q-particles appear contain an element with a Hamblin-type denotation: either a *wh*-word or a Hamblin-type set created by J.

Thus semantics provides the unifying aspect of the analysis, allowing us to capture important generalisations about the appearance of Q-particles crosslinguistically. However, there are also important language-specific differences in the distribution of Q-particles, as reiterated in the following section.

[5]The syntactic evidence involves the possible positions of *either*, which can occur both lower ([John **either** ate rice] or [he ate beans]) or higher (John **either** ate [rice] or [beans]) than the edge of the leftmost disjunct.

12.4 Syntactic features and language-specific restriction on Q-particles

As shown by Table 12.1, repeated below as Table 12.2, though there are many similarities between the languages examined in this study in terms of which environment Q-particles appear in, there are important differences as well. Additionally, Sinhala and Tlingit (in contrast to Japanese and Malayalam) employ Q-particles with different morphological forms in different syntactic environments.

	Old Sin	Class Sin	Lit Sin	Colloq Sin	Old Mal	Mod Mal	Tlin	Jap
y/n-ques.	(da)	(da)	da	də	-oo	-oo	gé	ka
wh-ques.	(da)	(da)	da	də	-oo	—	sá	ka
wh-indef.	—	—	hō (aff. & neg.), vat (neg.)	də (aff.), hari (aff.), vat(neg.)	-oo	-oo	sá	ka
decl. disj.	hō, heva(-t)	hō, heva(-t)	hō (aff. & neg.), vat (neg.)	də (aff.), hari (aff. & neg.), vat (neg.)	-oo	-oo	khach'u	ka
interr. disj.	da	da	da	də	-oo	-oo	gé... gwáa	[ka]

Table 12.2: Distribution of Q particles in various stages of Sinhala & Malayalam; Tlingit, Japanese (repeated)

The semantic analysis of Q-particles as denoting variables over choice functions successfully captures the recurrent crosslinguistic pattern of Q-particles appearing in *wh*-interrogatives, yes/no and alternative interrogatives, *wh*-indefinites, and non-interrogative disjunctions.

The explanation of which Q-particle appears in which environments in a particular language largely relies on the language-specific differences in formal syntactic features. Determining the distribution of Q-particles in a particular language depends then on the particular syntactic features borne by Q-particles and other elements such as the interrogative COMP head. For example, in modern colloquial Sinhala, the interrogative COMP head has the initial feature assignment [Q[ ], Int[+]], which means that it requires an element with a valued instance of the Q[ ] feature within its c-command domain. The Q-particle *də* has the feature assignment [Q[+]]. The Q-particle *hari* bears no relevant features, and thus cannot value the interrogative COMP's Q[ ] feature. This entails that all questions require the presence of *də*. In modern literary Sinhala, the interrogative COMP head again bears the initial feature assignment [Q[ ], Int[+]], but here *də* bears the assignment [Q[+],Int[ ]]. This accords with the bi-conditional nature of the relationship of interrogatives and *də* in the literary language: *də* must be present in all and only in questions.

The distribution of Q-particles in most of the remaining cases can be handled similarly (the complexity of the Tlingit pattern requires another feature, but the mechanisms are the same). The major exception is the inability of modern colloquial Sinhala *də* to appear in non-interrogative disjunctions.[6] In terms of the system of formal syntactic features described above, there is no means of restricting *də* only to interrogative disjunctions while at the same time allowing it to occur in the formation of indefinite pronouns.

In fact, *də* can be ruled out from non-interrogative disjunctions on pragmatic grounds. In the next section, I turn to the review of indefinite pronouns formed from *wh*-words+Q-particles, which appear to have a crosslinguistic tendency to form epistemic indefinites. Such indefinites bear presuppositions relating to the knowledge of the speaker[7], and the presupposition borne by *də* is incompatible with the assertion of a non-interrogative disjunction.

[6]The other exception relates to the lack of Q-particles in certain environments, such as in modern Malayalam *wh*-questions: this depends on whether an element (interrogative or indefinite pronoun) is semantically complex or semantically simple (see Chapter 9 for more details).

[7]In most cases the "ignorance component" of such indefinite relates to the speaker's knowledge; but it may relate to the addressee's knowledge, or the knowledge of some other person: see above Chapter 7.

12.5 Epistemic indefinites and pragmatic restrictions on Q-particles

As discussed in Chapter 7, in modern colloquial Sinhala we find, alongside of "plain" indefinites like that in (12a), two *wh*-based indefinites, both involving Q-particles: see (12b) and (12c).[8]

(12) a. sanat deyak gatta.
Sanath thing.INDEF buy.PAST.A
"Sanath bought a thing."

b. sanat monəva hari gatta.
Sanath what *hari* buy.PAST.A
"Sanath bought some thing."

c. sanat monəva də gatta.
Sanath what *də* buy.PAST.A
"Sanath bought some thing-or-other."

The sentences in (12a)–(12c) all have the same truth-conditional semantics, but they differ in terms of the pragmatic conditions in which they are felicitous. The two *wh*-based indefinites (12b) and (12c) both pragmatically signal that the identity of the referent of the indefinite is unknown. The precise pragmatic conditions under which (12b) and (12c) are felicitous differ, with the latter presupposing a greater degree of ignorance than (12b).

Considering the scenarios described below in (13), we find that the sentences above in (12) are felicitous in the environments shown in Table 12.3.

(13) a. I saw Sanath buy Jean-Baptiste Greuze's painting *The White Hat*.

b. I saw Sanath buy some piece of artwork, but I don't really know what it was. (I.e., I might be able to describe it, but I don't know what it's called, who painted it, etc.)

c. Sanath told me that he bought something, but I have no direct experience of the event.

	(12a) is felicitous	**(12b) is felicitous**	**(12c) is felicitous**
In scenario (13a)	✓		
In scenario (13b)	✓	✓	
In scenario (13c)	✓	?	✓

Table 12.3: Felicity conditions for indefinite constructions in Sinhala (repeated)

The pragmatics of Sinhala WH+*də* indefinite thus appear to be similar to those described by Alonso-Ovalle & Menéndez-Benito (2003, 2010) for Spanish *algún* NP, as well as those of the English structure *some NP-or-other*, while the pragmatics of Sinhala WH+*hari* are similar to those of English *some NP*. I describe the former type of indefinite as "intensionally-unknown" and the latter as "extensionally-unknown". "Extensionally-unknown" indefinites are felicitous where the speaker has no means of identifying an extension which uniquely satisfies his existential claim, and "intensionally-unknown" indefinites are felicitous where the speaker cannot even identify an individual concept which uniquely satisfies his existential claim.

The presuppositions of these two types of epistemic indefinites (which I suggest are responsible for the indefinites' "ignorance components") can be formalised in terms of basic intensional choice functions which approximate Important Predicates.

[8] In Old and Classical Sinhala we find only monomorphemic indefinites, *kisi*, *yam*, which are not synchronically related to *wh*-pronouns. In modern literary Sinhala, only *hō* (≈ *hari*) can form *wh*-indefinites. See further Chapter 11.2 for more detailed discussion of indefinites in early Sinhala.

In this account *hari*-type indefinites presuppose that there is no Important Predicate approximating basic intensional choice function f, such that the extension of the individual concept chosen by f both satisfies the proposition and does not vary between different epistemically-accessible worlds. In contrast, *də*-type indefinites presuppose that there is at least one epistemically-accessible world w′ for which there exists no Important Predicate approximating basic intensional choice function f such that the extension of the individual concept selected by f satisfies the proposition in w′.

Not only does this analysis correctly account for the distribution of epistemic indefinites like Sinhala WH+*hari* and WH+*də*, but it also accounts for the inability of the Q-particle *də* to appear in non-interrogative disjunctions. Given that the use of a declarative disjunction entails that the speaker asserts that the proposition is true in all epistemically-accessible worlds for at least one of the disjuncts, the inadmissibility of *də* in non-interrogative disjunctions is clear: *də* presupposes that there exists some epistemically-accessible world w′ for which it is the case that there is no individual concept known to the speaker whose extension satisfies the proposition. Thus the presupposition of *də* contradicts the assertion in non-declarative disjunctions, accounting for the inability of *də* to appear in such contexts—a restriction which we could not derive on the basis of formal syntactic features.

This is a welcome result, since the pragmatic analysis of *də* and *hari* is independently motivated by the distribution of indefinites in Sinhala, yet also provides a straightforward explanation for the inability of *də* to appear in non-interrogative disjunctions.

The distribution of Q-particles thus highlights the importance of linguistic analyses which consider various components of the grammar. The distribution of Q-particles cannot be accounted for completely utilising an analysis which relies on a single module of the grammar. Syntax, semantics, pragmatics, and morphology all play crucial roles in determining the distribution of Q-particles.

In the next section, I turn to the consideration of the diachronic changes in the syntax and semantics of focus and the origin and development of Q-particles in Sinhala.

12.6 Diachronic analysis of Sinhala focus and Q-particle constructions

12.6.1 The development of Sinhala focus constructions

In Old Sinhala, the precursor to the later "focus construction" does not appear to be inexorably linked to the presence of a focussed element, though it is compatible with it. At this stage, the construction is monoclausal, involving an "impersonal" use of the participle (nominalised verb) in the masculine/neuter nominative singular case-form (often appearing as *-e*).[9] That these structures are indeed monoclausal is evidenced by the fact that (1) displacement of any focussed elements is not obligatory, and (2) that a form of "to be" or other agreement element is not obligatory and even when it does appear, it always immediately follows the participle (and not the focussed element as in Classical and modern literary Sinhala).

By the stage of Classical Sinhala this construction has undergone reanalysis—quite possibly at least the partial result of Dravidian influence, given that structurally similar constructions appear in Tamil and Malayalam—as a biclausal "cleft" construction. That the structure is biclausal follows from the fact that (1) focussed elements are always displaced, either to the left-edge of the clause or else immediately following the participle, and (2) some copular element (form of "to be" or agreement clitic) obligatorily appears, and further always appears (3) immediately following the focussed element (not following the participle, as in Old Sinhala).

In Classical Sinhala and modern literary Sinhala, logical subjects (if present) appear in the accusative case,[10]

[9] Logical subjects of participles, if present, appear in genitive case.

[10] Except where the subject itself is focussed—see above Chapter 11.1 for more discussion.

which is morphologically identical to the Old Sinhala genitive (the genitive having been morphologically renewed by means of a postposition). In Old and Classical Sinhala the participle can also be used simply as a nominalised verb, but by the period of modern literary Sinhala, the masculine/neuter nominative singular case-form of the participle has become morphologically isolated, as such participles are no longer a general nominalisation strategy, but are rather confined to the verb of the "focus construction".

In modern colloquial Sinhala the focus construction has again been reanalysed as monoclausal—though it is structurally very different from the "impersonal" participial construction of Old Sinhala. In Old, Classical, and modern literary Sinhala, in "neutral" sentences (with no focus) we find subject/verb agreement morphology on the finite verb; in "focussing"/"impersonal" participial constructions, we find no subject/verb agreement morphology on the (participial) verb. However, by modern colloquial Sinhala all subject/verb agreement morphology has been lost. Additionally, though in earlier stages of Sinhala, logical subjects of participial/focussing constructions appear in the (simplex) genitive/accusative case, in modern colloquial Sinhala they appear in the normal case selected for by the verb (usually nominative). Both of these differences from earlier Sinhala further remove the *-e* verb form from any association with nominalisation. Further, and most importantly for the analysis of such structures as monoclausal, in modern colloquial Sinhala: (1) displacement of the focussed element is optional, and (2) no form of "to be" or agreement clitic is required to follow the focussed element.

12.6.2 The origin and development of Sinhala Q-particles

The Sinhala Q-particle *də* derives from Old Indo-Aryan *utā́ho*, which is itself ultimately composed of two particles, *ā́ho* and *utá*, see above Chapter 11.2.[11] In Classical Sanskrit *utāho* is used specifically in alternative questions, see a representative example in (14).

(14) kiṁ mama vadhopāyakramaḥ kubjasya vā_**utāho** anyasya vā kasyacit
Q me-GEN murder-plot.NOM.SG hunchback-GEN or_***utāho*** other-GEN or someone-GEN
'Is it I, against whom the murder-plot is laid, or is it the hunchback or somebody else?' (Panc. 332) (Speijer 1886: §415)

In Pāli, the cognate *udāhu* is employed in the same way, as in example (15).

(15) kiṇ amhehi saddhiṇ āgamissasi **udāhu** pacchā?
Q us with come-FUT.2SG ***udāhu*** later
'Will you come with us or later?' (DhA ii.96)

Pre-Sinhala *da* presumably had a similar distribution; but by the earliest texts we find *da* appearing sometimes also in yes/no and *wh*-questions.[12] Not until modern colloquial Sinhala does it begin to be used in the formation of *wh*-based indefinites.

[11] In earlier Vedic prose we in fact find alternative questions formed using the two particles separately, with *utá* appearing at the left edge of the first disjunct, and *utá* at the left edge of the second disjunct, as in the following example:

(i) **utá** _avidvā́n amúṁ lokáṁ prétya kaścaná gacchatī3 / **ā́ho** vidvā́n
uta one who does not know.NOM.SG yonder.ACC.SG worldACC.SG depart.GER anyone go.PRES.3SG / ***āho*** one who knows.NOM.SG
amúṁ lokáṁ prétya kaścit sámaśnutā3i /
yonder.ACC.SG world.ACC.SG depart.GER anyone reach.PRES.3SG /
'Does anyone who does not know, having died, go to yonder world, or does anyone who knows, having died, attain yonder world?' (Taittirīya Upaniṣad 2.6, cited from Böhtlingk & Roth 1855–1875)

[12] Note also the change from a proclitic-type element in Sanskrit/Pāli to an enclitic element by the time of the earliest attested Sinhala.

Hari is of more obscure origins; it replaces earlier *hō* (derived from a form of "to be"), originally restricted to the formation of non-interrogative disjunctions. *Wh*-based indefinites involving *hō*/*hari* appear earlier than do WH+*də* indefinites, presumably because the extension of the latter involved reanalysis of the formal syntactic features associated with *də* (which originally restricted it to interrogative contexts), while the extension of *hari* involved no such reanalysis. The development of these two particles is summarised in Table 11.7.

	Pre-Sinh 1	Pre-Sinh 2	Old Sinh	Class Sin	M Lit Sinh	M Collq Sinh
Da/Də						
alt-ques.	✓	✓	✓	✓	✓	✓
y/n-ques.		(✓)	(✓)	(✓)	✓	✓
wh-ques.			(✓)	(✓)	✓	✓
indef.						✓
Mono-morphemic indefs. (*yam*, *kisi*)	✓	✓	✓	✓	✓	
Hō/Hari						
indef.					✓	✓
decl. disj.	✓	✓	✓	✓	✓	✓

Table 12.4: Spread of particles *da*/*də*, *hō*/*hari* in Sinhala (repeated)

When *da*/*də* and *hō*/*hari* are extended to the formation of *wh*-based indefinites, we can derive the origin of the epistemic presuppositions which these bear (see above) as follows. The presupposition of "extensionally-unknown" WH+*hari* indefinites is the result of reanalysis of the conversational implicature associated with declarative disjunctions (i.e. speakers who say "X or Y did Z" implicate that they cannot say with certainty that "X did Z" or "Y did Z"), while the presupposition of "intensionally-unknown" WH+*də* indefinites is the result of reanalysis of the conversational implicature associated with *wh*-interrogatives (i.e. speakers who ask "who did Z?" implicate that they have no means of uniquely identifying "the person who did Z").

12.7 Implications and Directions for future research

The distribution of Q-particles in the languages examined herein, to wit, Japanese, Tlingit, Malayalam, and various stages of Sinhala, reveal a complex pattern, which cannot be explained within any one single component of the grammar.

The formal analysis which I have proposed in this thesis—namely that particles like Sinhala *də* and Japanese *ka* are predicted to occur in various syntactic environments (interrogative, disjunctions, indefinites) with the SAME semantic denotation (variable over choice functions)—has important theoretical implications for the syntax and semantics of both disjunction and conjunction more generally. Firstly, it entails that Q-particles are ubiquitous in disjunctions crosslinguistically, with the logical consequence that "conjunction markers" like English *or* should be analysed as Q-particles. Such an analysis finds support in the fact that *or* had a wider distribution in earlier English, occurring also as a "marker" of yes/no-questions. Additionally, the analysis of disjunction advanced here accords with the syntactic analysis of den Dikken (2006)—which accounts for the "*either*-too-high" and "*either*-too-low" phenomena—as well as with Alonso-Ovalle's (2006) Hamblin semantics treatment of disjunction—which offers various advantages over the traditional semantic analysis of disjunction.

Secondly, a parallel analysis is suggested for conjunction. Assuming that the role of the "junction" J is simply to

create Hamblin-type sets, then just as disjunction is here analysed as resulting from the application of an existentially-bound choice function variable to the Hamblin-type set created by "junction", so conjunction could be treated as involving the application of a UNIVERSALLY-QUANTIFIED choice function variable to the Hamblin-type set created by J. Such an analysis is also supported by data from languages like Japanese and Sinhala (as well as Gothic and Latin), where particles like Japanese *mo* and Sinhala *-t*, which serve as "conjunction markers", form universal quantifiers when composed with *wh*-words.

The choice functional account of Q-particles adopted in this study also leads to a novel analysis of epistemic indefinites—indefinites which signal a lack of information regarding who or what satisfies the existential claim. In Sinhala we find two morphologically- and pragmatically-distinct epistemic indefinites, both formed from the combination of a *wh*-word with Q-particle. The indefinite formed from WH+*də*, the "more unknown" indefinite, I analyse as an "intensionally-unknown" indefinite, similar in its pragmatics to English *some NP-or-other* and Spanish *algún NP*; the indefinite formed from WH+*hari*, the "less unknown" indefinite, I analyse as an "extensionally-unknown" indefinite, similar to English *some NP* pragmatically.

The presupposition posited for *də*, accounting for its particular epistemic qualities when used in the formation of indefinites, also provides an explanation for why *də* cannot be used to form non-interrogative disjunctions—a fact which could not be predicted by the formal syntactic treatment. This provides another example of the necessity of drawing on various components of the grammar in explaining complex linguistic phenomena.

The study of epistemic indefinites, and their relation to Q-particles, however, bears further investigation. Though I argue that the notions "existentially-unknown" and "intensionally-unknown" indefinites form coherent classes which are attested crosslinguistically, there are important differences between, say, English *some NP-or-other*, Sinhala WH+*də*, and Spanish *algún NP* as regards specificity. Given that there exists some amount of interaction between the properties of specificity and "unknownness", additional research on the properties of epistemic indefinites is desirable, especially in connection to Q-particles.

The historical investigations into the origins of Sinhala focus constructions and Q-particles raise questions about the interaction of language change, language contact, and language acquisition.[13] Both the development of the "cleft" focussing construction and the widespread use of (overt) Q-particles are likely to be—at least in part—the result of convergence with neighbouring Dravidian, which displays similar structural configurations. However, we also observe a number of divergent developments which take place in Sinhala: the biclausal "cleft" construction of Classical and modern literary Sinhala is reanalysed, in modern colloquial Sinhala, as monoclausal; and Sinhala shows a division of labour between three different Q-particles, *də*, *hari*, *vat*, which contrasts with the use of a single Q-particle, *-oo*, in Dravidian Malayalam. These patterns implying periods of convergence and divergence of Sinhala with Dravidian accord well with the results of Gair's examination of various aspects of Sinhala grammar with respect to Dravidian influence (see Gair 1976[1998], 1980, 1985[1998], 1986[1998]b), though as Gair (1985[1998]: 196) notes, it would be useful to examine the social and political history of Sinhala-Tamil interactions to see to what extent the apparent alternation of periods of strong Dravidian influence and periods of weak influence correlate with the political/social/religious history of Sri Lanka, especially that concerning the relations between Sinhala and Tamil speakers.

Some of the changes observed, particularly in the development of the focus construction in modern Sinhala (with respect to the loss of overt subject/verb agreement morphology on the verb, and the shifting case of logical subjects of focussed clauses) look likely to represent the result of reanalysis by child language learners. It would seem fruitful to further consider the relationship between language acquisition and language change in Sinhala. Further, there

[13]The similarities between the *kakari-musubi* construction of Old Japanese (see above Chapter 2.3) and "focussing" sentences/*wh*-interrogatives in modern Sinhala are also intriguing and it would be interesting to see to what extent the developments leading to these constructions are parallel.

appears to be as yet little research even on the question of the acquisition of "question particles" by child language learners. One exception is Okada & Grinstead (2003), who examine the acquisition of sentence-final particles in Japanese. It is interesting to note in their data that the Q-particles *no* and *ka* begin to be produced in child language from about age 2, but apparently only in yes/no-questions and not in *wh*-questions (*wh*-questions appear, but they provide no examples of *wh*-questions produced with Q-particles). Knowledge about child language acquisition of Q-particles could be useful for understanding not only the historical development of such particles, but also for our understanding of their synchronic properties.

The examination of the specific path of development of Q-particles also sheds light on the origin of the epistemic presuppositions we observe in the WH+Q indefinites of Sinhala. The presupposition of *hari* appears to derive from its earlier use in (declarative) disjunctions, while that of *də* from its earlier use in *wh*-questions. This observation emphasises the importance of philological and historical research for the understanding of certain synchronic data.

The distribution of Q-particles examined here also raises questions about the larger typology of *wh*-interrogatives and *wh*-indefinites. In Sinhala, Japanese, Tlingit, and early Malayalam, we find both the use of Q-particles in the formation of *wh*-interrogatives, and the appearance of the (same) Q-particles in the formation of *wh*-indefinites. This suggests that there may be some connection between the use of *wh*-words as indefinites and the use of Q-particles in the formation of *wh*-questions (and the availability of Q-particles more generally). However, Bruening (2007) shows that—through a typological survey and an in-depth analysis of Mandarin Chinese and Passamaquoddy—not only does *wh*-in-situ not correlate with the use of *wh*-indefinites (*pace* Cheng 1991, Cole & Hermon 1998), but neither does the use of (overt) Q-particles correlate with either of these properties. The Malayalam data examined here likewise rule out a correlation between the availability of *wh*-indefinites formed with Q-particles and the use of (overt) Q-particles in *wh*-interrogatives: in modern Malayalam (overt) Q-particles do not appear in the formation of *wh*-interrogatives, but they do appear in the formation of *wh*-indefinites.

The typology of Q-particles itself also is worthy of further investigation. In this study I have concentrated on only four (genetically-unrelated) languages; the close study of additional languages could add further to our knowledge of the properties of Q-particles. Further, a more complete and formalised examination of the role of Q-particles in the formation of relative clauses is needed (see Chapter 8 for some initial observations).

As demonstrated by this study, the evaluation of complex linguistic data such as the distribution of Q-particles crosslinguistically benefits greatly not only by drawing on analyses evaluating the roles played by different components of the grammar (syntax, semantics, pragmatics, morphology) and the interaction of these different modules of the grammar, but also by the consideration of synchronic data in the light of diachronic data obtained through philological investigation (and vice-versa, of course).

Bibliography

Abbott, Barbara. 2000. Attributive, referential, *de dicto* and *de re*. Ms., Michigan State University. http://semanticsarchive.net/Archive/DU3YTgyN/Attributive,%20referential,%20de%20dicto%20and%20de%20re.pdf.

Alonso-Ovalle, Luis. 2006. *Disjunction in alternative semantics*. Amherst, MA: University of Massachusetts, Amherst dissertation.

Alonso-Ovalle, Luis & Paula Menéndez-Benito. 2003. Some epistemic indefinites. In M. Kadowaki & S. Kawahara (eds.), *Proceedings of the North East Linguistic Society, 33*, 1–21. Amherst, MA: GLSA.

Alonso-Ovalle, Luis & Paula Menéndez-Benito. 2010. Modal indefinites. *Natural Language Semantics* 18. 1–31.

Annamalai, E. & Sanford B. Steever. 1998. Modern Tamil. In Sanford B. Steever (ed.), *The Dravidian languages*, 129–157. London: Routledge.

Aoun, Joseph & Yen-hui Audrey Li. 1993. *Wh*-elements in-situ: Syntax or LF? *Linguistic Inquiry* 24. 199–238.

Artstein, Ron. 2002. A focus semantics for echo questions. In Ágnes Bende-Farkas & Arndt Riester (eds.), *Workshop on information structure in context*, 98–107. Stuttgart: IMS, University of Stuttgart.

Asher, Ronald E. & T.C. Kumari. 1997. *Malayalam*. London: Routledge.

Baker, Carl Lee. 1970. Notes on the description of English questions: the role of an abstract question morpheme. *Foundations of Language* 6. 197–219.

Baker, Mark C. & Nadya Vinokurova. 2009. On agent nominalizations and why they are not like event nominalizations. *Language* 85(3). 517–556.

Bandhu, Vishva (ed.). 1963–1966. *Ṛgveda with the padapāṭha and commentaries of Skandasvāmin, Udgītha, Veṅkaṭamādhava and Mudgala* 8 vols. Hoshiarpur, Punjab: Vishveshvaranand Vedic Research Institute.

Bayer, Josef. 2004. Decomposing the left periphery: Dialectal and cross-linguistic evidence. In Horst Lohnstein & Susanne Trissler (eds.), *The syntax and semantics of the left periphery*, 59–95. Berlin: Mouton de Gruyter.

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

Beck, Sigrid & Shin-Sook Kim. 1997. On wh- and operator scope in Korean. *Journal of East Asian Linguistics* 6. 339–384.

Becker, Misha. 1999. The *some* indefinites. In Gianluca Storto (ed.), *UCLA Working Papers in Linguistics, vol. 3, October 1999 — Syntax at Sunset 2*, 1–13. Los Angeles: UCLA Department of Linguistics.

Bhat, D.N.S. 1981. Physical identification in Kannada. *Studies in the Linguistic Sciences* 11(2). 1–8.

Bittner, Maria. 1994. Cross-linguistic semantics. *Linguistics and Philosophy* 17. 53–108.

Boër, Steven E. & William G. Lycan. 1975. Knowing who. *Philosophical Studies* 28(5). 299–344.

Böhtlingk, Otto & Rudolph Roth (eds.). 1855–1875. *Sanskrit-Wörterbuch*. St. Petersburg: Kaiserlichen Akademie der Wissenschaften.

Bruening, Benjamin. 2007. *Wh*-in-situ does not correlate with *Wh*-indefinites or question particles. *Linguistic Inquiry* 38(1). 139–166.

Brugmann, Karl. 1904. *Die Demonstrativpronomina der indogermanischen Sprachen: Eine bedeutungsgeschichtliche Untersuchung*. Leipzig: Trübner.

van Buitenen, Johannes Adrianus Bernardus (ed.). 1981. *The Bhagavadgītā in the Mahābhārata*. Chicago: University of Chicago Press.

Cable, Seth. 2007. *The grammar of Q: Q-particles and the nature of Wh-fronting, as revealed by the Wh-questions of Tlingit*. Cambridge, MA: MIT dissertation.

Cain, Bruce Dwayne. 2000. *Dhivehi (Maldivian): a synchronic and diachronic study*. Ithaca, NY: Cornell University dissertation.

Campbell, Lyle. 1997. *American Indian languages: The historical linguistics of Native America*. Oxford: Oxford University Press.

Cheng, Lisa Lai Shen. 1991. *On the typology of wh-questions*. Cambridge, MA: MIT dissertation.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Robert Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays in Minimalist syntax in honor of Howard Lasnik*, 89–115. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–53. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belleti (ed.), *Structures and beyond*, 104–131. New York: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36. 1–22.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gartner (eds.), *Interfaces + Recursion = Language?: Chomsky's Minimalism and the view from syntax-semantics*, 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Cohen, Ariel. 2006. How to deny a presupposition. In Ken Turner & Klaus von Heusinger (eds.), *Where semantics meets pragmatics*, 95–110. Amsterdam: Elsevier.

Cohen, Ariel. 2009. No alternative to alternatives. *Journal of Semantics* 26(1). 1–48.

Cole, Peter. 1975. The synchronic and diachronic status of conversational implicature. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics, volume 3: Speech acts*, 257–288. New York: Academic Press.

Cole, Peter & Gabriella Hermon. 1998. The typology of *wh*-movement: *Wh*-questions in Malay. *Syntax* 1. 221–258.

Condoravdi, Cleo. 2005. Not knowing or caring who. Ms., PARC and Stanford University. http://www-csli.stanford.edu/~cleoc/AC/fc-han.pdf.

Dayal, Veneeta. 1997. Free relatives and *ever*. In A. Lawson (ed.), *Proceedings from Semantics and Linguistic Theory (SALT) 7*, 99–116. Ithaca, NY: CLC Publications, Cornell University.

De Alwis, James. 1852. *The Sidath Sangarawa, a grammar of the Sinhalese language translated into English*. Colombo: William Skeen, Goverment Printer.

den Dikken, Marcel. 2006. *Either*-float and the syntax of co-*or*-dination. *Natural Language and Linguistic Theory* 24(3). 689–749.

Donnellan, Keith S. 1966. Reference and definite descriptions. *Philosophical Review* 75. 281–304.

É. Kiss, Katalin. 1998. Identificational focus versus informational focus. *Language* 74(2). 245–273.

Elizarenkova, Tatyana. 1972. Influence of Dravidian phonological system on Sinhalese. *International Journal of Dravidian Linguistics* 1(2). 126–137.

Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595.

Embick, David & Rolf Noyer. 2007. Distributed morphology and the syntax/morphology interface. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 289–324. Oxford: Oxford University Press.

Evans, Gareth. 1982. *The varieties of reference*. Oxford: Oxford University Press.

Fairbanks, Gordon H., James W. Gair & M.W.S. De Silva. 1968. *Colloquial Sinhalese (Sinhala)*. Ithaca, NY: South Asia Program and Department of Modern Languages and Linguistics, Cornell University.

von Fintel, Kai. 2000a. Singleton indefinites (Schwarzschild 2000). Ms., MIT. http://web.mit.edu/fintel/www/singleton.pdf.

von Fintel, Kai. 2000b. Whatever. In B. Jackson & T. Matthews (eds.), *Proceedings from Semantics and Linguistic Theory (SALT) 10*, 27–40. Ithaca, NY: CLC Publications, Cornell University.

Fukutomi, Yasuyuki. 2006. Japanese alternative questions and intervention effects in DP. In Changguk Yim (ed.), *Minimalist views on language design: Proceedings of the 8th Seoul International Conference on Generative Grammar*, 53–60. Seoul: The Korean Generative Grammar Circle.

Gair, James W. 1968[1998]. Sinhalese diglossia. *Anthropological Linguistics* 10(8). 1–15. [Reprinted with additional notes in Gair 1998:213–223].

Gair, James W. 1970. *Colloquial Sinhalese clause structures*. The Hague: Mouton.

Gair, James W. 1976[1998]. The verb in Sinhala, with some preliminary remarks on Dravidianization. *International Journal of Dravidian Linguistics* 10(8). 259–273. [Reprinted in part, with additional notes, in Gair 1998:200–209].

Gair, James W. 1980. Adaptation and naturalization in a linguistic area: Sinhala focused sentences. In *Proceedings of the annual meeting of the Berkeley Linguistic Society*, 28–43. Berkeley, CA: Berkeley Linguistics Society.

Gair, James W. 1982[1998]. Sinhala, an Indo-Aryan isolate. *South Asian Review* 6(3). 51–64. [Reprinted, in Gair 1998:3–24].

Gair, James W. 1983[1998]. Non-configurationality, movement, and Sinhala focus. Paper presented at the Linguistic Association of Great Britain, Newcastle, September 1983. [Published in Gair 1998:50–64].

Gair, James W. 1985[1998]. How Dravidianized was Sinhala phonology? In Richard L. Leed & Veneeta Acson (eds.), *Festschrift for Gordon H. Fairbanks*, 37–55. Honolulu: Oceanic Linguistics, University of Hawaii. [Reprinted with additional notes in Gair 1998:185–199].

Gair, James W. 1986[1998]a. Sinhala diglossia revisted, or diglossia dies hard. In Bhadriraju Krishnamurti, Colin P. Masica & Anjani Kumar Sinha (eds.), *South Asian languages: Structure, convergence and diglossia*, 322–336. Delhi: Motilal Banarsidass. [Reprinted with additional notes in Gair 1998:224–236].

Gair, James W. 1986[1998]b. Sinhala focused sentences: Naturalization of a calque. In Krishnamurti, Bhadriraju and Masica, Colin P. and Sinha, Anjani Kumar (ed.), *South Asian languages: Structure, convergence and diglossia*, 147–164. Delhi: Motilal Banarsidass. [Reprinted with additional notes in Gair 1998:155–169].

Gair, James W. 1992. AGR, INFL, Case and Sinhala diglossia, or Can linguistic theory find a home in variety. In Braj Kachru, Edward C. Dimock & Bhadriraju Krishnamurti (eds.), *Dimensions of South Asia as a sociolinguistic area: Papers in memory of Gerald B. Kelley*, 179–197. Delhi: Oxford India Book House.

Gair, James W. 1995[1998]. Syntactic theory, AGR, and Sinhala diglossia. Published in Gair 1998:237–267.

Gair, James W. 1998. *Studies in South Asian linguistics: Sinhala and other South Asian languages.* New York: Oxford University Press.

Gair, James W. & W.S. Karunatilaka. 1974. *Literary Sinhala.* Ithaca, NY: South Asia Program and Department of Modern Languages and Linguistics, Cornell University.

Gair, James W. & John C. Paolillo. 1989[1998]. Sinhala non-verbal sentences and argument structure. In Srivastav, Veneeta and Gair, James W. and Wali, Kashi (ed.), *Cornell working papers in linguistics*, 39–77. Ithaca, NY: Cornell University. [Reprinted with additional notes in Gair 1998:87–107].

Gair, James W. & Lelwala Sumangala. 1991. What to focus in Sinhala. In Germán F. Westphal, Benjamin Ao & Hee-Rahk Chase (eds.), *ESCOL '91: Proceedings of the Eighth Eastern States Conference on Linguistics*, 93–108. Columbus: Ohio State University Working Papers.

Gair, James W. & S. Suseendirarajah. 1981[1998]. Some aspects of the Jaffna Tamil verbal system. *International Journal of Dravidian Linguistics* 10(2). 370–384. [Reprinted with additional notes in Gair 1998:170–181].

Geiger, Wilhelm. 1938. *A grammar of the Sinhalese language.* Colombo: Royal Asiatic Society. [reprinted, New Delhi: Asian Educational Services, 1995].

Geiger, Wilhelm (ed.). 1941. *An etymological glossary of the Sinhalese language.* Colombo: The Royal Asiatic Society, Ceylon Branch. [Reprinted, New Delhi: Asian Educational Services, 1997].

Geis, Michael L. & Arnold M. Zwicky. 1971. On invited inference. *Linguistic Inquiry* 2. 561–579.

Geist, Ljudmila & Edgar Onea Gáspár. 2007. Specificity and implicatures. In Maria Aloni, Paul Dekker & Floris Roelofsen (eds.), *Proceedings of the Sixteenth Amsterdam Colloquium, December 17–19, 2007*, 109–114. Amsterdam: University of Amsterdam Institute for Logic, Language and Computation.

Geldner, Karl Friedrich (ed.). 1951–1957. *Der Rig-Veda aus dem Sanskrit ins Deutsche übersetzt mit einem laufenden Kommentar versehen.* Harvard Oriental Series 33–36. Cambridge, MA: Harvard University Press. [Vols. 33–35 reissued in one volume with preface by Michael Witzel, Cambridge, MA: The Department of Sanskrit and Indian Studies, Harvard University, 2003].

Ginsburg, Jason Robert. 2009. *Interrogative features.* Tucson, AZ: University of Arizona dissertation.

Ginsburg, Jason Robert. 2010. Qu-morpheme positional variation in Sinhala yes/no constructions. In Yong-Se Kang, Jiun-Shiung Wu, Dong-Hu Choi, Jong-Yurl Yoon, Seongha Rhee, Kee-Ho Kim, Jongseon Hong, Kyoung-Ae Kim & Hye-Kyung Kang (eds.), *Proceedings of the 2010 Seoul International Conference on Linguistics (SICOL-2010): Universal Grammar and individual languages*, Seoul: Hankookmunhwasa.

Graßmann, Hermann. 1873. *Wörterbuch zum Rig-Veda.* Wiesbaden: Harrassowitz.

Grice, Herbert Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics, 3: Speech acts*, 41–58. New York: Academic Press.

Grimshaw, Jane. 1981. Form, function and the language acquisition device. In C.L. Baker & John McCarthy (eds.), *The logical problem of language acquisition*, Cambridge, MA: MIT Press.

Hagstrom, Paul. 1998. *Decomposing questions.* Cambridge, MA: MIT dissertation.

Haida, Andreas. 2007. *The indefiniteness and focusing of wh-words.* Berlin: Humboldt University dissertation.

Halle, Morris & Alec Marantz. 1993. Distributed morphology. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Hamblin, Charles Leonard. 1958. Questions. *Australasian Journal of Philosophy* 36(3). 159–168.

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

Han, Chung-hye & Maribel Romero. 2001. Negation, focus and alternative questions. In K. Megerdoomian & L.A. Bar-el (eds.), *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL) 20*, 262–275. Somerville, MA: Cascadilla Press.

Han, Chung-hye & Maribel Romero. 2002. Ellipsis and movement in the syntax of Whether/Q...or questions. In Masako Hirotani (ed.), *Proceedings of NELS 32*, 197–216. Amherst, MA: GLSA.

Harrison, Annette & William J. Ashby. 2003. Remodelling the house: The grammaticalisation of Latin *casa* to French *chez*. *Forum for Modern Language Studies* 39(4). 386–399.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Heim, Irene. 1994. Questions. Lecture notes, MIT, Fall 1994.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Hock, Hans Henrich. 1986. '*P*-oriented' constructions in Sanskrit. In Krishnamurti, Bhadriraju and Masica, Colin P. and Sinha, Anjani Kumar (ed.), *South Asian languages: Structure, convergence and diglossia*, 15–26. Delhi: Motilal Banarsidass.

Hock, Hans Henrich. 1986/1991. *Principles of historical linguistics*. Berlin: Mouton de Gruyter 1st/2nd edn.

Hock, Hans Henrich. 1988. Review article: Finiteness in Dravidian. *Studies in the Linguistic Sciences* 18(2). 211–231.

Hock, Hans Henrich. 1989. Review of Steever (1988): *The serial verb formation in the Dravidian languages*. *Language* 65. 398–405.

Hock, Hans Henrich. 2008. Dravidian syntactic typology: A reply to Steever. In Rajendra Singh (ed.), *Annual Review of South Asian Languages and Linguistics*, 164–198. Berlin: Mouton de Gruyter.

Hock, Hans Henrich. to appear. Some consequences of Proto-Indo-European verb-finality. In *Selected papers from the 20th International Symposium on Theoretical and Applied Linguistics (ISTAL 20)*, Thessaloniki, Central Macedonia Periphery: Monochromia.

Hock, Hans Henrich & Rajeshwari Pandharipande. 1978. Sanskrit in the pre-Islamic sociolinguistic context of South Asia. *International Journal of the Sociology of Language* 16. 11–25.

Hoekstra, Eric. 1993. Dialectal variation inside CP as parametric variation. In Werner Abraham & Josef Bayer (eds.), *Dialektsyntax* (Special issue no. 5 of *Linguistische Berichte*), 161–179. Opladen: Westdeutscher Verlag.

Horn, Laurence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.

Horvath, Julia. 2007. Separating "focus movement" from focus. In S. Karimi, V. Samiian & W. Wilkins (eds.), *Phrasal and clausal architecture*, 108–145. Amsterdam: John Benjamins.

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

Jayaseelan, Karattuparambil A. 2008. Question particles and disjunction. Ms., Hyderabad, English and Foreign Languages University. http://ling.auf.net/lingBuzz/000644.

Kariyakarawana, Sunil M. 1998. *The syntax of focus and wh-questions in Sinhala*. Colombo: Karunaratne & Sons Ltd.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1(1). 3–44.

Karunatillaka, W.S. 1977. The position of Sinhala among the Indo-Aryan languages. *Indian Journal of Linguistics* 4. 1–6.

Kawashima, Ruriko. 1994. *The structure of noun phrases and the interpretation of quantificational NPs in Japanese.* Ithaca, NY: Cornell University dissertation.

Kim, Shin-Sook. 2002. Focus matters: Two types of intervention effect. Paper presented at WCCFL 21, Santa Cruz, CA.

Kishimoto, Hideki. 1992. LF pied piping: evidence from Sinhala. *Gengo Kenkyu* 102. 46–87.

Kishimoto, Hideki. 1997. Wh-in-situ and null operator movement in Sinhala questions. Ms., Hyogo University of Teacher Education. Draft dated 21 October 1997.

Kishimoto, Hideki. 2005. *Wh*-in-situ and movement in Sinhala questions. *Natural Language and Linguistic Theory* 23(1). 1–51.

Klein, Jared S. 1974. *The particle* u *in the Rigveda.* New Haven, CT: Yale University dissertation.

Klein, Jared S. 1978. The system of coordinate conjunctions in the Rigveda. *Indo-Iranian Journal* 20. 1–23.

Klein, Jared S. & Nancy L. Condon. 1993. Gothic -(*u*)*h*: A synchronic and comparative study. *Transactions of the Philological Society* 91(1). 1–62.

Ko, Heejeong. 2005. Syntax of *why*-in-situ: Merge in [Spec, CP] in the overt syntax. *Natural Language and Linguistic Theory* 23(4). 867–916.

Kodagoda, Ṅaṇaloka Sthavirayan Vahanse (ed.). 1967. *Amāvatura (Gurulugomi).* Colombo: Gunasena.

Kratzer, Angelika. 1998. Scope or pseudo-scope? Are there wide-scope indefinites? In Susan Rothstein (ed.), *Events and grammar*, 163–196. Dordrecht: Kluwer.

Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate phrases: the view from Japanese. In Y. Otsu (ed.), *The Proceedings of the Third Tokyo Conference on Psycholinguistics*, 1–25. Tokyo: Hituzi Syobo.

Kripke, Saul. 1972. Naming and necessity. In D. Davidson & G. Harman (eds.), *Semantics of natural language*, 253–355, 763–769. Dordrecht: Reidel.

Kripke, Saul. 1977. Speaker's reference and semantic reference. In Peter A. French, Theodore E. Uehling, Jr. & Howard K. Wettstein (eds.), *Studies in the philosophy of language*, 255–276. Morris, MN: University of Minnesota.

Kripke, Saul. 1980. *Naming and necessity.* Cambridge, MA: Harvard University Press.

Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1. 199–244.

Kuroda, Sige-Yuki. 1965. *Generative grammatical studies in the Japanese language.* Cambridge, MA: MIT dissertation.

Lalith Ananda, Malawenna Gamage. 2008. *The cleft construction in Sinhala.* New Delhi, Jawaharlal Nehru University MPhil Thesis.

Lehmann, Thomas. 1998. Old Tamil. In Sanford B. Steever (ed.), *The Dravidian languages*, 75–99. London: Routledge.

Lightfoot, David W. 1999. *The development of language.* Oxford: Blackwell.

Lindholm, James M. 1972. Cleft sentences in Tamil and Malayalam. In V.I. Subramoniam & Elias Valentine (eds.), *Proceedings of the First All India Conference of Dravidian Linguists*, 297–306. Trivandrum, Kerala: Dravidian Linguistics Association.

Ludlow, Peter & Stephen Neale. 1991. Indefinite descriptions: In defense of Russell. *Linguistics and Philosophy* 14. 171–202.

Miyagawa, Shigeru. 1987. LF affix raising in Japanese. *Linguistic Inquiry* 18(2). 362–367.

Miyagawa, Shigeru. 1998. On islands. Ms., MIT.

Montague, Richard. 1970a. English as a formal language. In Bruno Visentini et al. (ed.), *Linguaggi nella società e nella tecnica*, 188–221. Milan: Edizioni di Comunità.

Montague, Richard. 1970b. Universal grammar. *Theoria* 36. 373–398.

Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In K.J.J. Hintikka, J.M.E. Moravcsik & P. Suppes (eds.), *Approaches to natural language*, 221–242. Dordrecht: Reidel.

Moore, Harumi. 2003. Japanese *ka*-indefinites and the semantic domain of *unknown to the speaker*. *Journal of Pragmatics* 35. 589–614.

Murray, James, John Simpson & Edmund Weiner *et al.* (eds.). 2011. *The Oxford English dictionary (online).* Oxford: Oxford University Press.

Narayanapilla, P.K. (ed.). 1971. *Prācīna-Malayāḷa-gadya-mātrkakaḷ.* Thiruvananthapuram: Kerala University.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar.* Dordrecht: Kluwer.

Nyman, Elizabeth & Jeff Leer. 1993. *Gágiwdul.àt: Brought forth to reconfirm. The legacy of a Taku River Tlingit clan.* Fairbanks, AK: Alaska Native Language Center.

Ogawa, Kunihiko. 1976. *Japanese interrogatives: A synchronic and diachronic analysis.* San Diego: University of California dissertation.

Ogawa, Kunihiko. 1977. Where diachronic and synchronic rules meet: A case study from Japanese interrogaties and kakari-musubi. *Papers in Japanese Linguistics* 5. 193–242.

Okada, Keiko & John Grinstead. 2003. The emergence of CP in child Japanese. In Juana M. Liceras, Helmut Zobl & Helen Goodluck (eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002): L2 links*, 213–218. Somerville, MA: Cascadilla Press.

Paolillo, John C. 1992. *Functional articulation in diglossia: A case study of grammatical and social correspondences in Sinhala.* Stanford, CA: Stanford University dissertation.

Paolillo, John C. 1994. The co-development of finiteness and focus in Sinhala. In William Pagliuca (ed.), *Perspectives on Grammaticalization*, 151–170. Amsterdam: John Benjamins.

Paranavitana, Senarat. 1956. *Sigiri graffiti, Sinhalese verses of the eighth, ninth and tenth centuries.* London: Oxford University Press.

Partee, Barbara H. 1986. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk, Dick de Jongh & Martin Stokhof (eds.), *Studies in discourse representation theory and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris. [Reprinted in Portner and Partee (2002), pp.357–381].

Partee, Barbara H., Alice ter Meulen & Robert E. Wall. 1990. *Mathematical methods in linguistics.* Dordrecht: Kluwer.

Pesetsky, David. 1982. *Paths and categories.* Cambridge, MA: MIT dissertation.

Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation*, 262–294. Amsterdam: John Benjamins.

Pillai, P.V. 1973. *Early Malayalam prose: A study.* Trivandrum, Kerala: University of Kerala.

Portner, Paul & Barbara H. Partee (eds.). 2002. *Formal semantics: The essential readings.* Oxford: Blackwell.

Potts, Christopher. 2007. Conventional implicatures, a distinguished class of meanings. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 475–501. Oxford: Oxford University Press.

Raman Pilla, C.V. 1918. *Rāmarājabahadur.* [Reprinted by Little Prince Publishers, Kottayam, Kerala. 1983].

Ramchand, Gillian Catriona. 1997. Questions, polarity and alternative semantics. In *Proceedings of NELS 27*, 383–396. Amherst: University of Massachusetts, GLSA.

Ratanajoti, Hundirapola. 1975. *The syntactic structure of Sinhalese and its relation to that of the other Indo-Aryan dialects.* Austin: University of Texas dissertation.

Reichenbach, Hans. 1947. *Elements of symbolic logic.* New York: Macmillan.

Reinhart, Tanya. 1997. Quantifier scope: how labor is divided between QR and choice functions. *Linguistics and Philosophy* 20. 335–397.

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

Reynolds, C.H.B. 1964. Participial forms in early Sinhalese prose. *Bulletin of the School of Oriental and African Studies, University of London* 27(1). 129–150.

Reynolds, C.H.B. (ed.). 1970. *An anthology of Sinhalese literature up to 1815: selected by the UNESCO National Commission of Ceylon.* London: George Allen and Unwin.

Reynolds, C.H.B. 1980. *Sinhalese: An introductory course.* London: School of Oriental and African Studies.

Rhys Davids, T.W. & William Stede (eds.). 1921–1925. *Pāli-English dictionary.* Chipstead, Surrey: Pali Text Society.

Rizzi, Luigi. 2006. On the form of chains: criterial positions and ECP effects. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Wh-movement: Moving on*, 97–134. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2007. On some properties of criterial freezing. *Studies in Linguistics* 1. 145–158.

Rohrer, Christian. 1973. On the relation between disjunction and existential quantification. In Maurice Gross, Morris Halle & Marcel-Paul Schützenberger (eds.), *The formal analysis of natural languagess: Proceedings of the First International Conference*, 224–232. The Hague: Mouton.

Romero, Maribel. 1999. Intensional choice functions for *which* phrases. In T. Matthews & D. Strolovich (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 9*, 255–272. Ithaca, NY: CLC Publications, Cornell University.

Romero, Maribel & Chung-hye Han. 2004. On negative *yes/no* questions. *Linguistics and Philosophy* 27. 609–658.

Rooth, Mats. 1985. *Association with focus.* Amherst, MA: University of Massachusetts, Amherst dissertation.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Rooth, Mats. 1996. Focus. In Shalom Lappin (ed.), *The handbook of contemporary semantics*, 271–297. Cambridge, MA: Blackwell.

Rooth, Mats. 1999. Association with focus or association with presupposition? In P. Bosch & R. van der Sandt (eds.), *Focus: Linguistic, cognitive, and computational perspectives*, 202–226. Cambridge: Cambridge University Press.

Rullman, Hotze & Sigrid Beck. 1998. Reconstruction and the interpretation of *which*-phrases. In Graham Katz, Shin-Sook Kim & Heike Winhart (eds.), *Reconstruction: Proceedings of the 1997 Tübingen Workshop* Arbeitspapiere des Sonderforschungsbereichs 340, Nr. 127, 233–256. Tübingen and Stuttgart: Universities of Tübingen and Stuttgart.

Russell, Bertrand. 1905. On denoting. *Mind* 14(56). 479–493.

Russell, Bertrand. 1910–1911. Knowledge by aquaintance and knowledge by description. *Proceedings of the Aristotelian Society* 11 (n.s.). 108–128.

Saito, Mamoru. 1998. Wh-licensing in Japanese: Q-WH binding and unselective binding. Talk given at MIT, May 4, 1998.

Sansom, George. 1928. *An historical grammar of Japanese.* Oxford: Oxford University Press.

Schwarzschild, Roger. 2002. Singleton indefinites. *Journal of Semantics* 19(3). 289–314.

Sharma, Devyani. 1999. Nominal clitics and constructive morphology in Hindi. *Online Procedings of the Lexical-Functional Grammar '99 Conference*, Manchester, UK. http://cslipublications.stanford.edu/LFG/4/lfg99.html.

Shimoyama, Junko. 2006. Indeterminate phrase quantification in Japanese. *Natural Language Semantics* 14. 139–173.

Slade, Benjamin. under review. Indefinite, unknown. *Semantics and Pragmatics* .

Speijer, J.S. 1886. *Sanskrit syntax.* Leiden: E.J. Brill. [Reprinted, Delhi: Motilal Banarsidass, 1973].

von Stechow, Arnim. 1996. Some remarks on choice functions and LF-movement. In Klaus von Heusinger & Urs Egli (eds.), *Proceedings of the Konstanz workshop "reference and anaphorical relations"*, Fachgruppe Sprachwissenschaft, Universität Konstanz.

Sternefeld, Wolfgang. 2001. Partial movement constructions, pied piping, and higher order choice functions. In C. Féry & W. Sternefeld (eds.), *Audiatur vox sapientiae: a festschrift for Arnim von Stechow*, 473–486. Berlin: Akademieverlag.

Strawson, P.F. 1974. *Subject and predicate in logic and grammar.* London: Methuen and Co., Ltd.

Traugott, Elizabeth Closs & Richard B. Dasher. 2005. *Regularity in semantic change.* Cambridge: Cambridge University Press.

Travis, Lisa deMena. 1984. *Parameters and effects of word order variation.* Cambridge, MA: MIT dissertation.

Tsai, Wei-Tien Dylan. 1994. *On economizing the theory of A-bar dependencies.* Cambridge, MA: MIT dissertation.

Turner, Ralph Lilley. 1962–1966. *A comparative dictionary of Indo-Aryan languages.* London: Oxford University Press. [Reprinted, Delhi: Motilal Banarsidass, 1999].

Watanabe, Akira. 1993. Subjacency and s-structure movement of *wh*-in-situ. *Journal of East Asian Linguistics* 1. 255–291.

Watanabe, Akira. 2002. Loss of overt wh-movement in Old Japanese. In David Lightfoot (ed.), *Syntactic effects of morphological change*, 179–195. Oxford University Press.

Whitman, John. 1997. Kakarimusubi from a comparative perspective. In Ho-min Sohn & John Haig (eds.), *Japanese/Korean linguistics 6*, 161–178. Stanford, CA: CSLI.

Whitney, William Dwight. 1879/1889. *Sanskrit grammar, including both the classical language and the older dialects of Veda and Brahmana.* Leipzig: Breitkopf and Härtel 1st/2nd edn. [reprinted, Delhi: DK Publications, 1995].

Wijemanne, Piyaseeli. 1984. *Amāvatura, a syntactic study.* Colombo: Ministry of Higher Education.

Wilson, George M. 1991. Reference and pronominal descriptions. *Journal of Philosophy* 88. 359–387.

Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20. 399–467.

Wold, Dag E. 1996. Long distance selective binding: the case of focus. In Teresa Galloway & Justin Spence (eds.), *Proceedings from Semantics and Linguistic Theory (SALT) 6*, 311–328. Ithaca, NY: CLC Publications, Cornell University.

Wright, Joseph. 1910. *Grammar of the Gothic language.* Oxford: Clarendon. [2nd edn. with a supplement to the grammar by O.L. Sayce, 1954].

Yanagida, Yuko. 2006. Word order and clause structure in early Old Japanese. *Journal of East Asian Linguistics* 15(1). 37–67.

Yatsushiro, Kazuko. 2001. The distribution of *mo* and *ka* and its implications. In Maria Christina Cuervo, Daniel Harbour, Ken Hiraiwa & Shinichiro Ishihara (eds.), *Formal approaches to Japanese linguistics 3*, Cambridge, MA: MIT Working Papers in Linguistics.

Yatsushiro, Kazuko. 2009. The distribution of quantificational suffixes in Japanese. *Natural Language Semantics* 17(2). 141–173.

Yoon, James Hye Suk. 1996. A syntactic approach to category-changing phrasal morphology: nominalizations in Korean and English. In Hee-Don Ahn, Myung-Yoon Kan, Young-Suck Kim & Sookhee Lee (eds.), *Morphosyntax in generative grammar: Proceedings of the 1996 Seoul International Conference on Generative Grammar*, 63–86. Seoul: Hankuk Publishing Company.

Yoshida, Keiko & Tomoyuki Yoshida. 1996. Question marker drop in Japanese. *ICU Language Research Bulletin* 11. 37–54.

Zaefferer, Dietmar. 1991. Weiß wer was? Wer weiß was? Wer was weiß... w-Interrogative und andere w-Konstruktionen im Deutschen. In Marga Reis & Inger Rosengren (eds.), *Fragesätze und Fragen*, 77–93. Tübingen: Niemeyer.

Zermelo, Ernst. 1904. Beweis, daß jede menge wohlgeordnet werden kann. *Mathematische Annalen* 59. 514–516.

Zilva Wickremasinghe, Martino de, H.W. Codrington & Senarat Paranavitana (eds.). 1912–1933. *Epigraphia Zeylanica: being lithic and other inscriptions of Ceylon.* London: H. Frowde for the Govt. of Ceylon.

Last modified: 2011-07-12 00:53:10 -0500

Author: Benjamin Slade

Git Hash: d2c464362d1316072a2594559da8bb48c0409ac0