Sinhala and Tamil: A Case of Contact-Induced Restructuring

Harold Dharmasenan Thampoe

Dissertation submitted in partial fulfilment of the requirements

of the regulations for the degree of

Doctor of Philosophy

at

School of English Literature, Language and Linguistics

Newcastle University

September 2016

Abstract

The dissertation presents a comparative synchronic study of the morphosyntactic features of modern spoken Sinhala and Tamil, the two main languages of Sri Lanka. The main motivation of the research is that Sinhala and Tamil, two languages of diverse origins—the New Indo-Aryan (NIA) and Dravidian families respectively—share a wide spectrum of morphosyntactic features. Sinhala has long been isolated from the other NIA languages and co-existed with Tamil in Sri Lanka ever since both reached Sri Lanka from India. This coexistence, it is believed, led to what is known as the contact-induced restructuring that Sinhala morphosyntax has undergone on the model of Tamil, while retaining its NIA lexicon. Moreover, as languages of South Asia, the two languages share the areal features of this region. The research seeks to address the following questions: (i) What features do the two languages share and what features do they not share?; (ii) Are the features that they share areal features of the region or those diffused into one another owing to contact?; (iii) If the features that they share are due to contact, has diffusion taken place unidirectionally or bidirectionally?; and (iv) Does contact have any role to play with respect to features that they do not share? The claim that this research intends to substantiate is that Sinhala has undergone morphosyntactic restructuring on the model of Tamil. The research, therefore, attempts to answer another question: (v) Can the morphosyntactic restructuring that Sinhala has undergone be explained in syntactic terms? The morphosyntactic features of the two languages are analyzed at macro- and micro-levels. At the macro-level, a wide range of morphosyntactic features of Tamil and Sinhala, and those of seven other languages of the region are compared with a view to determining the origins of these features and showing the large scale morphosyntactic convergence between Sinhala and Tamil and the divergence between Sinhala and other NIA languages. At the micro-level the dissertation analyzes in detail two morphosyntactic phenomena, namely null arguments and focus constructions. It examines whether subject/verb agreement, which is different across the two languages, plays a role in the licensing of null arguments in each language. It also examines the nature of the changes Sinhala morphosyntax has undergone because of the two kinds of Tamil focus constructions that Sinhala has replicated. It is hoped, that this dissertation will make a significant contribution to the knowledge and understanding of the morphosyntax of the two languages, the effects of language contact on morphosyntax, and more generally, the nature of linguistic variation.

Dedication

I dedicate this dissertation to God Almighty—the Father, Christ Jesus and the Holy Spirit—for His abundant grace, provision, strength and wisdom, and my wife Grace for her love, care, patience and understanding without whom I would never have been able to start, continue or complete successfully my PhD.

Acknowledgements

I would like to express my gratitude and appreciation to the following who have contributed to the completion of my Doctoral studies.

From the depths of my heart I thank Professor Anders Holmberg, one of my supervisors, for suggesting this fascinating topic. His inspiration, guidance, support and patience made this dissertation possible. I would also like to thank sincerely Dr. William van der Wurff, my other supervisor, for his encouragement, advice support and patience. His subtle insights and good humoured criticism have significantly enhanced the quality of my dissertation. I feel incredibly privileged to have you both as my supervisors, for you have made my research a rewarding and memorable experience.

I wish to extend my sincere gratitude and appreciation to Professor Maggie Tallerman, Newcastle University, and Prof. Andrew Nevins, University College London, my panel of examiners, for their very useful questions, comments and suggestions at the viva and their comprehensive report, raising diverse issues which helped me immensely to improve my dissertation.

I am ever grateful to Professor Anders Holmberg, Professor Noel Burton-Roberts, Professor Maggie Tallerman, Professor Martha Young-Scholten, Professor Karen Corrigan, Professor Steve Walsh, Dr. S. J Hannahs, Dr. Geoffrey Poole, Dr. Joel Wallenberg, Dr. William van der Wurff, Dr. Marie-Claude Tremblay, Dr. Cristina Dye, Dr. Hermann Moisl, Dr. Dawn Knight, Dr. Francis Jones and Mr. Scott Windeatt who taught me during the two year taught component of my PhD program.

I gratefully acknowledge the grant awarded to me under the Scholarship Programme of the Higher Education for the Twenty First Century (HETC) Project, Ministry of Higher Education, Sri Lanka, without which I would not have been able to do my PhD. A special word of thanks goes to Professor Lalith Munasinghe, HRD Consultant, and his efficient team for all their assistance. I am also very thankful to Dr. D. M. S. Munasinghe and Dr. M. Alfred who acted as sureties to the bond for my study leave.

I wish to express my heartfelt gratitude to my brothers, sisters, aunts and uncles in Sri Lanka and in the UK for their invaluable prayers, encouragement, support, help in all its forms and for just being there always.

I am greatly indebted to my wife, Grace, for the many sacrifices she has made for me. Her intuitions in the two languages have helped me analyze various morphosyntactic features of the two languages.

Finally, I owe a debt of gratitude to God for His manifold mercies, not only for giving me this great opportunity in life, but for all He has done to help me finish it successfully.

Table of contents

Chapter 1

Chapter 2

List of abbreviations

A	Agent
ABL	Ablative
A/Cat	Subtype of asymmetric negation connected to the marking of grammatical categories (such as tense, aspect, mood etc.)
ACC	Accusative
A/Fin	Subtype of asymmetric negation connected to finiteness of verbal elements
AGR	Agreement
A/NonReal	Subtype of asymmetric negation connected to the marking of reality status of events
APPL	Applicative
ASP	Aspect
ASS	Assertive
Asy	asymmetric
C/COMP	Complementizer
CAUS	Causative
CNPC	Complex Noun Phrase Constraint
COM	Comitative
CONC	Concessive
CONJ	Conjunctive coordinator
COP	Copula
CP	Complementizer Phrase
CVB/CONV	Converbs
D/Det	Determiner
DAT	Dative
DEG	Degree word
DEM	Demonstrative
DISJ	Disjunctive coordinator
DOM	Differential Object Marking
DPDL	Discourse Pro Drop Language
DP	Determiner Phrase
e	Ellipsis
EMPH	Emphatic

F/FEM	Feminine
FOC	Focus
FocP	Focus Phrase
FOFC	Final-Over-Final Constraint
FUT	Future
GB	Government and Binding
GEN	Genitive
GU	Greenbergian Universal
HON	Honorific
HORT	Hortative
HUM	Human
IMP	Imperative
INAN	Inanimate
INCL	Inclusive
INDF/INDEF	Indefinite
INF	Infinitive
INFER	Inferential
INS	Instrumental
INVL/INVOL	Involitive
IP	Inflectional Phrase
L1	First Language
L2	Second Language
LF	Logical Form
LOC	Locative
M	Model language
MAT	Matter replication
N	Neuter
NEG	Negative
NIA	New Indo-Aryan
NMLZ	Nominalized verb form
NOM	Nominative
NON-HUM	Non Human
NP	Noun Phrase
NSL	Null Subject Language

Num	Numeral
O	Object
P	Patient
PASS	Passive
PST	PAST
PAT	Pattern replication
PFV	Perfective
PERM	Permissive
PF	Phonetic Form
PL	Plural
POL	Polar
PNG/GNP	Person-Number-Gender/Gender-Number-Person
PRS	Present
PRO	Null pronoun ‘big PRO’
pro	Null nominative pronoun ‘little pro’
PRON	Pronoun/Pronominal
PST	Past
PTCP/PART/PPLE	Participle
Q	Question
QUANT	Quantifier
QP	Quantifier Phrase
QUOT	Quotative
R	Replica language
RC/REL	Relative Clause
RELAT	Relativizer
REL-PRON	Relative Pronoun
RECP	Reciprocal
REDUP	Reduplicated
REFL	Reflexive
S	Subject
SG	Singular
Spec	Specifier
Sym	Symmetric negation
SPL	Subject Prominent Language

TP	Tense Phrase
TOP	Topic
TopP	Topic Phrase
TPL	Topic Prominent Language
v	Light verb
VOL	Volitive
VP	Verb Phrase
WALS	World Atlas of Language Structures
X	Oblique (adverbial modifier or adjunct)
XP	Maximal projection/phrasal constituent
φ	Phi-features
Ø	Null argument
*	Ungrammatical
1	First person
2	Second person
3	Third person

List of tables

List of figures

Chapter 1

Introduction

1.1 Research aim and objectives

Sinhala and Tamil, despite their diverse origins, share a wide range of morphosyntactic features. The former is classified as a New Indo-Aryan (NIA) Language, while the latter belongs to the Dravidian family of languages. Sinhala has long been geographically isolated in Sri Lanka from the other main-land NIA Languges of South Asia and, more importantly, coexisted with Tamil in Sri Lanka over two millennia. This coexistence which resulted in the contact between the two languages has, it is believed, led to their convergence—'any process through which two or more languages in contact become more like each other' (Thomason 2001a: 89). This research undertakes a comparative synchronic study of the morphonsyntactic features of the two languages with the aim of demonstrating that modern spoken Sinhala has undergone what is known as contact-induced restructuring on the model of modern spoken Tamil in Sri Lanka. For the purposes of studying the nature of contact-induced restructuring that Sinhala has undergone, a wide range of morphosyntactic features of Sinhala and Tamil are examined at macro- and micro-levels.

The research addresses the following questions: (i) what morphosyntactic features do Sinhala and Tamil share, and what features do they not share?; (ii) do the features that they share result from ancient language contact between Indo-Aryan and Dravidian languages on the South Asian mainland or contact between Sinhala and Tamil in Sri Lanka?; (iii) if, as will be argued, the shared features are due to contact between the two languages, has the diffusion of these morphosyntactic features been bidirectional or unidirectional?; (iv) has contact had any role to play with respect to features that they do not share? The claim that will be substantiated in this dissertation is that the direction of diffusion is unidirectional, that is, from Tamil to Sinhala. An additional question then is, (v) can the morphosyntactic changes that Sinhala has undergone be explained in syntactic terms?

This research intends to achieve the following objectives:

(i) it will fill the gap of a detailed comparative morphosyntactic study of the two languages, which has been long overdue;

(ii) the analysis of these features will undoubtedly provide new insights into the morphosyntax and semantics of the two languages;

(iii) since the research involves language contact not only between the two languages but also among languages of the region, with special reference to the areal features of the region these languages share, it will make a significant contribution to the study of language contact in the region;

(iv) this research links two distinct yet crucial areas in the field of syntax, namely linguistic typology and generative syntax, the combination of which is taken to be a fruitful approach for research.

In what follows, Sections 1.2–1.4 discuss the information necessary for the understanding and analysis of contact-induced changes that have taken place in Sinhala. Section 1.2 focuses on the two main types of language classification relevant to this research, namely genetic and areal classification. Section 1.3 provides a brief introduction to contact-induced language restructuring and its kinds. Section 1.4 outlines the development in the subfield of linguistic typology with special reference to language universals; the two approaches used in the research, namely linguistic typology and generative syntax; areal typology which is the subdomain of linguistic typology; and the areal features in the languages of South Asia. Section 1.5 traces the history of the two languages under investigation and provides a brief review of the works on the two languages. Finally, Section 1.6 describes the methodology adopted in this research.

1.2 Language classification

Brown and Ogilvie (2010) distinguish four ways in which languages are classified: genetic, areal, lexicostatic and typological. Genetic classification groups languages into families on the basis of descent from a presumed common ancestor. Areal classification groups languages together on the basis of structural features shared across language boundaries within a geographical area. Lexicostatic classification uses word comparisons as evidence of language relationships. Typological or morphological classification, which was used prior to the other three types in the nineteenth century (Malmkjær 2002b), supposes a small set of language types, traditionally word types (isolating, agglutinating, fusional, and polysynthetic), to which languages can be assigned.[1]

---

[1] Note that the epithet ‘typological’ as used in this classification is different from the ‘typological’ used in the sub-discipline of linguistic typology to collectively refer to features in phonology, morphology, syntax etc. (see 1.3 and 1.4 below).

Of the four types of classification, types one and two are the most widely used in language classification studies. The classification of languages into language families indicate that they descend from a common parent language termed proto language, e.g., German, Dutch, Frisian, English among others belong to the West-Germanic family, all descended from a common parent language known as proto-West-Germanic language. Even if there are no records of the proto language, it can be reconstructed from records of languages which descend from it. Reconstruction is the process of 'Working out features of dead or unrecorded languages, or of unrecorded earlier stages of single languages' (Trask 2007: 243). Languages in a particular language family 'share certain observable linguistic characteristics, such as words, sounds and grammatical patterns' (Pereltsvaig 2012: 8). One distinct characteristic of the languages of the same genesis is that they have cognates, that is, lexical or gramamtical items which resemble each other both in form and meaning.

The fact that languages, genetically unrelated, share the same features shows that the features they share have diffused from one or more languages to others due to the contact among them. Thomason and Kaufman (1988: 37–39) distinguish two kinds of changes that languages undergo as a result of contact: 'borrowing' and 'substratum interference' (a.k.a 'interference through shift'). According to them, borrowing is a process of language maintenance whereby foreign features are incorporated into a group's native language by speakers of that language, while substratum interference results from imperfect group learning during a process of language shift, that is, 'when a group of speakers shifting to a target language, fails to learn that target language perfectly'. This imperfect learning often leads to 'The errors made by members of the shifting group in speaking the target language then spread to the target language as a whole when they are imitated by original speakers of that language' (ibid. 39). Two relative terms, namely 'substrate' and 'superstrate' are also used in contact situations in which languages do not share equal prestige or power: if a dialect/language of lower prestige (substrate) influences a dialect/language of a dominant dialect/language, it is called substratum interference (Thomason and Kaufman 1988); if a dominant dialect/language (superstrate) influences a dialect/language of lower prestige, it is called superstratum influence.

Heine and Kuteva (2005: 237–8) refer to borrowing as 'L2>L1 replication', that is, the second language (L2) influences the first language (L1) and the substratum interference as 'L1>L2 replication', that is, L1 influences L2. According to them, in borrowing situations L1 is maintained, whereas in situations involving substratum interference L1 is not maintained because the speakers of L1 shift to L2 either completely or partially. Language maintenance and language shift are two of the three broad kinds of language contact situations; the third kind is creation of new contact languages (Winford 2003). The Sinhala-Tamil contact situation can broadly be classified as a borrowing situation in which, it is assumed, linguistic features diffuse unidirectionally from Tamil (L2) to Sinhala (L1) owing to contact (see Coperahewa (2007) for a similar view; see 1.5).

A geographical area in which languages share distinct features on account of diffusion resulting from prolonged language contact is known as a *Sprachbund* which is a German word given different meanings in the literature such as, 'language alliance', 'language union', 'language league', 'linguistic area'and 'adstratum relationship'; (Campbell 2010, Thomason and Kaufman 1988, Malmkjær 2002a, Winford 2003, Trask 2007 among others). Thomason (2001a: 99) defines a linguistic area as 'a geographical region containing a group of three or more languages that share some structural features as a result of contact rather than as a result of accident or inheritance from a common ancestor.' A case in point is the South Asian Linguistic Area in which diverse languages as a result of contact among them share a variety of features. Myers-Scotton (2002: 173) observes that Sprachbunds are alternatively called convergence areas; convergence, according to her, implies 'the coming together (and therefore the gaining) of [structural] features in this context.' Structural features, as Campbell (2010: 62) notes, 'are not just loan words, but also shared phonological, morphological, syntactic and other traits.'

In genetic and areal classification, languages are classified on the basis of the similarities between/among them. The similarities in typological features between genetically unrelated languages are often used to establish the extent to which the language(s) has/have undergone morphosyntactic restructuring. This does not mean that the differences in typological features between languages are unimportant for language contact studies. The differences in typological features indicate that these features resist contact-induced restructuring or they too may have been induced by contact. Instances in which differences resulted from contact between Sinhala and Tamil are discussed in this dissertation.

Comrie (1989: 201) claims that there are in principle four reasons for the similarities between languages: i) they could be due to chance; ii) they could stem from the fact that the two languages are genetically related, and have inherited the common property from their common ancestor; iii) the two languages could be in areal contact;

one language could have borrowed the property from the other, or both could have borrowed it from some third language, either directly or through the mediation of yet other languages; and iv) the property could be a language universal, either absolute or a tendency. Should two languages share more structural features than would be expected when the features they share are due neither to chance nor to the features they share being language universals, then they either belong to one family or they have been in contact with each other. The languages of the same genesis do share similar features because they inherited them from the common ancestor or proto language, e.g., Proto-Dravidian. On the other hand, languages genetically unrelated share similar morphosyntactic (a.k.a 'typological'; see below) features owing to contact between/among them. Comrie's (1989: 204) observation is of relevance: there are instances in which 'languages are in such intimate contact that a wide range of similarities [at a level higher and more complex than that of lexical items] arise between them, often to the extent that they seem to share more similarities with one another than with languages to which they are genetically more closely related.'

Considering the Sinhala-Tamil contact situation in light of what is discussed above, two of Comrie's four reasons, namely that these similarities are due to chance and that the similar features may be language universals, can safely be ruled out because the morphosyntax of the two languages is strikingly similar, and the similarities between them are far too many to have been due to these two reasons. Although Tamil and Sinhala belong to two different families, a cursory glance at the morphosyntactic features which Sinhala possesses together with those which Tamil and other Dravidian languages such as Malayalam, Telugu etc. and other NIA languages such as, Bengali, Hindi, Gujarati etc. possess indicates that Sinhala shares at least as many features with Tamil and other Dravidian languages as with NIA languages. The only reason for the convergence of Sinhala towards Tamil in particular and the Dravidian family of languages in general is that Tamil and Sinhala have coexisted for a long time on the island of Sri Lanka and have, therefore, been 'in such intimate contact'. Tamil in Sri Lanka, on the other hand, has undergone little change owing to its coexistence with Sinhala. The coexistence of Tamil and Sinhala has, therefore, been consequent upon the diffusion of both lexical items and a wide range of morphosyntactic features from Tamil to Sinhala.

1.3 Contact-induced restructuring

The process whereby languages in contact have undergone changes to become morphosyntactically similar is referred to in the literature as ‘contact-induced restructuring’ (Winford 2003), ‘metatypy’ Ross (1996, 1999) or ‘typological convergence’ (Aikhenvald and Dixon 2001a). This section briefly introduces contact-induced restructuring, the different processes involved in it and its characteristics. Ross (2007: 116) defines metatypy as ‘a diachronic process whereby the morphosyntactic constructions of one of the languages of a bilingual speech community is restructured, that is, the existing structure changed, on the model of the constructions of the speakers’ other language.’ He uses metatypy to mean “ ‘change in type’ … where ‘type’ is used in the sense implicit in the term ‘typology’: SOV and SVO are clause ‘types’, NOUN + DETERMINER and DETERMINER + NOUN are noun phrase ‘types’, and so on” (ibid: 124). These typological features include features from phonology, morphology, syntax and semantics (see 1.4 and 1.6 below).

Ross (2007) uses ‘replica language’ and ‘model language’ to denote the language that undergoes metatypy and that which provides the model for the restructuring of the replica language respectively. They are also referred to as ‘modified language’ and ‘source language’. Ross’s conceptualization of metatypy is based on his research on the contact-induced restructuring of Takia, an Oceanic Austronesian language of Karkar island, on the model of Waskia, a Papuan language of the Trans-New Guinea phylum. Metatypy occurs only when speakers of a speech community are ‘polylectal’, that is, the inmates speak two or more ‘lects’—the term Ross (2001) uses to denote languages or dialects. These bilingual speakers have an ingroup (primary) lect which is emblematic of their identity, and one or more outgroup (secondary) lect(s) used for external communication with people outside their group. The replica language undergoes metatypy, while the model language, which provides the metatypic model, serves as an intercommunity language, a sort of local lingua franca (Ross 1996; also 1999).

The speakers of the modified language who form a sufficiently tightknit community are well aware of their language as a marker of their separate identity, and hence, they attempt to reduce the borrowing of lexical forms from the model language.

As Hickey (2010: 19) sums up, ‘Metatypy is the sharing of organizational structures across languages in a situation where social attitudes disfavor the replication of concrete word forms whose origin in another language is easily identifiable.’ Note also that, as pointed out above, the speakers’ emblematicity/identity is tied to the lexicon. Some bilingual speakers, however, begin to use the intercommunity language so extensively that they are more at home in it than in the emblematic language of the community (Ross 1999: 1). Note also that these bilinguals speak the intercommunity language with a phonology which resembles that of their ingroup or emblematic language. The increase in the use of the inter-community language will eventually lead to one language (the primary lect) being adapted 'morphosyntactically to the constructions of another (the secondary lect), with no change occurring in the latter' (Ross 2003: 183).

Prior to contact-induced restructuring, the replica language is subject to lexical calquing and grammatical calquing which facilitate the morphosyntactic restructuring that follows (Ross 2007). Calque means 'loan word' or 'loan translation' which has been extended to loan items at phrase or clause level. Note the difference between lexical calquing and lexical borrowing (appropriating words from the source language); the former is part of contact-induced restructuring, while the latter is not. Lexical borrowing is independent of 'syntactic borrowing'. According to Ross (2007: 122), lexical calquing consists of 'remodelling lexical ways of saying things on the model of [the source language]'. That is, the meaning range of each lexical item is matched to the meaning range of an item in the model language until the two vocabularies are readily intertranslatable. Often, lexical and grammatical items in the source language are adapted to conform to the morphophonological characteristics of the replica language (Lucas 2015). In some instances, the speakers, while modelling the new lexical items, produce items with the same morphophonological composition as those in the model language, which would result in items totally or partially homophonous with those of the model language.

As stated by Ross (2007: 122), 'Grammatical calquing consists of remodelling grammatical ways of saying things on the metatypic model', that is, the creation of paradigms of closed-set items of the replica language with meanings that match the corresponding items of the model language. These twin processes of lexical and grammatical calquing are followed by morphosyntactic restructuring. In effect, the speakers of the replica language restructure its morphosyntax by modelling it on the morphosyntax of the source language, exploiting the vocabulary/morphology of the replica language. What takes place via these processes, as Ross (2007: 116) sums up, is that 'The constructions of the replica language are changed through metatypy so as to match those of the model language in meaning and morphosyntax.'

Referring to this kind of changes made by bilingual speakers as 'spontaneous replication' or 'speaker innovation', Heine and Kuteva (2008) rightly point out that it is highly idiosyncratic and is often referred to as ‘speech errors’. On some occasions the replicated forms survive and, by constant use, may become part of the speech habits of a group of speakers, spreading to other groups of speakers and eventually to the speech community. According to Heine and Kuteva (2008), this process does not lead to linguistic change because these innovations may remain restricted to a specific period of time, but may subsequently be abandoned by the same speakers or the next generation of speakers. Only if such innovations acquire some stability across time, has ‘grammatical replication’ or contact-induced restructuring taken place.

Matras and Sakel (2007) distinguish two kinds of replication, resulting from contact, namely MAT(ter) and PAT(tern) replication. They define MAT replication or ‘replication of linguistic matter’ as direct replication of morphemes and phonological shapes from a source language; and PAT replication as the process whereby patterns of distribution, of grammatical and semantic meaning, and of formal-syntactic arrangement at various levels (discourse, clause, phrase, or word) are modelled on an external source. In the latter, they note, the formal substance or matter is not imported, but is taken from the inherited stock of forms of the recipient or replica language (see Matras 2010, 2011).

Following Weinreich (1964[1953] who uses the term ‘grammatical replication’ to characterize contact-induced changes, Heine and Kuteva (2011; see also 2005, 2008) argue that grammatical replication is one of the processes of transfer observed in language contact. The different kinds of linguistic transfer in language contact situations which Heine and Kuteva (2011) distinguish are shown in figure 1 (adapted from figure 23.1, p. 292, ibid. 292):

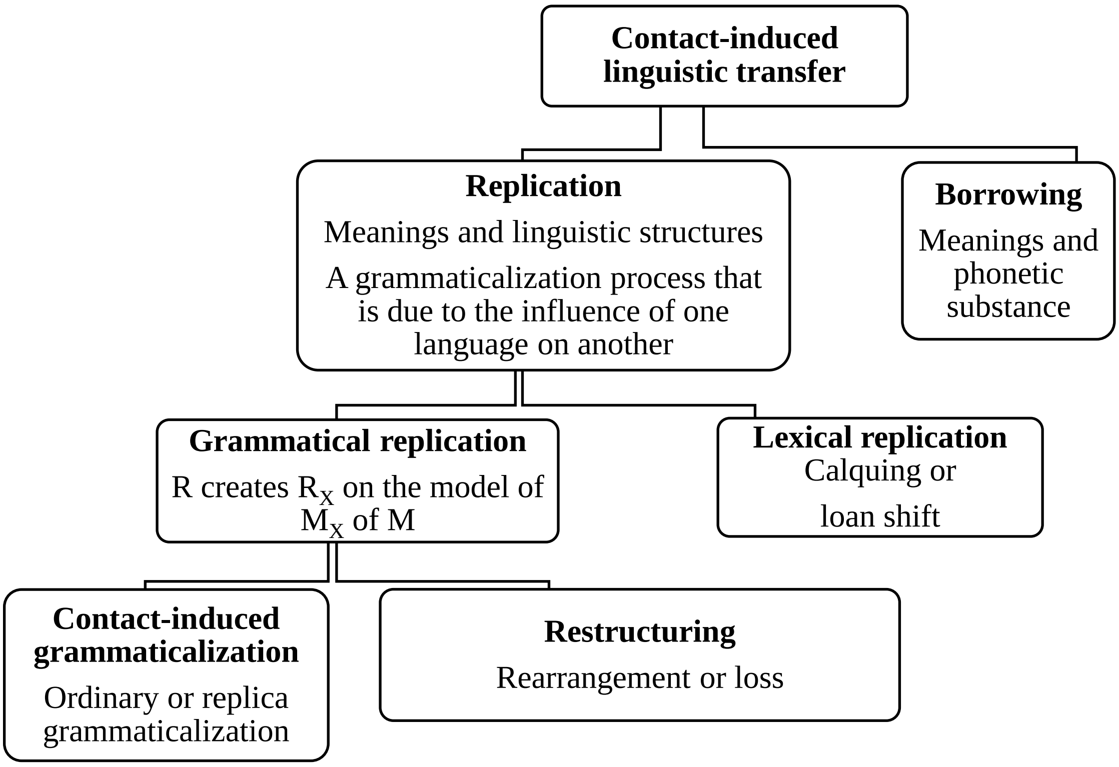

Contact-induced linguistic transfer
Replication
Meanings and linguistic structures
A grammaticalization process that is due to the influence of one language on another
Borrowing
Meanings and phonetic substance
Grammatical replication
R creates R_X on the model of M_X of M
Lexical replication
Calquing or loan shift
Contact-induced grammaticalization
Ordinary or replica grammaticalization
Restructuring
Rearrangement or loss

Figure 1.1: Main types of contact-induced linguistic transfer

Given below each type and sub-type of contact-induced linguistic transfer in figure 1.1 is a brief description of the processes involved. It is important to provide more details about ‘contact-induced grammaticalization’ and ‘restructuring’ to accurately characterize the contact-induced changes that Sinhala has undergone on the model of Tamil. Heine and Kuteva (2005) distinguish two kinds of contact-induced grammaticalization, namely ordinary and replica grammaticalization, which involve the mechanisms shown in table 1.1:

Ordinary grammaticalization	**Replica grammaticalization**
(a) Speakers [of R] notice that in language M there is a grammatical category Mx. (b) They create an equivalent category Rx in language R on the basis of the use patterns available in R. (c) To this end, they draw on universal strategies of grammaticalization, using construction Ry in order to develop Rx. (d) They grammaticalize Ry to Rx.	(a) Speakers [of R] notice that in language M there is a grammatical category Mx. (b) They create an equivalent category Rx in language R, using material available in R. (c) To this end, they replicate the grammaticalization process they assume to have taken place in M, using an analogical formula of the kind [My > Mx]: [Ry > Rx] (d) They grammaticalize Ry to Rx.

Table 1.1: Ordinary and replica grammaticalization processes

According to Heine and Kuteva (2005), the difference between ordinary and replica grammaticalization is that in the former, the category grammaticalized in the replica language is not modelled on the corresponding category in the model language, whereas in the latter, the category grammaticalized is modelled on the corresponding category in the model language (see (c) in table 1.1).

Restructuring, on the other hand, is a process whereby an existing structure is rearranged/replaced or lost as a consequence of language contact (other contact-induced restructuring related theories are discussed where necessary).

Heine (2007) provides a set of diagnostics for contact-induced restructuring, of which the ones relevant to Sinhala-Tamil contact situation are given below. Assume that in a contact situation which involves a model language M and a replica language R, there is a linguistic property P_R belonging to R that is believed to have been replicated from a corresponding property P_M in M.

(i) Intertranslatability: if a P_R in R is immediately intertranslatable with P_M in M or P_R and P_M display structural isomorphism—a property of two or more structures whose constituent parts are in a one to one correspondence with each other at a given level of abstraction (Crystal 2008).

(ii) Genetic patterning: if P_R is not found in other dialects or languages genetically related to R, while the corresponding category P_M of M is found in the languages genetically related to M. The diagnostic of genetic patterning can also be framed in terms of genetic inheritance, in that P_M can be reconstructed back to the earlier stages of P in M but P_R cannot similarly be reconstructed in R.

(iii) Paired structural similarity: if there is a set of two or more properties shared by M and R whose presence can neither be coincidental nor be due to shared genetic relationship. This kind of paired structural similarity can be seen in cases of shared polysemy. Polysemy is the capacity of a word or phrase to have multiple meanings which may be unrelated. Instances of auxiliation are often regarded as shared polysemy in that M and R display the same kind of main verb-auxiliary pairing, i.e., the same verb can be used as a verb and an auxiliary.

(iv) Paired grammaticalization: if two neighbouring but genetically unrelated languages M and R share not only one but two grammaticalization processes for the same general grammatical function.

The extent to which a replica language is subjected to contact-induced restructuring depends on the intensity of its contact with the model language. There is a continuum in the process of contact-induced changes ranging from relatively slight lexical borrowing under casual contact to extreme structural borrowing under very intense contact (Winford 2003). Matras (2009:156) provides the condensed version of Thomason and Kaufman's (1988:74–76) well-known borrowing scale, given in figure 1.2):

Casual contact	Category 1:	content words
↓	Category 2:	function words, minor phonological features, lexical semantic features
	Category 3:	adpositions, derivational suffixes, phonemes
	Category 4:	word order, distinctive features in phonology, inflectional morphology
Intense contact	Category 5:	significant typological disruption, phonetic changes

Figure 1.2: Thomason and Kaufman's (1988) borrowing scale

The Sinhala-Tamil contact situation demonstrates relatively high intensity of contact between Sinhala and Tamil which is close to or the same as Category 5 in that subtle items at the morphophonological level are replicated, as evidenced in (1) below (for ease of reference, in examples, Tamil and Sinhala forms are marked T and S respectively): [2]

(1) T. kumar than-da mahan-ai paaram-kudu-th-athu Mala-tta
kumar self-GEN son-ACC charge-give-PST-NMLZ Mala-LOC
S. kumar thaman-ge puthaa-wə bhaarə-dunn-e Mala-tə
kumar self-GEN son-ACC charge-give.PST-NMLZ Mala-DAT

‘It is/was to Mala that Kumar entrusted/handed over his son.’

The above cleft constructions indicate that one of the languages has replicated this construction from the other. The implications of 1T/S for the replication of the cleft construction by Sinhala from Tamil are analyzed below in terms of the diagnostics (stated within brackets). The constituent structure of (1) in the two languages is the same. Note that the two sentences can be translated word by word/morpheme by morpheme (intertranslatability), as shown in table 1.2:

	Features	**Tamil**	**Sinhala**
1	Possessive/genitive marker	*-da*	*-ge*
2	Anaphor ‘self’	*than*	*thaman*
2	Accusative marker	*-ai*	*-wə*
3	Past tense verb ‘entrust/hand over’	*paaram-kudu-th-*	*bhaarə-dunn-*
4	Nominalized verb suffix	*-athu*	*-e*
5	Locative/Dative marker	*-tta*	*-tə*

Table 1.2: Convergence between features of Tamil and Sinhala

In fact, this kind of cleft construction is distinctively Dravidian; the other major Dravidian languages have this construction (see Krishnamurti 2003: 425–426, Jayaseelan and Amritavalli 2005), while no other NIA language, according to Gair (1980; for details see 6.1), has this kind of cleft construction, which indicates that Sinhala has replicated it from Tamil (genetic patterning). Sinhala has also replicated

---

[2] In this dissertation, glossing is done in conformity with ‘The Leipzig Glossing Rules (2008)’, and the abbreviations used in glosses are those stipulated therein. In instances where abbreviations are not available, they are created with details given within brackets (see list of abbreviations). When an element is formally unsegmentable but has two clearly distinguishable meanings or grammatical properties, e.g., the stem verb and the tense in Sinhala, they are separated by periods.

another Dravidian focus construction which involves a focus particle (6.1) for the same semantic function (paired grammaticalization). Of importance in all features is the adoption of form-meaning units—matter replication—such as the anaphor *than* in Tamil and *thaman* in Sinhala 'self'. The verb for 'entrust: give someone the charge of/responsibility of looking after' in Tamil and Sinhala is a compound verb, consisting of *paaram* (1T) and *bhaarə* (1S) 'charge/responsibility' and *kuduthathu* (1T) and *dunne* (1S), the nominalized form of the verb 'give'.

The language that undergoes restructuring, the 'metatypised' language, maintains forms resembling those in its genetic relatives, but the meanings of these forms have undergone changes. As a result, when the restructured forms are compared with their functional equivalents in the model language, they "resemble each other quite systematically in their semantic organization, in the structures of their paradigm and in their morphosyntax" (Ross 1996: 182). A metatypised language, therefore, shows two types of resemblances i) correspondence in form and partial resemblances in meaning to its genetic relatives; and ii) more precise correspondences in meaning and resemblances in morphosyntax to its metatypic model (ibid. 182). The only difference between (1T) and (1S) is that the entrustee 'Mala' is marked by the locative postposition in the former, while it is marked by the dative case in the latter. However, the fact that the locative postposition *-tta* in Tamil and the dative case *-tə* in Sinhala are near-homophonous suggests that the latter may have been modelled on the former (see 3.6.1.1).

Heine and Kuteva (2011) point out that contact-induced change or transfer has an areal dimension because it creates areas of structural relationship. As a result of the change, the replica and model languages share a structural isogloss—the geographic boundary within which a particular linguistic item is used—that was not there before language contact. The ideal oft cited example is the morphosyntactic restructuring of varieties of Indo-Aryan Marathi and Urdu, and Dravidian Kannada in the Indian village of Kupwar. Gumperz and Wilson (1971: 154) point out that although the sentences in the three languages—in an example they give—are lexically distinct in almost every respect, they have identical grammatical categories and constituent structures and it is possible to translate one sentence into the other by simple morpheme by morpheme substitution.

In his taxonomy of language contact situations, Winford (2003: 11–24) distinguishes four kinds of convergence situations: (i) contiguous geographical location—associated with moderate structural diffusion, e.g., the Balkan *Sprachbund*;

(ii) intra-community multi-lingualism—associated with heavy structural diffusion, e.g., the influence of Marathi, Urdu and Kannada on one another in Kupwar. (iii) intense pressure on a minority group—also associated with high structural diffusion, e.g., Tibetan influence on Wutun; and (iv) intense inter-community contact (trade or exogamy)—associated with heavy lexical and/or structural diffusion, e.g., the languages of Northwest New Britain. As (1) indicates, the Sinhala-Tamil contact situation involves more than moderate structural diffusion. There are good reasons to assume that it conforms to Winford's situation (ii), that is, intra-community multi-lingualism, involving Sinhala and Tamil may have existed for a long time, probably from the initial stages of contact (more information in 1.5 below) till early post-independent/colonial period in Sri Lanka.

1.4 Linguistic and areal typology

As indicated in the preceding section, contact-induced restructuring involves changes in linguistic typological features. In most language contact studies (including the present research), whether languages have been in contact is determined by examining the convergence of the typological features in two or more languages. The outcome of language contact, in effect, the restructuring, is that the typological profile of the languages in contact, either in both or in one of them, changes and becomes similar to one another. Over the years, language contact studies have become symbiotically related to linguistic typological studies. Matras (2010: 68) rightly points out that 'language typologists identified contact as a potential trigger for typological change and sought to apply functional-typological models to explain the kind of change' while 'contact linguists […] recognized that grammaticalization was involved in many of the structural processes of change observed in contact situations.' Given the importance of linguistic typological features in language contact studies like the present research, it is useful to have a closer look at the linguistic typological features with regard to language contact.

Croft (2003: 1–2) provides three definitions of typology which he calls typological classification, typological generalization and functional-typological explanation. Typological classification refers to the classification of structural types across languages. A case in point is the morphological typology of the nineteenth and twentieth centuries (1.2). Typological generalization which began with Greenberg's (1963) discovery of language universals is the study of patterns that systematically occur across languages—recurring patterns which all or most languages share.

Functional-typological explanation which is allied to functionalism holds that ‘linguistic structure should be explained in terms of linguistic function’ Croft (2003:2). This theoretical approach became recognized in the 1970s (see Comrie 1989; Croft 2003; Haspelmath et al. 2005b). Croft (2003) claims that the three linguistic definitions of typology correspond to the three stages of any empirical scientific analysis: (i) the typological classification represents the observation of an empirical phenomenon (language) and the classification of what is observed; (ii) the typological generalization—language universals—is the formulation of generalizations over what has been observed; and (iii) the functional-typological approach constructs explanations of the generalizations over what has been observed (Croft 2003: 2).

Linguistic typology centres around language or typological universals. Cristofaro (2010: 227) defines typological universals as ‘empirically established generalizations that describe distributional patterns for particular grammatical phenomena across languages.’ ‘These distributional patterns,’ she notes, ‘are regarded as universal to the extent that they are found in all languages or in a statistically significant number of languages.’ Language universals are of two kinds. Firstly, there are absolute universals that are exceptionless and which hold for every member of their universe, e.g., in all languages there are stop consonants. Secondly, there are implicational universals which hold for a contextually delimited universe of languages. The implicational universals often come in pairs, e.g., ‘With overwhelmingly greater-than-chance frequency, languages with normal SOV order are postpositional’ (Greenberg’s 1963: universal 4) (for the classification of language universals, see Moravcsik 2011). Language typology has developed into a major area of research in the last three decades in linguistics, culminating in the publication of Haspelmath et al.’s (2005b) *The World Atlas of Language Structures* (WALS) together with its subsequent online version (Dryer and Haspelmath 2013; see 1.6 for details).

Also emerging, in the 1960s, more or less concurrently, with Greenberg’s approach to language universals was Noam Chomsky’s approach to universals posited in what is known as generative grammar. These two approaches are relatively different in terms of their attitude to evidence for and explanation of universals (Malmkjær 2002b). ‘The difference between the generative and typological approaches to universals,’ as Croft (2003: 4) claims, ‘can be traced to the traditions to which Chomsky and Greenberg responded. The generative approach emerged as a reaction against behaviourist psychology, while the typological approach emerged as a reaction against anthropological relativism.’ Chomsky’s main claim is that the innate universals of grammatical competence which a child is endowed with facilitate first language acquisition. The main argument of his theory is what he calls ‘poverty of stimulus’, that is, the notion that the output of the acquisition process (the adult languages) is underdetermined by the input (or stimulus)—the linguistic experience of the child. The poverty of stimulus argument posited by Chomsky is purely deductive and the generative approach is basically a rationalist approach to language. On the other hand, the universals in the typological approach posited by Greenberg are generalizations across languages, the type of argument employed is inductive ‘in keeping with typology’s empiricist approach to language’ (Croft 2003: 5).

Chomsky distinguished two kinds of universals, namely formal and substantive universals, some of which are features of all languages, while the others represent a set of features from which each language selects a subset (Malmkjær 2002b: 325). This distinction was made in what is identified as the first phase of generative syntax. These two types of universals were subsequently incorporated into the principles and parameters model postulated in Chomsky (1981), which is identified as the second phase.[3] Principles are abstract rules related to grammar common to all languages, while parameters are language-specific markers set for individual languages. One of the first principles posited in the tradition of principles and parameters is the principle of structure-dependency which means that the interpretation of linguistic expressions relies on the structural relationship between the constituents of the expressions rather than on the linear relationship between items. An important parameter postulated by Chomsky is the head parameter which stipulates the order of elements in a language: the head which is essential in any phrase occurs in languages either on the left of the phrase or on the right. The head parameter stipulates these two possibilities of which a language chooses one, in the ideal case consistently across categories, although in practice there will typically be some exceptions (see Baker 2008). In summary, generative grammar attempts to account for the surface structure of sentences in any language ‘with reference to certain highly abstract features which are shared by all languages because they are innate in humans’ (Malmkjær 2002b: 326).

These two approaches outlined above, namely linguistic typology and generative grammar, are used in this research. The former is used for the macro-level analysis of morphosyntatic features, while the latter is used for the micro-level analysis.

---

[3] Government and Binding Theory and the Minimalist Program, which are considered as the third and fourth phase of generative syntax, are not described as they are not relevant to the discussion here (see Brown 2002 for an overview of generative syntax).

From the point of view of historical linguistics, language universals play an important role in linguistic typology including its later subdomain, areal typology (see below), because these universals may reflect tendencies in languages to change towards preferable types of sound patterns, syllabic structures and even syntactic arrangements (Anderson 2002: 218). With the publication of Haspelmath et al.'s (2005b) and more importantly, the launch of its online database which made possible vast scale comparative studies of linguistic features across languages, there has been an increasing interest in the distribution of linguistic features in languages across contingent areas. According to Koptjevskaja-Tamm (2011: 577–8), areal typology is the study of the diffusion of typological features across languages in the same region. Discussing the difference between traditional linguistic typology and areal typology, she points out that the primary goal of the former is to describe and explain '(im)possible cross-linguistic variation in general and in using these results for getting insights into possible vs. impossible human languages,' while the latter 'holds that it can be of limited value to search for a possible human language without simultaneously investigating its genetically and areally determined manifestations and trying to uncover the possible historical reasons behind this variation.'

Following from the discussion in 1.2, the two main factors that account for the similarities between the typological features of languages are that either these languages are genetically related and inherited the similar features they share from the common ancestor or they are geographically located in the same region and acquired these features owing to contact. The features acquired via contact are deemed to be more important than those which are inherited from the common ancestor. Noonan (2010) is of the view that traditional genetic classification at higher taxonomic levels affords less information about the structure of a language than knowing where in the world that language is spoken. Knowing that Irish and Hindi are Indo-European languages, Noonan notes (fn. 3, p. 63, ibid.), tells us almost nothing about the structures of the two languages; knowing that Hindi is South Asian tells us a good deal more. The distinct features which languages of the same region share owing to their being in contact are called areal features. These features tend to delimit areas of structural relationship, resulting in the languages in contact sharing a structural isogloss. The branch of linguistics that studies these features is areal linguistics. Koptjevskaja-Tamm (2011: 577) observes that traditional concerns of areal linguistics have been the 'Diffusion of structural features across linguistic boundaries, identification of convergence areas, and, in general, similarities between geographically contiguous languages.'

The languages spoken in South Asia, the region to which the two languages under study belong, are either genetically related or in contact, being geographically contiguous. Ebert (2010: 995) notes that the roughly 450 languages of the region belong to four different language families, namely Indo-European, Dravidian, Austroasiatic (e.g., Khasi and Munda) and Sino-Tibetan. In terms of the speakers, those of Indo-Aryan and Dravidian constitute approximately 98% of the total population of the region, while those of Austroasiatic and Tibeto-Burman together do not constitute more than 2%. The region is regarded as a rich area for linguistic studies, especially language contact studies because, as Munshi (2010: 522) points out:

> In South Asia, from region to region, the accents, dialects and languages change. The dividing lines between languages are often distorted by large numbers of overlapping and interwoven dialects. It is extremely rare that speakers use only one language; multilingualism is the norm. Speakers often switch between two or more languages.

These trends of the region indicate that widespread language contact among languages predominantly from the two major language families, i.e., NIA and Dravidian, in the South Asian region is common. The South Asian region is classified as a linguistic area (see Ebert 2010, Emeneau 1956) and the languages which belong to the region irrespective of their origin are claimed to share the following areal features: retroflex consonants, subject-object-verb (SOV) order, postpositions, absence of prefixes, presence of dative-subject construction, absence of the verb ‘to have’, conjunctive or absolutive participles, morphological causatives, explicator compound verbs and sound symbolic forms based on reduplication (Campbell 2010: 62–63; see also Ebert 2010; Velupillai 2012; Subbarao 2012).

The Sinhala-Tamil contact situation is relatively complex. Here are two languages believed to belong to two different families, originating in one area, South Asia, where they acquired the areal features of the region. Subsequently they were relocated to another narrow area, Sri Lanka, where they have been in contact since then. In their place of relocation one language, Tamil, has not undergone any major changes which make it differ morphosyntactically from its mother language in South India, while the other, Sinhala, has undergone considerable restructuring on the model of Tamil. There are, therefore, two kinds of similarities between Tamil and Sinhala: (a) the areal features and others common to all South Asian languages given their origin in the South Asian mainland; and (b) the features which Sinhala shares with Tamil because of the former’s contact with the latter. As will be shown in Chapter 4, most of the areal features of South Asia are claimed to have diffused from Tamil/Dravidian and it is difficult to determine when Sinhala acquired the areal features mentioned, whether in South Asia through other languages or in Sri Lanka through Tamil.

1.5 Languages in contact: Sinhala and Tamil

This section provides an outline of the history of Sinhala and Tamil in Sri Lanka with respect to the contact between the two languages as recorded in the existing literature. Sinhala (a.k.a Sinhalese) and Tamil are the two national and official languages in Sri Lanka. Sinhala is classified as an NIA language along with others such as Marathi, Oriya, Gujarati, Bengali, Hindi etc. (Masica 1991). It is the mother tongue of the Sinhalese, the largest ethnic group that makes up nearly 75 percent of the Sri Lankan population. In 543 BC ‘a band of North-Indian wanderers’ (Gair 2007: 847) either from Bengal or Gujarat reached Sri Lanka, which was called Ceylon then. They spoke and hence brought to Sri Lanka ‘an archaic Parakritic dialect, similar to Pali or to the language of the Asoka inscriptions’ (Geiger 1937: 20). According to Hettiaratchi (1959: 36), ‘The language of the earliest extant records, viz. the cave and the early rock inscriptions, resembles closely Middle Indo-Aryan, and possesses the characteristics of Parakrit.’ In the prehistoric period, Sri Lanka was mainly inhabited by indigenous people called Yakkas and Nagas who spoke *Hela* (a.k.a *Elu*, *Helu*) (de Silva 2001). It is believed that Sinhala came into being as a consequence of the merging of the Parakrit spoken by the Aryan immigrants and *Hela* spoken by the aborigines. There are no significant works in the literature which attempt to tease apart the Hela element from the Indo-Aryan element in Sinhala.

The other inhabitants of Ceylon at the arrival of Aryan immigrants to Ceylon are those who belong to the tribe called Vaddas also Veddas meaning ‘hunter gatherers’ who spoke a non-Aryan language. Hettiaratchi (1959: 34) notes that ‘When the Vaddas came in contact with the Sinhalese [North-Indian immigrants], they first adopted from them a number of Aryan words.’ As a consequence, the Sinhala loan words increased in the Vadda language, and the aboriginal language was gradually replaced by colloquial (modern spoken) Sinhala. According to Gair (1998a: 4), though the Vadda language is considered to be a dialect of Sinhala, it has had a substrate influence on Sinhala which may account for the existence of a number of items in the Sinhala vocabulary which cannot be traced to either Indo-Aryan or Dravidian languages. He adds that the role that some indigenous language(s) played in the formation of Sinhala has not really been investigated, but it may well have been more than is generally recognized.

Tamil, on the other hand, is one of the four major Dravidian languages; the other three are Malayalam, Kannada and Telugu (Krishnamurti 2003). Tamil is spoken by Tamils predominantly from the north, east and central parts of Sri Lanka, and the majority of the Muslim community in Sri Lanka. Tamil found its way into Sri Lanka via various South Indian invasions of Ceylon from 2nd Century BC and subsequently developed in Sri Lanka into a distinct variety of Tamil. Following South Indian Tamil invasions, several Tamil kings from South India reigned in Sri Lanka. By virtue of being the language of the rulers, Tamil was the language of power. Disanayaka (2007) distinguishes the then Tamil in Sri Lanka as the donor language and Sinhala as the recipient language and notes that the political power and prestige of Tamil has influenced Sinhala. As a consequence, the ordinary Sinhala folk aspired 'to use this language in interpersonal communication, even at the expense of their own to attain the social status of the day' (Disanayaka 2007: 6). In the development of Sinhala, Coperahewa (2007: 140) observes, 'for many centuries it was constantly influenced by the dialects spoken by new immigrants who came to the island from different parts of India, especially from South India, for commercial and religious purposes' and hence Tamil might be described as a 'contact language' for the Sinhalese after the tenth century.[4]

Coperahewa is of the view that in addition to being the language of power and prestige, Tamil played a key role in trade and business along the Indian and Sri Lankan coasts and was a main language used in commercial communication. He further notes that geo-linguistically, Dravidian languages became the closest languages to Sinhala. Being confined to Sri Lanka, separated by the Dravidian belt, Sinhala has lost contact with its genetic relatives, the other NIA languages. This geographical proximity of Sinhala to Dravidian languages, especially to Tamil, according to Coperahewa (2007: 141), has led to cultural contact between the Sinhalese and Tamil communities for many centuries, which results in these two communities—their ethnic differences notwithstanding—having 'similarities in social customs, manners, practices, beliefs and linguistic patterns.'

From the point of view of literature, Tamil was treated on a par with Sanskrit and Pali as an important literary language of the day by the Sinhala Poets of the 15th and 16th centuries who considered those who were not versed in these languages to be 'ignorant' (Coperahewa 2007: 141). Moreover, there is evidence in Sinhala literary

---

[4] Here, in what sense Coperahewa (2007) refers to Tamil as 'contact language'—(i) a language which Sinhala was in contact with or (ii) a local lingua franca—is not known.

works that Tamil language and its literature were studied in colleges of medieval Sri Lanka. Works like Godakumbura (1946), discusses the Tamil literary trends in Sinhala literature. The factors discussed above account for the ‘huge influx of Tamil vocabulary into Sinhala, mainly connected with daily life, warfare, arts and administration’ (Coperahewa 2007: 141). Among the lexical items, kinship terms clearly indicate the Tamil influence on Sinhala lexicon. The most common kinship terms in the two languages are given in table 1.3:

Sinhala	**Tamil**	
amma		mother
akka		elder sister
mama		uncle/father-in-law
aachchi		grandmother
thaaththa/appachchi	*appa/aiya*	father
aiya	*anna*	elder brother
malli	*thambi*	younger brother
nangi	*thangachchi*	younger sister
nanda	*maamy*	aunt/mother-in-law
athamma	*amamma*	maternal grandmother
	amappa	paternal grandmother
aatha/seeya	*apamma*	maternal grandfather
	apappa	paternal grandfather

Table 1.3: Kinship terms

Note that though Tamil and Sinhala share only 4 kinship terms, the others in Sinhala are either derived from the Tamil ones, e.g. *appachchi* ‘father’ or used to refer to some other person, e.g., *aiya* to refer to ‘elder brother’ in Sinhala. Distinguishing words such as *akka* (elder sister), *amma* (mother), *attamma* (grandmother), *appachchi* (father), *aiya* (elder brother), *nangi* (younger sister) *mama* (uncle) as loans from Tamil, Coperahewa (2007: 145) points out that ‘they are so much part of Sinhala usage today that Sinhala speakers hardly realize that they are borrowings from Tamil.’

If Coperahewa’s (2007: 141) view that ‘a vast number of Tamil loan words’ found their way into Sinhala is correct, then ‘borrowing’, one of Thomason and Kaufman’s (1988: 37) two kinds of contact-induced changes (1.2), clearly explains the Sinhala-Tamil contact situation because ‘in a borrowing situation, the first foreign elements to enter the borrowing language are words… If there is long-term cultural pressure from the source-language speakers on the borrowing language-speaker group, then structural features may be borrowed as well … and even (though more rarely) features of the inflectional morphology.' Thus, what began with the borrowing of lexical items ended with vast scale morphosyntactic restructuring in Sinhala.

'Interference through imperfect learning', the other kind of contact-induced change, 'does not begin with vocabulary: it begins instead with sounds and syntax' (ibid. 39). Though rather unlikely, yet it is not totally impossible to describe the Sinhala-Tamil contact situation as interference through shift. Tamil speakers (comprising Tamils and members of the Muslim community that spoke Tamil), the minority group, aspired to establish ties with the Sinhalese, the majority group, and during their initial stages of contact may have learned Sinhala, but learned imperfectly. The resultant imperfect form of Sinhala, based on Tamil grammar but with Sinhala lexicon, a substrate form, may subsequently, perhaps due to a shift of power relations between the two communities, have become the standard form of Sinhala (see 1.2).[5] This view presupposes a rather complex pattern of language contact and shift and has in fact not been advocated anywhere in the literature. This dissertation, therefore, advances the view that the Sinhalese who learned Tamil, the then prestigious language, during the initial stages of contact, borrowed considerable number of words and subsequently restructured their grammar on the model of Tamil grammar, retaining their Sinhala lexicon.

A significant event in the history of Sri Lanka with regard to changes in the two languages is the colonial occupation of Sri Lanka by three nations, namely the Portuguese (1506–1658), Dutch (1658–1796) and English (1815–1948). Other than a limited set of lexical items from these three languages, which came into Sinhala and Tamil, none of these languages has induced any morphosyntactic changes in Sinhala or Tamil. A small community of Portuguese and Dutch descendants remained in Sri Lanka. They are called Burghers who are of mixed Portuguese/Dutch and Sri Lankan descent and speak a Portuguese creole. The only other language which Tamil and Sinhala have come into contact with was Malay, which came to Sri Lanka when Malay soldiers from Indonesia were posted in Sri Lanka by the Dutch while Indonesia was a Dutch colony. Like the Portuguese, the descendants of the Malays remained in Sri Lanka and continued to speak Malay. Once again only some Malay words entered Tamil and Sinhala, and Malay has not induced any morphosyntactic changes in either

---

[5] Sinhalese emerged as the majority community with considerably more social power than the Tamil speakers.

Sinhala or Tamil. Other than these four languages, Sinhala and Tamil in Sri Lanka have not come into contact with any other language. It may, therefore, be concluded that almost all the morphosyntactic changes in Sinhala had taken place prior to the arrival of the Portuguese in Sri Lanka and the morphosyntactic changes that Sinhala has undergone have been modelled on Tamil.

Although none of the four languages, Portuguese, Dutch, English and Malay have exerted influence on Sinhala or Tamil, the opposite has occurred; since Malay and Portuguese have been in contact with Sinhala and Tamil for more than three or four centuries in Sri Lanka, the former, referred to as Modern Sri Lankan Malay and Modern Sri Lankan Portuguese respectively, have undergone significant metatypy owing to their contact with Sinhala and Tamil. According to Bakker (2006: 135), 'The creolized and hence relatively isolating languages became agglutinative in probably just a few generations, and both created new forms for almost all the semantic categories expressed in the previously creolized local Indic and Dravidian verbs, nouns and discourse structure' (for details, see Bakker (2006) and the references therein).

Coperahewa (2007) also records that in the recent history of Sri Lanka, the importation of workers from South India for the tea plantation industry during the 19th century by the colonial rulers saw a large Tamil population from South India settling down in the Sinhalese-populated central part of Sri Lanka, creating complex Sinhala-Tamil contact situations. These contact situations 'paved the way for Sinhala-Tamil bilingualism among the Sinhalese in that area' (Coperahewa 2007: 140). It is assumed, along the lines of Ross (2003 and his other works mentioned above), that a bilingual Sinhalese community which used Tamil as the intercommunity language emerged. Note that bilingualism has been a common feature in Sri Lanka ever since the two speech communities came into contact with each other, always, or at least for long periods, and that the morphosyntactic changes that will be described and analyzed in the following chapters have come about as an effect of the bilingualism which has, as will be discussed below, led to changes in Sinhala modelled on Tamil, but not vice versa.

The lexicon of Sinhala which is argued to be of Sanskrit and Pali (Indo-Aryan) origin has by and large remained unchanged except for a considerable number of Tamil (see Coperahewa 2007) and English words, and also words from other languages found in Sinhala.[6] Though the two languages have distinct kinds of lexicon, Dravidian and Indo-Aryan, the majority of morphosyntactic features, as will be discussed in the dissertation, are similar. A detailed study of the wider context, including comparison of

---

[6] According to Coperahewa (2007), there are about 900 Tamil words in Sinhala usage.

Tamil and Sinhala with Dravidian and NIA languages more generally, as will be done in this dissertation, will show that the changes in the morphosyntactic features of Sinhala are due to influence of Tamil on Sinhala.

As for Tamil in Sri Lanka, it has not undergone any significant morphosyntactic changes on the model of Sinhala despite their long coexistence. However, there is one exception: one of the dialects of Tamil spoken by a small bilingual community in Negombo on the western coast of Sri Lanka 'has undergone significant amounts of borrowing and grammatical realignment under the influence of colloquial Sinhala, an unusual instance of South Asian contact-induced change in the direction from Dravidian to Indo-Aryan' (Bonta 2010: 310). The Sri Lankan Tamil lexicon has remained unchanged, except for some English words and words from other languages—which Tamil, like Sinhala, was in contact with—found their way into it. Evidence for words of Sinhala origin in Tamil has not been found. The changes, especially in morphosyntax, that have occurred in Sri Lankan Tamil have largely followed changes that have occurred in the Tamil language used in South India.

There are, however, phonological, lexical and subtle morphological differences between the Tamil spoken in South India and the one in Sri Lanka. The Tamil in South India, it is believed, has not undergone any significant changes, language contact-induced or otherwise, over the years. This may in part be due to the purist movement that emerged in South India to prevent any Sanskrit words from entering Tamil. This movement propagated a pure style of Tamil which 'is characterized by the avoidance of Sanskrit words, the use of indigenous Tamil words, the use of grammatical forms of Old and Middle Tamil, and in particular a rigorous application of morphophonemic (sandhi) rules' (Lehmann 1989: viii). This trend in the South Indian Tamil accounts at least in part for the absence of any significant changes in Sri Lankan Tamil.

Sociolinguistically, language dominance plays an instrumental role in changes which languages undergo due to contact. With respect to dominance owing to prestige or power that languages enjoy, language contact situations can be divided into two kinds: i) those in which the languages in contact are on a par with each other; and ii) those in which the languages are not on a par with each other. The two kinds—referred to as 'balanced and displacive' language contact respectively (Velupillai 2012)—take place in bi-/multilingual settings. In a balanced language contact situation, the bi-/ multilingualism is mutual in the sense that the speakers and the languages in contact influence each other and speakers of both or all languages are bi-/multilingual. Hence, the diffusion of features is multi-/bidirectional. An example is the balanced language contact in the Vaupés river basin area of Amazonia at the Colombian/Brazilian border.

In a displacive language contact situation, the influence tends to be one-way: only the speakers of the dominated language are bi-/multilingual, while the speakers of the dominating language are not. For example, the majority of Welsh speakers in Wales, as speakers of the dominated language, are bilingual in English, whereas a vast number of English speakers who are monolingual only in English, the dominating language, have no knowledge of Welsh. Instances of extreme dominance due to enforcement, such as laws against using minority languages or lack of socio-economic opportunities for the speakers of the dominated language, may ultimately lead to language loss or attrition where speakers end up shifting to the dominating language (cf. interference through shift).

The Sinhala-Tamil contact situation may be classified as the displacive kind for the following reasons:

(i) The contact and the changes induced by the contact have never been mutual or reciprocal, but have always been one sided. A comparison of Sri Lankan Tamil and Indian Tamil would reveal that Sri Lankan Tamil does not diverge in any major way from Indian Tamil, which shows that Sri Lankan Tamil has not undergone any restructuring under the influence of Sinhala except for the instance mentioned above.

(ii) In terms of power and therefore of dominance, Tamil and Sinhala are not on a par with each other. It is Tamil which was the dominating language during the early stages of contact between the two languages and has thus exerted influence on Sinhala (see Disanayaka 2007).[7]

Despite the numerous Tamil loan words in Sinhala, Sinhala evolved as a language with distinct NIA lexicon which is largely distinct from the Tamil/Dravidian lexicon, which implies that there may not have been coercion or extreme dominance over the Sinhalese by Tamil speakers.

Generally scholars account for the similarities between the morphosyntactic features of the two languages as being the result of ‘Dravidian influence on Sinhala’ (Gair 1998, 2007, Geiger 1937, Coates 1972 and Hettiaratchi 1959). Although the other Dravidian languages are geographically contiguous to Sinhala, and speakers of these

---

[7] The situation is diametrically opposite in the present day Sri Lanka in which the Sinhalese are the majority and Tamil speakers are the minority.

languages may have been in contact with the Sinhalese, it is Tamil with which Sinhala has been in direct contact. Evidence for influence of other Dravidian languages on Sinhala has not been found. The phrase ‘Dravidian influence on Sinhala,’ therefore, does not adequately describe the complex restructuring that a wide range of morphosyntactic phenomena in Sinhala have undergone, modelled on Tamil. One possible reason for the use of the epithet ‘Dravidian’ to denote Tamil influence on Sinhala is that Tamil is formally referred to as *dravida bashawa* ‘Dravidian language’ in Sinhala. More importantly, Disanayaka (1994: 3950) points out that “Owing to its contact with Tamil, Sinhala shares some linguistic features of Dravidian, thus providing an excellent example of linguistic convergence.” It is W. F. Gunawardhana’s (1924) claim cited in Disanayaka (2007: 4) that sums up the core of the issue:

> I found that some of the maian [sic] features of Sinhalese grammar are quite out of tune with their Aryan analogues, while all principles of fundamental importance harmonize with their Tamil equivalents … This forced on me the conclusion that the structural foundations of the Sinhalese language are Dravidian while its superstructure, i.e., the vocabulary is Aryan.

Disanayaka (2007), who cites Gunawardhana (1924: 4), notes that the above observation aroused a debate among scholars, which ‘came to an end proving that Sinhala is an Indo-Aryan language, of course, with many Dravidian elements that have found their way to Sinhala not due to genetic affinities but due to the language-in-contact situation.’

Scholars trace the origin of Sinhalese to Sanskrit and Pali: the former is the acclaimed mother of all Indo-Aryan languages, while the latter is the medium of the majority of Buddhist texts. They also acknowledge what they call the Dravidian influence on Sinhala. Their main intention, as it is evident from different works, is to assert that Sinhala is of Indo-Aryan origin. A case in point is Gair (2007: 848), who notes that ‘Though geographically isolated from all the mainland Indo-Aryan languages, Sinhala is without doubt an Indo-Aryan language… Consideration of the basic vocabulary, as well as a number of fundamental grammatical properties leave no doubt as to its membership in the IA [Indo-Aryan] family.’ This is the view that is disseminated within academia. It is, therefore, reasonable to assume that the present day Sinhalese, except for the information conveyed in works such as Coperahewa (2007), Disanayaka (2007), Gair (1985, 1998d , 2007), GnanaPrakasar (1937) among others, may not know the extent to which the core morphosyntax of Sinhala has undergone restructuring on the model of Tamil.

There are works on diverse aspects of Tamil/Dravidian in Sinhala available in the literature. J. W. Gair is the most prolific and well-known among writers on Sinhala language and linguistics. His work (1985) examines the influence of Dravidian phonology on Sinhala. He has also written on a variety of topics in Sinhala grammar (see references). GnanaPrakasar (1937) traces the morphological and etymological features of Tamil/Dravidian origin in the lexical items in Sinhala.[8] In the literature reviewed, there is no comprehensive comparative study of the morphosyntactic features of Sinhala and Tamil, illustrating these features with real data from the two languages by juxtaposing them so as to show the similarities/differences between the features, and thereby the restructuring of the Sinhala features on the model of the corresponding Tamil features, as done in this research. Gair in many of his works notes the dearth of a comparative study of the two languages in the literature. It is believed that the current research will fill this major gap in the domain of morphosyntax of Tamil and Sinhala.

1.6 Methodology

Since this research attempts to study the nature of the contact-induced restructuring of the morphosyntactic features that modern spoken Sinhala has undergone on the model of modern spoken Tamil, the method adopted in this research is to compare typological features of the two languages by illustrating them with natural data and examine them to study the nature of restructuring that has taken place. To this end, a wide range of morphosyntactic features of these two languages are examined. The comparative analysis is at macro- and micro-levels. At macro-level, the morphosyntactic features of the two languages are compared in terms of the typological features related to morphosyntax listed in Dryer and Haspelmath (2013). Included also for comparison are certain features that are not included in Dryer and Haspelmath (2013). At micro-level, two morphosyntactic phenomena in the two languages are examined.

Dryer and Haspelmath (2013; WALS hereafter) provides a viable criteria on the basis of which comparison of these features can be made. WALS 'shows […] structural features, i.e., abstract features of the language system that can be compared across unrelated languages' (Haspelmath et al. 2005a: 1). It distinguishes cross-linguistically 144 typological features, and these features are divided into twelve sections. These sections are further divided into subsections, covering the most important subdomains of these structural domains (see ibid. for details). Determined for each feature is a set of values which are the diverse ways in which a feature is manifested in languages. Each

---

[8] The current research does not include phonological and lexical features.

feature is described comprehensively in what is referred to as chapters, specifying the criteria used to formulate the values and illustrating each value with examples from the languages which possess the respective feature value. The study of feature values is undertaken in samples comprising a considerable number of languages. The exact number varies depending on the feature that is investigated, but ranges approximately from 100 to 1500. WALS has data from 2679 languages, at the last count, and provides the feature values of these languages determined ‘primarily from previously-published descriptions of each individual language, typically in the form of a reference grammar’ (Haspelmath et al. 2005a: 2). The geographical distribution of the feature values for each feature is displayed on maps.

WALS affords a remarkable resource for this research in that the typological features it makes available provide effective criteria for the comparison of diverse morphosyntactic features of Sinhala and Tamil.[9] In this research, 51 WALS typological features 81A–144S (excluding 129A–142A) related to morphosyntax and 13 other features are selected for comparison.[10] The 64 features compared in Chapters 2–4 are given in table 1.4:[11]

---

[9] There are, however, some limitations in WALS (see Holmberg (to appear); Koptjeveskaja-Tamm (2011)), which are not relevant to the research. Therefore, they are not discussed owing to constraints of space.

[10] WALS features 129A–142A which are divided into three sections, namely Lexicon (129A–138A); Sign Languages (139A–140A) and Other (Writing Systems and Paralinguistic Usages of Clicks) (141A – 142A) are non-morphosyntactic; hence, these features are not included in this research.

[11] For consistency and ease of reference, the typological features are indicated using ‘WALS and the relevant number given to each feature, e.g., WALS 81, ‘Order of Subject, Object and Verb’.

Chapter	Sections	WALS No.	No. of features	
2	Word order	81A–97A		17
3	Simple clauses i) Case and agreement	98A–105A	08	
	ii) Valence and voice	106A–111A	06	
	iii) Negation and questions	112A–116A, 143A, E & 144A	08	
	iv) Predication	117A–121A	05	
				27
4	Complex sentences	122A–128A		07
	Miscellaneous			13
	Total			64

Table 1.4: Morphosyntactic features

Chapters 2, 3 and 4 begin with a section in which the information necessary for an understanding of the features compared in that chapter is discussed; this is followed by a section in which the relevant features of the two languages are briefly described, illustrated with examples with the aim of comparing them; in the next section, based on the findings from the comparison of the typological features, the typological profiles of the two languages are produced. Also discussed are the implications—which arise from the comparative analysis—for the possible contact-induced changes that Sinhala may have undergone. The final section summarizes the findings in each chapter.

The similarities and differences between the features of the two languages identified from these typological profiles are examined to determine whether the features of Sinhala are NIA, areal or those adopted from Tamil. To this end, the 51 feature values of Tamil and Sinhala are compared with the corresponding feature values of two Dravidian languages, namely Malayalam and Telugu and five NIA languages, namely Marathi, Gujarati, Oriya, Bengali and Hindi; the feature values of these languages are given in the Appendix.[12] 13 features from diverse domains of morphosyntax which are not included in the 51 features are also analyzed and, where possible, compared with the respective features of other Dravidian and NIA languages. Further, any other features relevant to those compared in that chapter are also examined.

To determine the feature values and formulate the data from the two languages used in the research, I used my intuition as a native speaker of Tamil and my knowledge

---

[12] The feature values of Tamil and Sinhala are those determined via research, whereas the feature values of the other seven languages are those available in WALS (Dryer and Haspelmath 2013).

of Sinhala. Also used for these two purposes are descriptive grammars available on the two languages. In instances where the feature values determined in the three chapters are different from those of Tamil and Sinhala available in WALS, the differences/ anomalies are accounted for.

With regard to the scope of this research, the following facts need to be stated at the outset:

(i) this research, as mentioned in 1.1, is a comparative synchronic study of the morphosyntactic features of Tamil and Sinhala, and not a diachronic study which is beyond the scope of this research owing to the long history of the two languages (over two millennia). It is hoped that this synchronic study will provide a comprehensive understanding of the contact-induced restructuring of Sinhala because the present morphosyntactic features of Sinhala have resulted from the changes that have occurred over time.

(ii) only the default forms of the morphosyntactic features of the two languages are compared; these exclude forms which encode other ways of expressing these features in the two languages. However, where deemed necessary, the other possible forms are also discussed.

WALS features related to word order, simple clauses, and complex sentences are compared in Chapters 2, 3 and 4 respectively (see table 4.1). Also compared in Chapter 4 are 13 other features not included in WALS. In the two remaining Chapters (5 and 6), two important morphosyntactic phenomena in the two languages, namely the occurrence of null arguments (Chapter 5) and focusing (Chapter 6) are analyzed in detail. Subject/ Verb (SV) agreement is a major difference between the two languages. In many languages which allow null arguments, especially null subjects, agreement is claimed to play an important role. Tamil and Sinhala allow null arguments under the same contextual conditions despite the difference in SV agreement between them. In Chapter 5, null arguments in the two languages are examined on the basis of recent studies on null arguments to determine i. whether agreement plays any role in their occurrence; ii. what kind of null argument languages Tamil and Sinhala are; and iii. the conditions under which they allow null subjects. Chapter 6 examines the two kinds of focus—cleft and focus particle—constructions which assign exhaustive focus to a constituent in a sentence. These constructions in Sinhala provide clear evidence for contact-induced restructuring/replica grammaticalization modelled on the corresponding constructions in Tamil. These two focus constructions which Sinhala replicated from Tamil triggered changes in other morphosyntactic phenomena in Sinhala such as SV agreement and, polar and wh-questions which are also examined in this chapter. Chapter 7 is the conclusion of the dissertation.

Chapter 2

Features related to word order

2.1 Preliminaries

In this chapter, word-order related features of Tamil and Sinhala are compared to identify the similarities and differences between these features of the two languages. The aim is to develop the typological profiles of the features related to word order, which will help establish the restructuring of Sinhala on the model of Tamil, and to examine the morphosyntactic changes that the features of Sinhala have undergone. This chapter compares 17 WALS features (81A–97A) which belong to the section titled 'Word Order' in Dryer and Haspelmath (2013; see section 'F' in Haspelmath et al. 2005). These 17 features deal with the order of diverse constituents on both clause/sentence and phrase levels.

Basically word or constituent order is the order in which the three main constituents of a sentence—subject, object and verb—occur in a language. The use of the terms 'subject' and 'object' in a sentence is determined vis-à-vis the three notions: A(gent) and P(atient) are the 'more agent-like' and 'more patient-like' arguments respectively in a transitive clause, while S(ubject) is the single argument in an intransitive clause (Dryer 2013a). The transitive clause has a transitive verb which can take a direct object, whereas the intransitive clause has an intransitive verb which cannot take a direct object. The first two features are about the basic word order of a language, that is, the order in which the transitive/intransitive verb and its arguments are arranged in a declarative sentence. The other 15 features are concerned with the order of other constituents in declarative and interrogative sentences.[13]

Section 2.2 compares the 17 features related to word-order in the two languages. Each of these features and their values are described briefly and the feature values of Tamil and Sinhala are illustrated with examples to determine whether that feature is similar or different in the two languages. In Section 2.3 titled Discussion, the main findings of the comparative study in Section 2.2 are used to establish the typological profiles of the word order related features, and the implications arising from the findings are discussed with the aim of studying the nature of the changes in these features of Sinhala induced by its contact with Tamil.

---

[13] In 2.2 below and the sections of Chapters 3 and 4 below, only the relevant features are compared; all the other details related to the features are discussed in the section titled Discussion in these chapters.

2.2 Features related to word order: a comparison

2.2.1 Order of subject, object, and verb

This feature (WALS 81A, Dryer 2013a) is about the ordering of subject, object and verb in a transitive clause, more specifically declarative clauses in which subject and object involve a noun (and not just a pronoun). The seven values of this feature are: 1. SOV; 2. SVO; 3. VSO; 4. VOS; 5. OVS; 6. OSV; and 7. those which 'lack a dominant order.[14] The dominant order of Tamil and Sinhala is SOV (value 1), as in (1) below (alternative orders are possible here and in features compared in 2.2.2–2.2.4 below; for examples, see 2.3):[15]

(1) T. kumar oru paattu paad-in-aan
Kumar one(INDF) song sing-PST-3SGM
S. kumar sinduw-ak kiwwa
Kumar song-INDF tell.PST

'Kumar sang a song.'

2.2.2 Order of subject and verb

This feature (WALS 82A, Dryer 2013b) is about the dominant order of lexical (or non-pronominal) subject and verb. There are three values for this feature: 1. subject precedes verb (SV); 2. subject follows verb (VS); and 3. both orders with neither dominant. In Tamil and Sinhala the dominant order is SV (value 1), as in (2) below:

(2) T. rani siri-th-aal
Rani laugh-PST-3SGF
S. rani hinawuna
Rani laugh.PST

'Rani laughed.'

2.2.3 Order of object and verb

This feature (WALS 83A, Dryer 2013c) refers to the order of lexical (non-pronominal) object and verb in a transitive clause. The three values of this feature are: 1. object

---

[14] In comparing the morphosyntactic features here and below (Chapters 2–4) only the information necessary to illustrate the features of Tamil and Sinhala is discussed (for further information about features, see the work cited in the subsection on each of these features).

[15] Note that in (1T) the subject of the sentence agrees with the verb, whereas in (1S) the subject does not agree with the verb (for agreement features in the two languages, see 3.6.1.2).

precedes verb (OV); 2. object follows verb (VO); and 3. both orders with neither dominant. The dominant order of object and verb in Tamil and Sinhala is OV (value 1), as in (3) below:

(3) T. avan pingaan-ai kaluv-in-aan
He plate-ACC wash-PST-3SGM
S. eyaa pingaanə heeduwa
He plate wash.PST

‘He washed the plate.’

2.2.4 *Order of object, oblique, and verb*

This feature (WALS 84A, Dryer and Gensler 2013) is concerned with the dominant order of lexical object, oblique phrase and verb. An oblique phrase is a noun phrase or adpositional (prepositional or postpositional) phrase, excluding indirect objects (recipients with ditransitive verbs); they function as an adverbial modifier (or “adjunct”) of the verb (indicated by ‘X’). Of the six values of this feature—1. VOX; 2. XVO; 3. XOV; 4. OXV; 5. OVX; and 6. more than one order with none dominant—Tamil and Sinhala have XOV (value 3) as their dominant order, as in (4) below:

(4) T. kamala penn-al kakitha-(thth)ai thira-nth-aal
Kamala pen-INS letter-ACC open-PST-3SGF
S. kamala pææn-en liumə kæduwa
Kamala pen-INS letter break.PST

‘Kamala opened the letter with a pen.’

2.2.5 *Order of adposition and noun phrase*

This feature (WALS 85A, Dryer 2013d) is about the order of adposition and noun phrase. An adposition may either be a word or an affix (often marking case on nouns) which combines with a noun phrase and indicates a syntactic or semantic relationship of that noun phrase to the verb in the clause. There are five values for this feature: 1. Postpositions; 2. Prepositions; 3. Inpositions; 4. More than one adposition type with one dominant; and 5. No adpositions. Tamil and Sinhala have postpositions (value 1), as in (5) below:[16]

---

[16] The second and third person singular finite verbs are also marked for honorificity in Tamil, glossed as HON stated within brackets, to show respect, whereas the finite verbs in modern spoken Sinhala are not marked for honorificity.

(5) T. amma doctor-itta poo-n-aa
Mother doctor-LOC go-PST-3SGF(HON)
S. amma dostorə-gaavə-tə giya
Mother doctor-LOC-DAT go.PST

‘Mother went to (see) the doctor.’

2.2.6 Order of genitive and noun

This feature (WALS 86A, Dryer 2013e) refers to the order of a genitive or ‘possessor’ noun phrase with respect to the head or ‘possessee’ noun. The construction itself is known either as a genitive construction or a possessive construction. The three values of this feature are: 1. genitive-noun (GenN); 2. noun-genitive (NGen); and 3. both orders with neither order dominant. Tamil and Sinhala have GenN (value 1) as their order, as in (6) below:

(6) T. mala-inda puthakam
Mala-GEN book(N)
S. mala-ge pothə
Mala-GEN book(N)

‘Mala’s book’

2.2.7 Order of adjective and noun

This feature (WALS 87A, Dryer 2013f) is about the distribution of the two possible orders of modifying adjective and noun. The four values of this feature are: 1. modifying adjectives precede noun (AdjN); 2. modifying adjectives follow noun (NAdj); 3. both orders of noun and modifying adjective, with neither dominant; and 4. adjectives do not modify nouns, occurring as predicates in internally-headed relative clauses, that is, the adjective is the predicate in an internally-headed relative clause. Tamil and Sinhala have AdjN (value 1) as their order, as in (7) below:

(7) T. sivappu poo
Red(ADJ) flower(N)
S. rathu malə
Red(ADJ) flower(N)

‘Red flower’

2.2.8 Order of demonstrative and noun

This feature (WALS 88A, Dryer 2013g) is about the order of demonstrative and noun. Of the six values of this feature —1. demonstrative word precedes noun (DemN); 2. demonstrative word follows noun (NDem); 3. demonstrative prefix on noun; 4. demonstrative suffix on noun; 5. demonstrative simultaneously before and after noun; and 6. two or more of above types with none dominant—Tamil and Sinhala have DemN (value 1) as their order, as in (8) below:

(8) T. antha veedu
That(DEM) house(N)
S. ee gee
That(DEM) house(N)

‘That house’

2.2.9 Order of numeral and noun

This feature (WALS 89, Dryer 2013h) is concerned with the order of cardinal numerals with respect to nouns they modify. There are four values for this feature: 1. numeral precedes noun (NumN); 2. numeral follows noun (NNum); 3. both orders of numeral and noun with neither order dominant; and 4. numeral only modifies verb. The order of cardinal numerals and the nouns they modify in Tamil is different from the order in Sinhala in that Tamil has NumN (value 1), whereas Sinhala has NNum (value 2), as in (9) below:

(9) T. moondu pallikoodan-gal
Three school-PL
S. iskoolə thun-ak
School.PL three-INDF

‘Three schools’

2.2.10 Order of relative clause and noun

This feature (WALS 90A, Dryer 2013i) is about the order of a relative clause and the noun it modifies. There are seven values for this feature: 1. relative clause follows noun (NRel); 2. relative clause precedes noun (RelN); 3. internally-headed relative clause; 4. correlative relative clause; 5. adjoined relative clause; 6. double-headed relative clause; and 7. mixed types of relative clause with none dominant. The order of relative clause and the head noun in Tamil and Sinhala is RelN (value 2), as in (10) below:

(10) T. [rani _____ asin-a] manushan nettru inge va-nth-aan
Rani scold.PTCP-RELAT man yesterday here come-PST-3SGM
S. [rani_____ bænəp-u] miniha iiyee mehe aawa
Rani scold.PTCP-RELAT man yesterday here come.PST

‘The man who Rani scolded came here yesterday.’

2.2.11 Order of degree word and adjective

This feature (WALS 91A, Dryer 2013j) is concerned with ‘the position of degree words with respect to the adjective that they modify.’ For the purposes of this feature, ‘degree words’ are defined as words with meanings like ‘very’, ‘more’, or ‘a little’ ‘that modify the adjective to indicate the degree to which the property denoted by the adjective obtains.’ The three values of this feature are: 1. degree word precedes adjective (DegAdj); 2. degree word follows adjective (AdjDeg); and 3. both orders occur with neither order dominant. Tamil and Sinhala have DegAdj (value 1) as their order, as in (11):

(11) T. koncham oothai
A little (DEG) dirty(ADJ)
S. tikak kilutui
A little (DEG) dirty(ADJ)

‘A little dirty’

2.2.12 Position of polar question particles

This feature (WALS 92A, Dryer 2013k) refers to the position of question particles in questions which elicit a ‘yes’ or ‘no’ response. There are six values for this feature: 1. question particle at beginning of sentence; 2. question particle at end of sentence; 3. question particle in second position in sentence; 4. question particle with other position; 5. question particle in either of two positions; and 6. no question particle. In terms of the criteria discussed above, Tamil has no question particle (value 6); it uses the polar question clitic *-aa* instead, whereas Sinhala uses a common question particle *də* (value 3), as in (12b). Note that the declarative forms of (12b) are given in (12a):

(12) a. T. nandini poona kilamai pillaikal-(u)kku inglish padipp-ich-aa
Nandini last week students-DAT English teach-PST-3SGF(HON)
S. nandini giyə sumane lamay-tə ingriisi igænnuwa
Nandini last week students-DAT English teach.PST

‘Nandini taught English to the students last week.’

(12) b. T. nandini poona kilamai pillaikal-(u)kku inglish padipp-ich-aa**-(v)aa**
Nandini last week students-DAT English teach-PST-3SGF(HON)-Q
S. nandini giyə sumane lamay-tə ingriisi igænnuwa **də**
Nandini last week students-DAT English teach.PST Q

'Did Nandini teach English to the students last week?'

2.2.13 Position of interrogative phrases in content questions

This feature (WALS 93A, Dryer 2013l) is about the position of interrogative phrases in content questions which are those that contain an interrogative phrase such as 'who', 'what' etc, eliciting a specific answer other than 'yes' or 'no'. The three values of this feature are 1. interrogative phrases obligatorily initial; 2. interrogative phrases not obligatorily initial; and 3. mixed, some interrogative phrases obligatorily initial; some not. The interrogative phrases in Tamil and Sinhala are not obligatorily initial (value 2), as in (13) below (for convenience, the questions below are formed from (12a) above):

(13) T. nandini poona kilamai pillaikal-(u)kku enna padipp-ich-aa
Nandini last week students-DAT what teach-PST-3SGF(HON)
S. nandini giyə sumane lamay-tə mokak də igænnuw-e
Nandini last week students-DAT what Q teach.PST-NMLZ

'What did Nandini teach the students last week?'

2.2.14 Order of adverbial subordinator and clause

This feature (WALS 94A, Dryer 2013m) is concerned with the position of adverbial subordinators (separate words referred to by traditional grammar as 'subordinating conjunctions', e.g., *because, although, when, while* and *if* in English) and morphemes which mark adverbial clauses for their semantic relationship to the main clause. There are five values for this feature: 1. adverbial subordinators which are separate words and which appear at the beginning of the subordinate clause; 2. adverbial subordinators which are separate words and which appear at the end of the subordinate clause; 3. clause-internal adverbial subordinators; 4. suffixal adverbial subordinators; and 5. more than one type of adverbial subordinators with none dominant. Tamil and Sinhala have suffixal adverbial subordinators (value 4) which appear at the end of the subordinate clause, i.e., clause final, as in (14):[17]

---

[17] Note that Ø used in this dissertation stands for null constituent, that is, a constituent which is not phonologically realized.

(14) T. appa town-(u)kku poon-aal Ø bank-il kaasu edu-pp-aar
Father town-DAT go.PTCP-if bank-LOC money take-FUT-3SGM(HON)
S. thaaththa towmə-tə giy-oth, Ø banku-(w)en salli ganiyi
Father town-DAT go.PTCP-if bank-ABL money take.FUT

'If father goes to town, he will get money from the bank.'

WALS 95A, 96A, and 97A (2.2.15–2.2.17 below) deal with the relationship between the values of two typological features which intersect to define four possible values (1–4) in each of the three features. The fifth value of these three features is the language type which does not belong to—i.e., is different from—the other four types.

2.2.15 Relationship between the order of object and verb and the order of adposition and noun phrase

The two features which intersect in this feature (WALS 95A, Dryer 2013n) are order of object and verb (2.2.3) and order of adposition and noun phrase (2.2.5). The five values of this feature are: 1. object-verb and postpositional (OV&Postp); 2. object-verb and prepositional (OV&Prep); 3. verb-object and postpositional (VO&Postp); 4. verb-object and prepositional (VO&Prep); and 5. languages not falling into one of the preceding four types. Of the four types which result from the intersection of the two values each of the two features, types 1 and 4 are very common, while types 2 and 3 are relatively uncommon. Tamil and Sinhala are consistently OV&Postp (value 1), as in (15) below:

(15) T. leela nimal-odu padam paar-th-aal
Leela Nimal-COM film watch-PST-3SGF
S. leela nimal ekkə chithrəpatəyə bæluwa
Leela Nimal COM film watch.PST

'Leela watched the film with Nimal.'

2.2.16 Relationship between the order of object and verb and the order of relative clause and noun

The two features which intersect in this feature (WALS 96A, Dryer 2013o) are order of object and verb (2.2.3) and order of relative clause and noun (2.2.10). The five values of this feature are: 1. object-verb and relative clause-noun (OV&RelN); 2. object-verb and noun-relative clause (OV&NRel); 3. verb-object and relative clause-noun (VO&RelN); 4. verb-object and noun-relative clause (VO&NRel); and 5. languages not falling into one of the preceding four types. Of the four types which result from the intersection of the two values each of the two features, types 1, 2 and 4 are common, while type 3 is uncommon. Tamil and Sinhala are consistently OV&RelN (value 1), as in (16):

(16) T. latha ravi kudu-thth-a sattai-ai poo-t-aal
Latha Ravi give.PTCP-RELAT frock-ACC put-PST-3SGF
S. latha ravi diip-u gaumə ænda
Latha Ravi give.PTCP-RELAT frock wear.PST

‘Latha wore the frock that Ravi gave.’

2.2.17 Relationship between the order of object and verb and the order of adjective and noun

The two features which intersect in this feature (WALS 97A, Dryer 2013p) are order of object and verb (2.2.3) and order of adjective and noun (2.2.7). The five values of this feature are: 1. object-verb and adjective-noun (OV&AdjN); 2. object-verb and noun-adjective (OV&NAdj); 3. verb-object and adjective-noun (VO&AdjN); 4. verb-object and noun-adjective (VO&NAdj); and 5. languages not falling into one of the preceding four types. Of the four types which result from the intersection of the two values each of the two features, all types (1–4) are common. Tamil and Sinhala are consistently OV&AdjN (value 1), as in (17):

(17) T. ravi oru karuppu kaar vang-in-aan
Ravi one(INDF) black car buy-PST-3SGM
S. ravi kalu kaar ekak gaththa
Ravi black car one(INDF) buy.PST

‘Ravi bought a black car.’

The 17 features compared above are the features classified under word order in Dryer and Haspelmath (2013). In 2.3 below, the features illustrated in 2.2 and others relevant to them are examined to show the features that have undergone restructuring in Sinhala owing to its contact with Tamil.

2.3 Discussion

In this chapter 17 features related to word order in Tamil and Sinhala are compared (2.2). The implications for the possible contact-induced changes that Sinhala may have undergone which arise from the comparative analysis are discussed below.

2.3.1 Features related to word order: implications of the comparison

The typological profiles established on the basis of the comparison of the values of the 17 word order-related features of Tamil and Sinhala are given in table 2.1:[18]

---

[18] For consistency, here and below (Chapters 2–4.2) the feature values in the typological profiles are stated as in Dryer and Haspelmath (2013).

	WALS features	**Tamil(T)**	**Sinhala(S)**	**Similar/ Different**
2.2.1	Order of subject, object, and verb (81A)	SOV	SOV	similar
2.2.2	Order of subject and verb (82A)	SV	SV	similar
2.2.3	Order of object and verb (83A)	OV	OV	similar
2.2.4	Order of object, oblique, and Verb (84A)	XOV	XOV	similar
2.2.5	Order of adposition and noun phrase (85A)	postpositions	postpositions	similar
2.2.6	Order of genitive and noun (86A)	genitive-noun	genitive-noun	similar
2.2.7	Order of adjective and noun (87A)	adjective-noun	adjective-noun	similar
2.2.8	Order of demonstrative and noun (88A)	demonstrative-noun	demonstrative-noun	similar
2.2.9	Order of numeral and noun (89A)	numeral-noun	noun-numeral	different
2.2.10	Order of relative clause and noun (90A)	relative clause-noun	relative clause-noun	similar
2.2.11	Order of Degree Word and Adjective (91A)	degree word-adjective	degree word-adjective	similar
2.2.12	Position of polar Question Particles (92A)	final (a clitic)	final	different
2.2.13	Position of interrogative phrases in content questions (93A)	not obligatorily initial interrogative phrase	not obligatorily initial interrogative phrase	similar
2.2.14	Order of adverbial subordinator and clause (94A)	suffixal adverbial subordinators	suffixal adverbial subordinators	similar
2.2.15	Relationship between the order of object and verb and the order of adposition and noun phrase (95A)	OV & postpositions	OV & postpositions	similar
2.2.16	Relationship between the order of object and verb and the order of relative clause and noun (96A)	OV & RelN	OV & RelN	similar
2.2.17	Relationship between the order of object and verb and the order of adjective and noun (97A)	OV & AdjN	OV & AdjN	similar

Table 2.1: Typological profiles of features related to word order

Thomason (2001b) points out that contact-induced change can have a profound effect on the typological profile of the receiving language. She adds that the most obvious examples and also the ones that are easiest to find are changes in basic sentential word order. The above typological profiles of Tamil and Sinhala related to word-order features show that except for two features that are different, all the feature values in Tamil and Sinhala are similar, that is, Sinhala shares with Tamil 15 of the 17 features compared, which indicates that Sinhala converges towards Tamil on a significantly high number of word-order related features.

These features should not, however, be quantitatively used as indicators of Sinhala's convergence with Tamil as the direct consequence of contact between the languages because the seven languages compared with the two languages, too, have the same feature values (see Appendix). It is important to note that Tamil and Sinhala and, by extension, the other languages of South Asia whose dominant word order is SOV share many of the 17 features listed above because SOV word order is associated with certain universal tendencies as Greenberg's (1963) word order universals (see 1.4) stipulate. The word order universals (Greenberg 1963) are used to show below that as a result of their SOV word order, Tamil and Sinhala together with seven other SOV languages (see Appendix) share certain distinct features compared above:

(i) Since Tamil and Sinhala together with other languages of South Asia have SOV order (WALS 81A; 2.2.1), the order of subject and verb (WALS 82A; 2.2.2) is by default SV.

(ii) On the basis of Greenbergian Universal (GU) 4 ('With overwhelmingly greater-than chance frequency … languages with normal SOV order are postpositional'), it is expected that the value of the order of adposition and noun phrase (WALS 85A; 2.2.5) in Tamil and Sinhala is postposition.

(iii) The conclusion of (ii) above together with GU 2 ('…in languages with postpositions [the genitive] almost always precedes [the governing noun]') predicts that the value of the order of genitive and noun (WALS 86A; 2.2.6) is genitive-noun (GenN) and the value of the order of adjective and noun (WALS 87A; 2.2.7), by implication, is 'modifying adjectives precedes noun (AdjN)' (contra GU 5 'If a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun.').

As shown by Dryer (1988), the correlation between object-verb order and adjective-noun order is, in fact, not universal, but an areal feature, where adjective-noun order is an areal feature of most of Eurasia. Dryer (1992) also shows that the order of demonstrative and noun does not correlate with OV order, on a universal basis. Still, even if adjective-noun order and demonstrative-noun order are disregarded, the above discussion indicates that the set of word order features shared by Tamil and Sinhala can be largely explained as a result of universal tendencies, and, therefore, cannot be used as evidence of convergence due to contact between these two languages.

What is important to note is that the fact that languages of South Asia have predominantly SOV type word order with mainly head-final phrases—either adopted via contact or inherited from a common ancestor as languages of a family or acquired by coincidence—may have facilitated the restructuring of other features owing to the intense contact among these languages of the region. As Thomason (2001a) observes, borrowing is facilitated if the languages are typologically similar or congruent to each other (also Koptjevskaja-Tamm 2011). That Tamil and Sinhala belong to the same Greenbergian type would thus have facilitated the restructuring of Sinhala on the model of Tamil. A brief introduction to the SOV properties of Tamil and Sinhala is in order. Subbarao (2010: 2) enumerates the following 'word order universals pertaining to verb-final (SOV) languages in sentences with their unmarked word order': i) it has postpositions (2.2.5); ii) the auxiliary verb follows the main verb (4.3.4); iii) indirect object precedes direct object ((12a) in 2.2.12); iii) the genitive precedes the governing noun (2.2.6); iv) time adverbs precede place adverbs; and v) time and place adverbials occur in a descending order, that is, the superordinate segment, e.g., when stating dates, the year precedes month which precedes date or when referring to a place, the continent/country precedes province/district which precedes town/village etc. The features (i–v) hold true for Tamil and Sinhala.

Herring and Paolillo (1995: 166) rightly point out that Tamil and Sinhala are 'harmoniously head-final'. These head-final languages have a left-branching structure in which modifiers/adjuncts and/or complements linearly precede its head which is modified or complemented. This head final structure determines the order of most of the other constituents.[19] The two languages are also classified as configurational languages with a relatively free word order. Although in Tamil and Sinhala the verb, by default, occurs at the end of the clause, as in (18a), since they display free word order (or

---

[19] The order of noun and numeral in Sinhala is an exception in that the head precedes the numeral (see below for discussion).

scrambling), other word orders (18b–f) are also possible (see Sarma 2003 for Tamil and Gair 1998c, 2007 for Sinhala):[20]

(18) T. a. ravi oru veedu vaangi-n-aan
Ravi one(INDF) house buy-PST-3SGM
'Ravi bought a house.'

b. oru veedu ravi vaanginaan
c. oru veedu vaanginaan ravi
d. ravi vaanginaan oru veedu
e. vaanginaan ravi oru veedu
f. vaanginaan oru veedu ravi

S. a. ravi gey-ak gaththa
Ravi house-INDF buy.PST
'Ravi bought a house.'

b. geyak ravi gaththa
c. geyak gaththa ravi
d. ravi gaththa geyak
e. gaththa ravi geyak
f. gaththa geyak ravi

Of the six orders in (18) in Tamil and Sinhala, (a,b), are more common than (c–f). Claiming that 'all orders of the major constituents' are possible in Sinhala, Gair (2007: 871) notes that 'there will be differences in topic, emphasis, afterthought etc. accompanied by intonation in these orders.' This claim holds true for Tamil too (Sarma 2003).

Though Sinhala is claimed to be harmoniously head-final, like Tamil, there are a few features in Sinhala which are not typical of the other head-final languages of the region. 2.3.1.1 below discusses facts in support of this claim.

2.3.1.1 Other word order issues: numerals, indefinite marker and negative marker

One of the 'structural effects of contact-induced grammaticalization on the replica language' which Heine and Kuteva (2005: 124) propose is 'coexistence' where there has been some equivalent grammatical category in the replica language and, as a result of contact, the inherited (old) and innovated (new) structures encoding this category coexist side by side. According to Heine and Kuteva (2005: 130) coexistence may take

---

[20] From the perspective of generative linguistics, it may be assumed that free word order is a consequence of greater freedom of movement, including scrambling, that is, movement of arguments in the TP/IP domain.

two different forms: (i) the new and the old categories are combined and co-occur in the same construction, thus resulting in double marking; (ii) the two coexist as alternative constructions available to speakers of the replica language. I will illustrate these two forms of coexistence in this section. Lehmann (1978) and Chandralal (2010) (the latter on the basis of the arguments developed by the former) claim that Sinhala, which at one time included a number of head-initial characteristics comparable to classical Sanskrit and Middle Indic dialects like Pali, has subsequently become a highly consistent head-final language through contact with Dravidian languages, as have the NIA languages in general.

One area in which a presumed head-initial feature is still found in Sinhala is what Lehmann (1978) calls the teen numerals (eleven to nineteen), most of which have adopted the Dravidian head-final order in which the decimal number (modifier/dependent) precedes the respective single digit (head) in keeping with the head-final order, e.g., ‘thirteen’ is *daha-thunə* = ten-three. However, eleven, twelve and fifteen which Chandralal (2010) refers to as ‘aberrant numerals’ follow the head-initial pattern of the single digit preceding the decimal, e.g., ‘eleven’ is *ekolaha* = one-ten. These three numerals, as Chandralal (2010) points out, are the remnants of the old Sinhala system which is a mixture of head-final and head-initial features. This mixture of two kinds in the numeral system in Sinhala indicates the second form of coexistence discussed at the beginning of this section. The retention of some of the numerals which follow the old head-initial pattern would be one instance in which the VO/head-initial feature surfaces in Sinhala.

Two other instances involving quantification where Sinhala differs from Tamil is the order of numeral and noun (2.2.9) and indefinite marker and noun. The order of numeral and noun in Tamil is NumN (9T), whereas in Sinhala it is NNum (9S).[21] The order of numeral and noun is one of the two differences between the word-order related features of the two languages. Simson and Syed (to appear: 3) note that in Bengali, the N(oun) P(hrase) in a Q(uantifier) P(hrase) which contain numerals can undergo ‘nominal-internal movement’ over the numerals to encode a definite reading to the QP: ‘this movement is possible over low numerals [1–4], but may not take place when higher numerals are present’. However, in Sinhala all numerals whether lower or higher follow the nouns, as in (19S) (see also (9S):

---

[21] The feature value of numeral and noun (2.2.9) for Sinhala based on Geiger (1938: 121–122) given in Dryer and Haspelmath (2013) is no ‘dominant order’ which is inaccurate. Gair (2007: 872) notes that ‘Numerals [in Sinhala] generally follow the head noun, and bear the definiteness and case inflection.’

(19) S. digə meesa hathak/namayak/vissak
Long table.PL seven/nine/twenty
‘Seven/nine/twenty long tables’

Like numerals, quantifiers too precede the noun in Tamil, as in *konjam seeni,* lit. ‘(a) little sugar’ and follow the noun in Sinhala, as in *seeni chuttak*, lit. ‘sugar (a) little’. Though the default order of numeral/quantifier and noun in Tamil is Quantifier/NumN, the former can optionally follow the latter, as in *pallikoodangal moondu,* lit. ‘schools three’, *seeni konjam,* lit. ‘sugar (a) little’. In Sinhala, on the other hand, the order of quantifier and noun is not obligatory; the quantifier can also precede the noun, as in *chuttak seeni*, lit. ‘(a) little sugar’, whereas the noun-numeral order (NNum) is obligatory. Of the nine languages compared, Sinhala is the only language which obligatorily requires this order (NNum) as the only order (see Appendix); note, however, that in Tamil, the default order is NumN, but it may also have NNum.

Numeral ‘one’ in Sinhala is an exception in that it may either precede or follow the noun or may be covert, as in (20aS–cS) respectively:

(20) a. S. mee gam-ee ekə iskoolə-(y)ak thiyenəwa
this village-LOC one school-INDF have.PRS
(20) b. S. mee gam-ee iskoolə ek-ak thiyenəwa
this village-LOC school one-INDF have.PRS
(20) c. S. mee gam-ee iskoolə-(y)ak thiyenəwa
this village-LOC school-INDF have.PRS

‘This village has a school/There is a school in this village.’

In (20aS), the numeral ‘one’ precedes the noun, while the indefinite marker follows the noun; in (20bS), the numeral follows the noun, like the other numerals in Sinhala; and in (20cS) the numeral does not occur, but the indefinite marker follows the noun. Note also that in both (20aS, cS) the indefinite marker needs to be suffixed to the noun, whereas in (20bS) the noun form of numeral ‘one’ (see below and also 2.3.1.2) contains the indefinite marker.

(21a,bT) below are the two Tamil equivalents of (20aS–cS):

(21) a. T. intha kiraama-(th)il oru pallikoodam iru-kku-thu
This village-LOC one/INDF school exist-PRS-3SGN
(21) b. T. intha kiraama-(th)il pallikoodam ondu iru-kku-thu
This village-LOC school one/INDF exist-PRS-3SGN

‘This village has a school/There is a school in this village.’

Lehmann’s (1989) account of numeral ‘one’ in Tamil is instructive and it explains the use of numeral ‘one’ in Sinhala too. According to Lehmann, the numeral ‘one’ in Tamil has an adjectival form *oru* and a nominal form *ondu.* The numeral ‘one’ in Sinhala too has an adjectival form *ekə* and a nominal form *ekak.*[22] Lehmann notes that the adjectival form of numeral ‘one’ in Tamil *oru* occurs prenominally (21aT), whereas its nominal form *ondu* occurs postnominally as a transposed noun modifier (21bT). Similarly, in Sinhala *ekə* which is the adjectival form of numeral ‘one’ occurs prenominally (20aS), whereas *ekak* which is its nominal form occurs postnominally as a transposed noun modifier (20bS).[23] It seems plausible to conclude that Sinhala has adopted the distinction between the two forms of numeral ‘one’, prenominal (adjectival) and postnominal (nominal) from Tamil, but more research is needed to establish the occurrence of this distinction among other Dravidian and NIA languages. As for the other numerals (from ‘two’), in Tamil, they precede the noun in the unmarked case but may optionally follow the noun, whereas in Sinhala, they obligatorily follow the noun (see (19S)).

Another feature that surfaces in (20) is the order of the indefinite marker and noun in a D(eterminer) P(hrase)/NP. Both languages derive the indefinite marker from numeral ‘one’: in Tamil *oru* the adjectival form of ‘one’(21aT) is used (Lehmann 1989), whereas in Sinhala the nominal numeral suffix, *-ak/-ek*, is used (Chandralal 2010). However, in Tamil, the indefinite marker obligatorily precedes the noun, as in (21aT), whereas in Sinhala it obligatorily follows the noun (20a,cS). What is important to note is that in utterances like (20aS) which occur frequently in Sinhala, indefiniteness is double marked both by the adjectival form *ekə* and the indefinite marker *-ak*, the latter is obligatory even if the former occurs; the latter may also occur alone, as in (20cS), but the utterances with the former alone (without the latter) are ungrammatical. This kind of double marking of indefiniteness in utterances like (20aS) describes the first form of coexistence, i.e., the new and the old categories co-existing in the same category—a structural effect of contact-induced grammaticalization.

It is also interesting to note that the ordinal numbers in the two languages are formed in the same way: in Tamil, the suffix *-aavathu* is added to the cardinal numeral, as in *eel-aavathu* ‘seventh’; *onpath-aavathu* ‘ninth’, whereas in Sinhala *-wæni* is added

---

[22] Note that in counting, the nominal form, e.g. *ondu* ‘one’, *rendu* ‘two’, is used in Tamil, whereas the adjectival form, which Chandralal (2010: 58) refers to as the citation form, e.g. *ekə* ‘one’, *dekə* ‘two’ is used in Sinhala.

[23] Surprisingly, only *eke*, the adjectival form of ‘one’ is used prenominally (see Masica 1991: 248), as in (20aS), whereas the adjectival forms of the other numerals cannot be used in this way.

to the base form of numerals (Chandralal (2010: 60), as in *hath-wæni* ‘seventh’; *namə-wæni* ‘ninth’; the ordinal ‘first’ in the two languages is an exception in that it is formed by adding the same suffixes to a special form of ‘one’, namely *muthal* and *paləmu,* as in *muthal-aavathu* and *palə(mu)-wæni* ‘first’ in Tamil and Sinhala respectively. These suffixes have been derived from one of the two copulas *aaku* in Tamil and *wendə* in Sinhala ‘be/become’ (3.6.4): *-aavathu* is ‘the future neuter participial noun form ...’ (Lehman 1989: 116), while *-wæni* may be a derivation of the future form *weyi* which consists of the verb stem *we* ‘be’ + *yi*, an alternant of *yæ*, the predicate (agreement) marking form in classical Sinhala (Paolillo 1994: 160, Gair 1998e: 242; see 6.4).

As for the order of constituents in QPs and DPs, the numeral/quantifier and the indefinite marker by default follow the noun in Sinhala unlike the numeral/quantifier and the indefinite marker which by default precede the noun in Tamil (see also 2.3.1.2 below). While these features in Sinhala may be remnants of an earlier Indo-Aryan grammatical system, as Lehmann’s (1978) claim about the aberrant teen numerals in which the single digit (head) precedes the decimal number (modifier/dependent), the fact that they are only found in Sinhala of all the NIA and Dravidian languages in the sample used makes it rather more likely that they are the result of independent, endogenous changes. It is also possible that these are substrate features of the earlier indigenous languages, Hela or the Vadda language mentioned in 1.5.

Finally, the default order of the verb and the negative marker in the two languages is the verb preceding the negative marker (see 3.4.1–3.4.3), as in (22b):

(22) a. T. avan poo-v-aan
He go-FUT-3SGM
S. eyaa yai
He go.FUT

‘He will go.’

(22) b. T. avan poo-ha-villai
He go-INF-NEG
S. eyaa yann-e nææ
He go.FUT-NMLZ NEG

‘He won’t go.’

Sinhala also has another negative construction in which the ‘verbal prefix’ *no-* is used as the ‘negator’ (Gair 2007: 885), as in (23S):

(23) S. eyaa no-yai
S/he NEG-go.FUT
‘He won’t go.’

In (22bT/S), the order of the NEG(ative) marker and the verb is VNEG (head-final), whereas in (23S) it is NEGV (head-initial). The preverbal (head-initial) negatives frequently used in Sinhala is a feature of NIA languages like Hindi and Gujarati, unlike Dravidian which have only clause-final negative particles. Note that the clause-final negatives (22bT/S) is the default negative form in Tamil and Sinhala. Referring to Southworth's (1971: 264) finding that Marathi alone among the Indo Aryan languages has developed a whole set of negative auxiliaries in the final position on the Dravidian pattern, Gair (1998d: fn. 7, p. 332) argues that Sinhala has the verb plus negative auxiliary which 'is essentially the functional equivalent of the Tamil verb plus *illai.*'[24] The pre- and postverbal negatives in Sinhala describe the second form of coexistence in which two constructions, as Heine and Kuteva (2005: 130) 'coexist as alternative constructions available to the speakers of the replica language.'

The discussion shows that though Sinhala is a consistently head-final language like Tamil, it diverges from Tamil with respect to some word order features, such as the aberrant teen numerals with the single digit preceding the decimal; the numeral and the indefinite marker following the head noun; and the negator preceding the verb. The aberrant teen numerals may be a remnant of an ancient Indo-Aryan order. The numeral/indefinite marker following the noun, which is a feature unique to Sinhala among the South Asian languages, could have conceivably been acquired from the indigenous languages, Hela or Vadda languages of Sri Lanka (1.5) before its contact with Tamil. Alternatively, it may be a result of an endogenous (language-internal) change in Sinhala. The preverbal negation, however, is a common NIA feature. To conclude, although there is considerable divergence between the two languages with respect to the features examined in this subsection, the coexistence of two forms in some of the features of Sinhala indicate that they have undergone contact-induced changes. Gair (1998a) claims that the overall close resemblance between Sinhala and the South Dravidian languages in regard to branching direction (left branching/head final order) indicates that Dravidian influence was a factor in Sinhala's acquisition of its present syntactic character, even more than in most other NIA languages. 2.3.1.2 below examines how (in)definiteness is marked in the two languages.

---

[24] Southworth (1971: 264) also notes that the Marathi negative forms also consist of 'the inherited Indo Aryan material (including the initial morpheme *na-*, of Indo-European origin) but have clearly been remodeled on the prevailing Dravidian pattern.'

2.3.1.2 (In)definiteness

As shown in 2.3.1, Tamil and Sinhala mark indefiniteness using the indefinite markers which are derived from numeral ‘one’. This is crosslinguistically very common and a well-attested contact-induced change (Heine and Kuteva 2008). [25] However, the indefinite marker precedes the head noun in Tamil, but follows the head noun in Sinhala (see examples in 2.3.1.1).

If indefiniteness is overtly marked in the two languages, definiteness is covertly marked, that is, there is no definite marker, in the two languages, as in (24):

(24) T.	athipar	indai-kku	pallikoodath-(u)kku	var-a-(v)illai	
	Principal	today-DAT	school-DAT	come-INF-NEG	
S.	viduhalpathi-thuma	adə	iskoole-tə	aaw-e	nææ
	Principal-HON	today	school-DAT	come.PST-NMLZ	NEG

‘The principal didn’t come to school today.’

Note that in (24) the principal, being the only one in a school, needs to have definite reference, but it is not marked for definiteness. In the absence of a definite article, the two languages often use the distal demonstrative ‘that’ *antha* in Tamil and *ee* in Sinhala as the definite marker. The use of the distal demonstrative as the definite marker is also a well-attested contact-induced change (Heine and Kuteva 2008: 60–62). It may be a fairly common phenomenon among languages of South Asia. [26]

Masica (1991:248) distinguishes two types of ‘morphological definiteness marking’ in NIA languages: ‘the Eastern (Bengali-Assamese-Oriya) type and the Sinhalese type’. In both, he claims, it is a matter of specifying, through suffixes, the status of a noun in discourse as identified (previously mentioned or known = “definite”) or as unidentified (new = “indefinite”). The Eastern type uses specifier suffixes called “articles” or “enclitic definitives or numeratives” derived from a numeral-classifier system. Added to nouns, these suffixes specify identified status; added to numerals preceding nouns, they specify unidentified status (Masica 1991: 250). The Sinhalese type involves the suffixed “indefinite articles” *-ek* (animate) and *-ak* (inanimate, sometimes also feminine). Nouns without them are deemed to be definite: *laməya*/*laməyek* ‘the child/a child’. The numerals and quantifiers follow the nouns (2.3.1.1); hence these indefinite markers ‘impart the same sense of Specified-Indefinite

---

[25] In the two languages, the numeral ‘one’ is also used in the formation of inclusive generic pronoun (human), *oruvan* (masculine)/*oruthi* (feminine) in Tamil and (*ek*)*kenek* in Sinhala.

[26] Marathi uses the demonstrative as the definite marker (Nayudu 2008: 16).

(Unidentified) to numerals [and quantifiers]’, e.g. *lamay denna*/*lamay dennek* ‘the two children/two children’ or *lamay hungədenek* ‘many children’ (Masica 1991: 248–250).

Masica (1991: 372) proposes the following structure of the Sinhalese NP, given in (25S):

(25) S. Det^1 + (Adj) + N + (Quant) + Det^2, where Det^1 (= Definite, i.e. Demonstratives and ZERO) and Det^2 (= Indefinite, i.e. *-ek*/*-ak*) are mutually exclusive.

What Masica (1991) fails to include in (25) is the numeral which occurs in the (postnominal) position of the quantifier. Note that the numeral and quantifier are in complementary distribution; the numeral obligatorily follows the noun, whereas the quantifier optionally follows the noun. What is significant to note is that in Sinhala, definiteness (via the use of demonstratives) is syntactically realized pre-nominally, whereas indefiniteness is realized post-nominally (see 2.3.1.1 above).

The structure of the Tamil NP, based on (25S) is given in (26T):

(26) T. Det/Quant/Num + (Adj) + N, where Det (=Definite, i.e. Demonstratives and ZERO and Indefinite i.e. *oru*) are mutually exclusive.

As mentioned above, in Tamil the numeral and quantifier precede the noun in the unmarked case, but may also follow the noun. Note also that while quantifier and numeral are mutually exclusive, determiner and quantifier/numeral are not mutually exclusive; they can co-occur, as in *antha rendu pillaikal* ‘those two children’, which holds true for Sinhala, as in *ee lamay denna* ‘those two children’.[27] However, unlike in Sinhala, both definiteness and indefiniteness are syntactically realized pre-nominally in Tamil.

Based on Paramasivam’s (1983) claim that there is interdependence between the accusative case and the deictic category of definiteness in Tamil, Lehmann (1989: 27) notes that when the direct object NP of a transitive verb is definite—either it consists of a personal noun or contains a determinative noun modifier such as possessive pronoun or the demonstrative adjective— the accusative case suffix obligatorily occurs on the object NP (for cases in the two languages, see 3.6.1.1). Even if the object NP does not contain any noun modifier, the accusative case suffix marks the NP as definite, as in

---

[27] Unlike the English demonstratives, the demonstratives in Tamil and Sinhala are number-invariant.

(27aT) below. If a direct object NP is not marked with an accusative marker, it is not interpreted as definite. However, in most cases the indefinite direct object NPs too require the accusative case, as in (27bT). If object NPs have generic reference, as in (27cT), they are not marked for accusative.[28]

As for Sinhala inanimate object NPs —whether they have definite or indefinite reference—they are not marked for accusative, as in (27a,bS), while definite animate object NPs are obligatorily marked for accusative (27aS) and indefinite animate object NPs are optionally marked for accusative (27bS). Finally, like Tamil, if object NPs have generic reference, as in (27cS), they are not marked for accusative.

(27) a. T. avan ravi-ai/ (than-da/ antha) kulanthai-ai/ naai-ai / petti-ai
He Ravi-ACC (self-GEN that) baby-ACC dog-ACC box-ACC
thed-in-aan
search-PST-3SGM
S. eyaa ravi-wə (thaman-ge/ee) baba-wə / balla-wə / pettiya(*-wə)
He Ravi-ACC (self-GEN that) baby-ACC dog.ACC box
hewwa
search.PST

‘He searched for Ravi/the baby/the dog/the box (he searched for Ravi/(his/that)baby/dog/box).’

(27) b. T. avan oru kulanthai-ai/ oru naai-ai/ oru petti-ai
He one(INDF) baby-ACC one(INDF) dog-ACC one(INDF) box-ACC
kan-d-aan
see-PST-3SGM
S. eyaa bab-ek(-wə)/ ball-ek(-wə) / petti-ak(*-wə) dakka
He baby-INDF-ACC dog-INDF-ACC box-INDF see.PST

‘He saw a baby/dog/box.’

(27) c. T. avan oru kulanthai-ai / oru naai/ oru petti
He one(INDF) baby one(INDF) dog one(INDF) box
thedu-rr-aan
search-PRS-3SGM
S. eyaa bab-ek/ ball-ek/ petti-ak hoyənəwa
He baby-INDF dog-INDF box-INDF search.PRS

‘He is searching for a baby (to adopt)/dog (to raise)/ a box (to use).’

Paramasivam’s (1983) and Lehmann’s (1989: 27–28) claim that the accusative case marks definiteness in Tamil needs further research because of the irregular correlation

---

[28] However, this cannot be completely ruled out because sometimes object NPs with generic reference are also marked for accusative.

between the accusative case and definiteness as evidenced in (27bT) (for the opposite view see Krishnamurti (2003: fn. 8, p. 432). Nevertheless, if Paramasivam (1983) and Lehmann (1989) are correct, Tamil has what is known as differential object marking (DOM) common in languages with overt case marking of direct objects. In DOM languages, some object NPs are overtly case marked, whereas the others are not, depending on semantic and pragmatic features of the object (Aissen 2003). In Tamil, definiteness seems to determine DOM, which, however, cannot be established as some indefinite object NPs are accusative case marked. Sinhala too has DOM in that only animate object NPs may be accusative case-marked, while inanimate object NPs are not. In Sinhala, therefore, DOM is determined not by definiteness, but by animacy.

Noting that the object is not distinguished from subject by case in many NIA sentences, Masica (1991: 365) opines that both may be in the nominative in keeping with their split ergative/absolutive system. Unlike Sanskrit or Dravidian, ‘there is generally no distinctive accusative case in NIA’ languages—except for the accusative suffix -*wə* in Sinhala and distinctive accusative pronominal suffixes in NIA languages like Poguli (fn. 14, p 478, ibid.)—because historically it is merged with the nominative. In the absence of the accusative case, the objects may take case in the form of the dative marker. This marker which he calls ‘dative-accusative’ performs two functions which, in his view, are more pragmatic than syntactic in that (i) in the case of non-human nouns, it generally indicates a definite object, that is, to show that the object is already known; (ii) in the case of human nouns, it stresses their patienthood, a marked status (human NPs normally being agents). Masica (1991) notes that even though function (i) is minimized in languages which have other means of marking definiteness, it, nevertheless, exists in NIA languages except in Sinhala (27a,bS). Further Masica (1991: 366) claims that the human-patient marking (function (ii) above) too is general in all the NIA languages except Sinhala.

In conclusion, with respect to the marking of (in)definiteness, Tamil and Sinhala—the latter, unlike the Eastern type NIA languages—use the same strategies:
(i) the indefinite marker is derived from numeral ‘one’ which, in the unmarked case occurs pre-nominally in Tamil, whereas it obligatorily occurs post-nominally in Sinhala; (ii) there is no distinctive definite marker; and (iii) the distal demonstrative occurs optionally as the definite marker. In instances like (27aT) the accusative case suffix, as Parmasivam (1983) and Lehmann (1989) claim, marks definiteness pragmatically on the object NP in Tamil, while in Sinhala it marks definiteness pragmatically on the animate object NP in instances like (27aS). What is important though is that in Tamil and

Sinhala there is no distinctive definite marker, whereas the Eastern type NIA languages such as Bengali, Oriya and Assamese among others use definite markers referred to as 'specifier suffixes' or 'enclitic definitives' in the literature (Masica 1991: 250). The discussion shows that Sinhala shares with Tamil the same strategies of marking (in)definiteness on NPs. However, whether this is an effect of its contact with Tamil or an areal feature of South Asia cannot be determined. 2.3.1.3 below provides evidence for the restructuring of another important morphosyntactic phenomenon, namely relative clauses.

2.3.1.3 Prenominal relative clauses

The order of relative clause and noun in Sinhala and Tamil is the relative clause preceding the noun (head) ((10) in 2.2.10).[29] This prenominal relative or adjectival participial clause is distinctively Dravidian, while some NIA languages use correlative clauses (see 2.3.1.4). Distinguishing between the relative clause and the adjectival participle as the salient means of embedding in NIA languages, Masica (1991: 408) claims that Sinhala, 'like its neighboring Tamil has only the latter.' In South Asia, the two NIA languages that have prenominal relative clauses are Sinhala and Marathi, while others like Hindi, Bengali and Oriya (see Appendix) have correlative clauses. Bengali also uses a participial kind of relatives (pc. van der Wurff). Gair (1998a, fn. 5, p 306) notes that Sinhala appears to have lost the correlative construction very early. It is rarely used in modern colloquial Sinhala in which the only extant type is the prenominal RelN type (see also 4.2.1, 4.2.2 and 4.3). As for Marathi and Sinhala which have prenominal relatives, Marathi is believed to have adopted it from Kannada, another Dravidian language, which it has been in contact with (1.3), while Sinhala may have adopted it from Tamil. Sinhala may originally have used some participles attributively, resulting in relative-like constructions similar to English *a recently discovered manuscript* or *a fast disappearing custom*, which also exist in some other NIA languages. But modern Sinhala has gone well beyond such simple patterns and developed a complete grammatical system for prenominal relatives. The model for this is likely to have been Tamil which shares the prenominal relative with the other Dravidian languages. The next subsection examines correlative clauses in the two languages.

---

[29] The value of this feature for Sinhala, based on Gair (1970), given in Dryer and Haspelmath (2013) is NRel, which is incorrect. The correct value of this feature for the two languages is RelN. In his later work, Gair (2007: 894) states that 'Relative clauses always precede the noun head.'

2.3.1.4 Correlative clauses

According to Krishnamurti (2003: 447), ‘Many Dravidian scholars think that correlative clauses occur in Dravidian through diffusion from Indo-Aryan.’ A possible reason for this conclusion is that many NIA languages have correlatives as their default form, while Dravidian languages have prenominal relatives as their default form. However, rejecting the claim that ‘the correlative constructions were borrowed in Dravidian from Indo-Aryan’, Steever (1988: 33) argues that ‘correlative clauses are found throughout the Dravidian family with much the same structure: the principles of comparative reconstruction would dictate their presence in the proto-language’ (for other reasons in support of the Dravidian origin of correlatives, see Steever 1988; also Krishnamurti 2003).

Correlative clauses in Tamil and Sinhala are formed using two clauses—a relative pronoun of the first (embedded) clause, co-referring with a correlative demonstrative (expressed or implied) in the correlative main clause (Slade 2013: 2), as in (28T) (Annamalai & Steever 1998 cited in ibid. p. 2, ex. (3)) as the Tamil correlative clause: [30]

(28) T. [**yar** anke mutalil vantu ceru-v-ar] $_{\text{RC}}$-**oo**
[**REL-PRON**.NOM there first come-CONV arrive.FUT.3PL] $_{\text{RC}}$-***oo***
[avar tikettu vankalam] $_{\text{CC}}$
they.NOM.PL ticket.NOM buy.PERM[missive]] $_{\text{CC}}$
“Let whoever reaches there first buy the tickets.”

The relative clause in (28T) begins with the relative pronoun *yaar* ‘who(ever)’ and ends with the common ‘Dravidian clitic *-oo*’.[31]

According to Slade (2013), this kind of correlative clauses which are well-attested in classical and modern literary Sinhala, like Tamil, are made up of a relative clause headed by a relative pronoun *yam*, and ending either with the question particle *də* or the conditional particle *nam* together with the correlative main clause, as in (29S) (Gair and Karunatillake 1974: 295 cited in Slade 2013, p. 2, ex. (2)):

---

[30] For consistency, (28)–(30) below are glossed, as in Slade (2013).

[31] Note that Slade (2013) glosses the third person singular honorific suffix at the end of the finite verb in the relative clause in (28T) as ‘3PL’ to mark the honorific status of the subject (indefinite pronoun) *yar* ‘whoever’, which is one way of marking honorificity in Tamil (see also the correlative main clause in (30T) below). Another way of expressing honorificity in Tamil is to add the plural suffix to the finite verb.

(29) S. [**yam** kumariyak ohu duṭuvā]$_{RC}$ **da**
[**REL-PRON** princess.INDEF him see.PAST.3SG.FEM] $_{RC}$ ***da***
[oo ohu kerehi piḷinda sit ætikara gattāya]$_{CC}$
[she him towards connected mind developed get.PAST.3SG.FEM]$_{CC}$
"**Whatever princess** saw him fell in love with him." [Literary Sinhala]

As mentioned above, Sinhala has apparently lost the correlative clause and adopted the prenominal relative clause. In modern colloquial Tamil, on the other hand, correlative clauses like (30T) are frequently used:

(30) T. [**entha** pediyan nettu paattu paad-in-aan]$_{RC}$-**oo**
[**REL-PRON** boy yesterday song sing-PAST-3SGM]$_{RC}$-*oo*
[antha pediyan-ai teecher pukal-nth-aar] cc
[that boy-ACC teacher praise-PAST-3PL] cc
'The teacher praised the boy who sang a song yesterday.'
'**Whichever boy** sang a song yesterday, that boy, the teacher praised.'

Steever (1988), Krishnamurti (2003) and Subbarao (2010) note that Dravidian languages have both prenominal relatives (2.3.1.3) and correlatives which Subbarao (2010) calls 'gap relative clauses' and 'sentential relative clauses' respectively. The majority of the NIA languages also have correlatives as their default form (see Appendix). Thus, while it may be the case that Sinhala replicated its correlative from Tamil (Slade 2013), it is at least as likely that the Sinhala correlative clause predates its encounter with Tamil in Sri Lanka. 2.3.1.5 below shows that Sinhala has modelled its question particle on the Tamil/Dravidian correlative particle.

2.1.3.5 The Sinhala question particle: a diachronic study

In his diachronic study of question particles and relative clauses in Sinhala, Slade (2013) claims that the Sinhala question particle *da* (hereafter *də*) 'originated from the Sanskrit alternative question conjoiner *utaho*,' occurring 'at the front of the second part of the disjunction'. Its counterpart in Pali is *udahu.* Slade observes that the original environment of *də* in Sinhala is specifically interrogative disjunction (alternative questions), which subsequently began to be used in yes/no questions. Once fully established in alternative and yes/no questions, it appears to have gradually been generalized to all interrogative contexts, including wh-questions; hence, it appears more fully established in yes/no questions than in wh-questions in Sinhala. In recent history, *də* subsequently spread to show up in indefinites. Distinguishing four 'stages/forms of Sinhala', namely (1) Old Sinhala [OS; ca 8th –10th c. A. D); (2) Classical Sinhala [CS; 12th–15th c. A. D]; and two modern varieties, (3) Modern Literary Sinhala [LS] and (4)

Modern Colloquial Sinhala [MCS] (ibid. 2), Slade provides the following hypothetical pattern of expansion of *də* in table 2.2 (table 2 in Slade 2013: 15) from its origin as a particle specific to alternative questions.

	pre-pre-OS	pre-OS	OS	CS	LS	MCS
alternative question		X	X(?)	X	X	X
yes/no-question		(X)	(X)	X	X	X
wh-question			(X)	(X)	X	X
wh-indefinite						X
rel. clauses w/*yam* [relative pronoun]				X	X	n/a

Table 2.2: Appearance of *də* in various syntactic contexts in the history of Sinhala

The fact that Sinhala has apparently lost the correlative clause accounts for n(ot)/a(pplicable) specified under MCS. In addition to its extension to other interrogatives, *də* exhibits other syntactic changes which made it diverge from its Sanskrit and Pali precursors. Most importantly, it has undergone change from a proclitic-type element to an enclitic-type element. The motivation for this change in *də*, as Slade assumes, is the Dravidian influence—given that the Dravidian question particles are post-clausal enclitics. In addition to correlative clauses, the Dravidian particle -*oo* (common to all Dravidian languages, Subbarao 2010) occurs in polar interrogatives (2.2.12, also 2.3.1.6), indefinitives (3.6.3; existential quantifiers, 4.3.7) and disjunctive coordination (4.3.5).[32] Note that Malayalam uses the same marker -*oo* as its polar question clitic, while Tamil and Kannada uses -*aa* as their question clitic which may have evolved from -*oo*. [33] Slade argues that *də* may have been modelled on the Dravidian 'final "clause-closing" particle' -*oo* which occurs in the correlative clauses believed to be Dravidian in origin (see 2.3.1.4).

According to Slade (2013: 16), 'The use of *də* in classical and modern literary Sinhala relative-correlative constructions parallels the employment of the common Dravidian question particle -*oo*[/-*aa*], and the appearance of *də* in this syntactic environment is likely due to Dravidian influence, although the evidence is not entirely

---

[32] It is not known in which construction—the correlatives, the polar interrogatives, the indefinitives or the disjunctive coordination—the Dravidian particle -*oo* originated.

[33] Sinhala uses the same clitic -*oo* in indefinites (3.6.3; and existential quantifiers, 4.3.7) and its slight variant -*ho* as the disjunctive coordinator (4.3.5).

clear.’ He adds that the use of Dravidian correlatives appears to date to an early period in Dravidian which precedes the appearance of *də* in Sinhala correlatives. Sinhala may have modelled its question particle on the Tamil correlative particle for the following reasons: i) unlike its Sanskrit/Pali proclitic precursors and the question particles in some NIA languages which occur clause-initially, the particles occur clause-finally; ii) the Dravidian correlatives pre-date Sinhala correlatives (see 2.3.1.4); and iii) the two languages use the same correlative marker—the clitic *-oo* (and its variant *-aa* in polar questions) in Tamil and *də* in Sinhala—among others, in their correlatives, polar questions, and only in Sinhala, its content/wh-questions. In 2.3.1.6 below, the interrogatives in the two languages are examined to determine whether interrogatives in Sinhala have undergone any contact-induced changes.

2.3.1.6 Interrogatives

The second feature among the 17 features, which is different between the two languages is the position of polar question particles (2.2.12). The two languages do not share the same feature value because Tamil does not have a question particle. Instead, it has a question clitic *-aa*, whereas Sinhala has a question particle *də* (12). Dryer (2013k) does not include interrogative affixes on verbs as a value of WALS 92A, although he notes that the distinction between interrogative affixes and separate interrogative particles is often hard to make. The polar question clitic/particle in Tamil and Sinhala, respectively, exhibit the same properties in polar questions.

It can be argued that Sinhala polar questions have been modelled on Tamil polar questions. The first reason in support of this claim is that the question clitic in Tamil (12bT) and the question particle in Sinhala (12bS) occupy the same clause-final position in default polar questions. All the Dravidian languages have clause-final question clitics, while some NIA languages like Hindi-Urdu, Punjabi etc. do not have clause-final question markers (4.4.2; see also Biberauer et al. 2009). Secondly, in constituent polar questions, the polar question clitic/particle occurs adjacent to a distinct constituent, lending a focus reading, as in (31a,b) (based on (12a/b)):

(31) a. T. nandini-**aa** poona kilamai pillaikal-(u)kku inglish padippi-th-athu/
Nandini-Q last week students-DAT English teach-PST-NMLZ/
padippi-ch-aa
teach-PST-3SGF(HON)
S. nandini **də** giyə sumaane lamay-tə ingriisi igannuw-e
Nandini Q last week students-DAT English teach.PST-NMLZ

'Was it Nandini that taught English to students last week?'

(31) b. T. nandini poona kilamai(**-aa)** pillaikal-ukk(**-aa)** inglish(**-aa)**
Nandini last week(-Q) students-DAT(-Q) English(-Q)
padippi-ch-aa
teach-PST-3SGF(HON)
S. nandini giyə sumaane **(də)** lamay-tə **(də)** ingriisi **(də)**
Nandini last week (Q) students-DAT (Q) English (Q)
igannuwe
teach.PST-NMLZ

'Was it last week that Nandini taught English to students?'
'Was it to students that Nandini taught English last week?'
'Was it English that Nandini taught to students last week?'

In all constituent polar questions in Sinhala, the finite verb is changed into its nominalized form (31a,bS). In Tamil only when the logical subject of the sentence is questioned, the finite form is changed into its nominalized form, while the finite form of the verb is also used (31aT) above; when the other constituents are questioned, the finite form of the verb is used (31bT) (this will be discussed in more detail in Chapter 6). The two languages have modelled their constituent polar constructions on their focus particle construction which is one of the two kinds of constructions in which exhaustive focus is assigned to a constituent in a sentence in the two languages; the other kind is the cleft construction in which the two languages use the nominalized (NMLZ) form of the verb. Sinhala replicated the two kinds of focus constructions from Tamil (see 6.1, 6.2). The positions of the polar question marker in polar and constituent polar questions in the two languages and the fact that Sinhala has extended the NMLZ form from the cleft construction to the constituent polar questions indicate that Sinhala polar and constituent polar questions have been modelled on those in Tamil (see 2.3.1.5 above, for the origin of the Sinhala question particle).

Though the feature value of the position of interrogative wh-phrases in content questions in the two languages (2.2.13) is the same, there are two major differences between wh-questions in Tamil and Sinhala. First, in Tamil, wh-phrases such as *yaar* 'who', *enna* 'what', *enge* 'where' etc. do not have an overt question clitic or particle (see (32T) below). In Sinhala, as Kishimoto (2005: 4) notes, 'wh-words are indefinite pronouns.' The wh-words such as *kau*(*ru*) 'who', *mokak* 'what', *kohe* 'where' combine with the common question particle *dә* to form wh-phrases *kau dә* 'who', *mokak dә* 'what', *kohe dә* 'where' in wh-questions, as in (32S) (formulated from (13) in 2.2.13 above):

(32) T. nandini poona kilamai yaar-ukku inglish padipi-ch-aa
Nandini last week who-DAT English teach-PST-3SGF(HON)
S. nandini giyә sumaane kaa-tә **dә** ingriisi igannuw-e
Nandini last week who-DAT Q English teach.PST-NMLZ

'To whom did Nandini teach English last week?'

The second difference between the wh-questions in the two languages is that Tamil uses the finite form of the verb in wh-questions (32T) except in the question which requires the logical subject of the sentence as the answer, as in (33T); the finite form of the verb is rarely used in this kind of wh-question. Sinhala, on the other hand, uses the NMLZ form of the verb in all wh-questions, as in (33S) also (32S):

(33) T. yaar poona kilamai pillaikal-ukku inglish padippi-th-athu
Who last week students-DAT English teach-PST-NMLZ
S. kau **dә** giyә sumaane lamay-tә ingriisi igannuw-e
Who Q last week students-DAT English teach.PST-NMLZ

'Who taught English to the students last week?'

Here again, Sinhala has extended the NMLZ form to wh-questions.

Gair (2007: 887) observes that the wh-phrases always co-occuring with the question particle in Sinhala is 'a way in which it differs from other languages, such as the South Dravidian ones and current Japanese in which they are in complementation [sic] [complementary distribution].' By default, the question particle in Sinhala wh-questions occurs immediately following the interrogative phrase. However, in a few instances the question particle comes at the end of the question, in which the particle is claimed to be functioning as the scope marker (for more information, see 6.5), as in (34S) below. In Tamil, on the other hand, wh-questions do not have a question particle as in (34T):

(34) T. nandini oru maasath-(u)kku vahuppu-kku evvalavu edu-kkir-aa
Nandini one month-DAT class-DAT how much take-PRS-3SGF(HON)
S. nandini maas-ek-(ə)tə panthiyə-tə kiiy-ak gannəwa **də**
Nandini month-INDF-DAT class-DAT how much-INDF take.PRS Q

‘How much does Nandini charge for the class for a month?’

In (34S) the particle occurs at the end of the clause and the finite form of the verb is used instead of the NMLZ form of the verb.[34] The occurrence of the question marker in these two positions (that is, clause internally adjacent to the wh-phrase or clause finally following the finite verb)—also considering it in light of the criteria Zwicky and Pullum (1983) propose to distinguish between clitics and affixes—shows that it is a particle and not a clitic or affix.

The question particle with its occurrence in two positions clause internally and clause finally and the use of NMLZ form of the verb in Sinhala wh-questions make them differ appreciably from Tamil wh-questions. The wh-questions in the two languages are similar in that, as expected in SOV languages, the wh-phrases are in situ (Greenbergian Syntactic Universal 12). The extension of the NMLZ form to the wh-questions in Sinhala may, however, be an effect of the cleft construction that Sinhala has replicated from Tamil. In the next subsection, the subordinate clauses in the two languages are analyzed.

2.3.1.7 Subordinate clauses

The two languages converge on subordinate clauses, too, in that the subordinators occur clause-finally in subordinate clauses which by default precede the main clause ((14) in 2.2.14). The majority of the subordinators in Tamil and Sinhala are suffixes, while some are particles, as shown in table 2.3:[35]

---

[34] The form in which the question particle is adjacent to the wh-phrase is also possible, as in (iS), which obligatorily requires the NMLZ form of the verb, unlike the finite form of the verb in (34S) (see 6.5.2 for more information).

(i) S. nandini maas-ek-tə panthiyə-tə kiiy-ak-**də** gann-e
Nandini month-INDF-DAT class-DAT how much-INDF-Q take.PRS-NMLZ
‘How much does Nandini charge for the class for a month?’

Tamil, however, has only one form (34T) because it does not have a question particle.

[35] The feature value for Sinhala given in WALS 94A ‘Order of adverbial subordinator and clause’ is ‘mixed’ based on Geiger (1938) who may have arrived at this conclusion owing to the two kinds of subordinators, namely suffixes and particles in Sinhala. This would hold true for Tamil, as well.

	Tamil	**Sinhala**
(35)	ravi nalla padith-**aal**, Ravi well study.PTCP.if Ø sothanai pass-pannu-v-aan exam pass-do-FUT-3SGM	ravi hondətə paadamkər-**oth**, Ravi well study.PTCP.if Ø vibhage pass-karai exam pass-do.FUT
	‘If Ravi studies hard, he will pass the examination.’	
(36)	ravi nalla padithth-**um**, Ravi well study.PTCP-though Ø sothanai pass-pann-a-(v)illai exam pass-do-INF-NEG	ravi hondətə paadamkərəla-**th**, Ravi well study.PTCP.though Ø vibhage pass-kara-e nææ exam pass-do.PST.NMLZ NEG
	‘Though Ravi studied hard, he did not pass the examination.’	
(37)	ravi nalla padikira-(th)**aala** Ravi well study.PTCP-because Ø nalla marks edu-kir-aan good marks get-PRS-3SGM	ravi hondətə paadamkərənə **nisa** Ravi well study.PTCP because Ø hondə lakunu gannəwa good marks get.PRS
	‘Because Ravi studies hard, he gets good marks.’	
(38)	ravi padikira **poothu** Ravi study.PTCP when amma avan-ai koopi-tt-aa mother he- ACC call-PST-3SGF(HON)	ravi paadamkərənə **kotə** amma Ravi study.PTCP when mother eyaa-tə andəgahuwwa he-DAT call.PST
	‘While Ravi was studying, mother called him.’	
(39)	ravi padikira **maathiri** Ravi study.PTCP like avan-ra thambi-um he-GEN (younger)brother-INCL padi-kir-aan study-PRS-3SGM	ravi paadamkərənəwa **wage** Ravi study.PRS like eyaa-ge malli-th he-GEN (younger) brother-INCL paadamkərənəwa study.PRS
	‘Ravi’s younger brother studies like Ravi does.’	
(40)	ravi padiththa-((th)-ukku) **piraku** Ravi study.PTCP-DAT after Ø vilaiaad-a poonaan play-INF go-PST-3SGM	ravi paadamkaraa-tə **passe** Ravi study.PTCP.DAT after Ø sellaŋkərandə giya play.INF go.PST
	‘After Ravi studied, he went to play.’	

Table 2.3: Subordinator suffixes

The order of adverbial subordinator and clause (2.1.14) in the two languages and the two kinds of subordinators show that Sinhala has modelled its subordinate clauses on Tamil subordinate clauses. Gair’s (1998d: 208) conclusion confirms this claim:

> … subordinate verbal structures as a whole are of a strikingly Tamil and Dravidian character. The exclusive use of preposed adjectival clauses as the equivalent of relativization stands out here, as do the conjunct affixes such as the conditional and concessive (i.e., Sinhala *giyot*, Tamil *poonaal* 'if one goes'; Sinhala *giyat*, Tamil *poonaalum* 'even if one goes') as well as the use of particles or other forms following verbal adjectives so as to form adverbial clauses (Sinhala *yana kota*, Tamil *pookira polutu* or colloquial *poora appa* 'when one goes'). Whatever the problem in accounting for specific similarities, the cumulative effect is nothing short of overwhelming, particularly considering the lack in Sinhala of the alternate structures found in northern Indo-Aryan, such as correlatives and clause-initial conjunctions.

The last segment of the quotation seems to suggest that Sinhala shares more features with Tamil than it does with other NIA languages which it is genetically related to.

2.4 Conclusion

The word-order related features compared in this chapter show extensive convergence between Tamil and Sinhala in that the values of 15 of the 17 features related to word order are similar in the two languages. Since most of the languages of the region belonging to both the NIA and the Dravidian language families share the same values, the results obtained may not adequately reflect the restructuring that Sinhala has undergone owing to its contact with Tamil, though. Moreover, the main areal feature of these languages, the SOV word order, correlates with the order of the other constituents from a universal perspective, which may in part account for the similarities between these features.

However, some of the features discussed in this chapter provide reliable evidence for the restructuring of Sinhala on the model of Tamil. First, the two kinds of coexistence of the following features in Sinhala discussed in 2.3.1.1 demonstrate the contact-induced restructuring Sinhala has undergone: (i) the aberrant (possibly NIA) and unmarked (Dravidian) teen numerals; (ii) the double marking of the adjectival form of 'one' and the indefinite marker (20aS); and (iii) the pre-(NIA) and post-verbal (Dravidian) negatives. Secondly, Sinhala adopted the prenominal relative clause from Tamil and lost the correlative clause which the other NIA languages have. Thirdly, although the Sinhala question particle *də* had its origin in the Sanskrit/Pali alternative question conjoiner, it has been modelled on the common Dravidian correlative particle -*oo* and extended subsequently to other interrogative contexts from an early period. The fact that Dravidian correlative clauses predate the Sinhala correlatives is consistent with the claim that Sinhala may have replicated the correlatives also from Tamil. Fourthly, though interrogatives in Sinhala appear to differ from those in Tamil, there are some similarities between them as well. The position of the question clitic/particle in polar and constituent polar questions in the two languages is the same which suggests that they have been modelled on Tamil polar and constituent polar questions. Moreover, Sinhala has extended the use of the nominalized form of the verb—also used in the cleft construction (6.1) which it replicated from Tamil—as its default verb form in wh-questions. Finally, it is fairly evident that Sinhala has modelled its subordinate clause on the Tamil subordinate clause because of the clause-final subordinators in the two languages and the two kinds of subordinators (suffixes and particles) that both have.

Biberauer et al. (2009) remark that in the South Asian linguistic area, the more rigidly head-final Dravidian languages have a long history of contact with disharmonic NIA languages. Close examination of the word-order related features discussed in this chapter and others like the clause final complementizer (see 4.2.7 and 4.3.10) amply demonstrate that Sinhala, because of its contact with Tamil and the consequent restructuring, has become a rigidly head final language like Tamil/Dravidian (unlike other major NIA languages like Hindi and Bengali).

The two languages in contact, Tamil and Sinhala, fulfil the diagnostics of contact-induced restructuring/replication (Heine 2007). Firstly, the examples given in the two languages to illustrate the features analyzed in this chapter show that a considerable number of constructions are mutually intertranslatable word by word, if not, morpheme by morpheme. Secondly, a number of features that Sinhala shares with Tamil, it shares with the other Dravidian languages too, e.g. the prenominal relative clause, but does not share with the other NIA languages. The discussion lends support to the claim that the morphosyntactic phenomena in Sinhala discussed in this chapter have undergone considerable changes vis-à-vis the corresponding phenomena in Tamil.

Chapter 3

Features related to simple clauses

3.1 Preliminaries

Chapter 2 examined the convergence between Tamil and Sinhala on features related to word order. This chapter examines features related to simple clauses in the two languages with a view to determining the features that are similar and those that are different. For the purposes of discussing the relevant features, a simple clause is defined as the smallest grammatical unit that can convey a proposition. In a clause, the verb, which expresses an event, action, process, situation or state, occurs with the arguments which it requires (Tallerman 2015). The arguments that a verb requires depend on the type of verb, e.g. intransitive, transitive etc. The arguments, which are determiner/noun phrases (DPs/NPs), play distinct semantic roles within the clause, and are assigned case to indicate the syntactic relationship among them. The number of arguments a verb requires is usually referred to as its valence, which can be increased or decreased by way of morphosyntactic operations. In many languages verbs agree with one or more of their arguments, that is, the various properties of the DP/NP arguments such as person, number, and gender or noun class are marked on the verb (Tallerman 2015: 47).

Further variation is found when the proposition expressed in a simple clause is questioned or negated. Simple clauses/sentences are traditionally analyzed as being made up of a subject, the doer of the action, and a predicate, the constituent that has all the obligatory constituents other than the subject. Broadly speaking, the predicate usually assigns a property to the subject, or locates it in some space. These morphosyntactic processes are called predication. The current chapter covers the above features as they are found in simple clauses in the two languages.

The 27 features compared in this chapter constitute a crucial part of morphosyntax because they are concerned with how aspects of morphology and syntax establish relations between meaningful units, i.e., constituents, performing different functions on both clause and phrase levels. The section titled ‘Simple Clauses’ in Dryer and Haspelmath (2013; see section ‘G’ in Haspelmath et al. 2005) contains 24 WALS features (98A–121A), subdivided into four subsections: (i) Simple Clauses (WALS 98A–105A); (ii) Valence and Voice (WALS 106A–111A); (iii) Negation and Questions (WALS 112A–116A); and (iv) Predication (WALS 117A–121A). In addition to these 24 features, three features related to negation 143A, 143E and 144A which were included later in Dryer and Haspelmath (2013) are also compared with other features related to negation (iii) in this chapter; hence, the total number of features compared in this chapter is 27. Note also that the first subsection containing eight features related to case and agreement (WALS 98A–105A) are not given a title; these features are simply listed under the title of the section (see Haspelmath et al. 2005a).

Section 3.2 compares features in the two languages related to case marking on the core arguments of the verb and person marking on the verb. Section 3.3 deals with features related to valency and operations reducing valency (e.g., reflexives, passive and antipassive) and those increasing valency (e.g., applicative and causative). Features related to clausal negation and polar questions in the two languages are compared in Section 3.4. This includes the ways in which negation is encoded, the position of the negative element with regard to the other constituents of a clause and the methods by which languages indicate that an utterance is a question. In Section 3.5, features related to predication in the two languages are compared; included in this section are questions regarding how predicative possession, and nominal, adjectival and locational predicates are encoded; whether a zero copula is allowed in nominal predicates; and the construction used to make comparison. In Section 3.6, the main findings of the comparison of the features under Sections 3.2–3.5 are used to establish the typological profiles of those features. The implications arising from the findings are then discussed with the aim of studying the nature of possible changes that have occurred in these features of Sinhala owing to its contact with Tamil.

3.2 Features related to case and agreement: a comparison

3.2.1 Alignment of case marking of full noun phrases

This feature (WALS 98A, Comrie 2013a) is concerned with the ways in which core argument noun phrases are marked by means of morphological case or adpositions to indicate which particular core argument position they occupy. There are six types of alignment of case: 1. neutral; 2. nominative-accusative (standard); 3. nominative-accusative (marked nominative); 4. ergative-absolutive; 5. tripartite; and 6. active-inactive. Tamil and Sinhala have standard nominative-accusative marking, in which either just the accusative or both nominative and accusative are marked (value 2), as in (1):

(1) T. sri pillai-ai nulli-n-aan
Sri child-ACC pinch-PST-3SGM
S. siri laməya-wə keniththuwa
Siri child-ACC pinch.PST

‘Sri/Siri pinched the child.’

3.2.2 Alignment of case marking of pronouns

This feature (WALS 99A, Comrie 2013b) concerns the alignment of case marking of pronouns. Pronouns are treated separately because in many languages (including English) pronouns have a different case marking system from full noun phrases (see WALS 98A above). There are seven values for this feature: values 1–6 are identical with 1–6 in WALS 98A (3.2.1), while value 7 is referred to as ‘none’.

As it happens, there is no difference between the alignment of case marking of full noun phrases and that of pronouns in Tamil and Sinhala: both have the same system, nominative-accusative (standard) (value 2), as in (2):

(2) T. naan avan-ai anuppi-n-een
I he-ACC send-PST-1SG
S. mamə eyaa-wə yæwwa
I he-ACC send.PST

‘I sent him.’

3.2.3 Alignment of verbal person marking

This feature (WALS 100A, Siewierska 2013a) is concerned with the alignment of person markers, that is, how the two core arguments of the transitive verb, A(gent) and P(atient), align with the sole core argument, S(ubject), of the intransitive verb. There are six possible values: 1. neutral alignment (no verbal person marking); 2. accusative alignment; 3. ergative alignment; 4. active alignment; 5 hierarchical alignment; and 6. split alignment. Both Tamil and Sinhala have accusative alignment (value 2) in which A and S have the same treatment while P has a different treatment, as in (3a) (transitive) and (3b) (intransitive):[36]

---

[36] The title, ‘Alignment of Verbal Person Marking,’ is misleading in that it does not refer to the marking of agreement features on the verb (see WALS 102A in 3.2.5 below). This feature is about the alignment of case, that is, the grammatical relation between the arguments of the verb.

(3) a. T. polis aval-ai pidi-th-aarkal
Police she-ACC catch-PST-3PL
S. polis eyaa-wə alləgaththa
Police she-ACC catch.PST

‘The police caught her.’

(3) b. T. aval oodi-n-aal
She run-PST-3SGF
S. eyaa diwwa
She run.PST

‘She ran.’

3.2.4 Expression of pronominal subjects

This feature (WALS 101A, Dryer 2013q) is about simple sentences with a pronominal subject in which the only expression of the subject is a pronominal morpheme, such as an independent pronoun or an affix on the verb. The six values of this feature are: 1. pronominal subjects are expressed by pronouns in subject position that are normally if not obligatorily present; 2. pronominal subjects are expressed by affixes on verbs; 3. pronominal subjects are expressed by clitics with variable host; 4. pronominal subjects are expressed by subject pronouns that occur in a different syntactic position from nominal subjects; 5. pronominal subjects are expressed only by pronouns in subject position, but these pronouns are often left out; and 6. more than one of the above types with none dominant. This feature is different in Tamil and Sinhala: in Tamil, pronominal subjects are expressed by affixes on verbs (value 2), as in (4T), whereas in default Sinhala sentences with finite verbs, the only expression of pronominal subjects involves pronouns in subject position, but such pronouns are optional (value 5), as in (4S):

(4) T. (avan) puthakath-ai vaasi-th-aan
He book-ACC read-PST-3SGM
S. (eyaa) pothə kiyewwa
He book read.PST

‘He read the book.’

Both have independent pronouns as in (4), while allowing pronominal subjects to be null, i.e., they are pro-drop languages. This feature is discussed more fully in Chapter 5. Tamil and Sinhala have different values for this feature in that in Tamil, the pronominal subject is expressed also by ‘morphemes coding semantic or grammatical features of the pronominal subject' (Dryer 2013q) in the verb, whereas in Sinhala, these features of the pronominal subject are not encoded in default finite verbs (see 6.3 for details).

3.2.5 Verbal person marking

This feature (WALS 102A, Siewierska 2013b) is concerned with the number and identity of the arguments of a transitive clause which displays person marking on the verb. There are five values for this feature: 1. no person marking of any argument; 2. person marking of only the A argument; 3. person marking of only the P argument; 4. person marking of the A or P argument; and 5. person marking of both the A and P arguments. This feature too is different in Tamil and Sinhala: in Tamil, the person of A is marked on the verb (value 2), as in (5T), whereas in Sinhala, there is no person marking of any argument (value 1), as in (5S):

(5) T. mohan kriket vilaiyaadu-v-aan
Mohan cricket play-FUT-3SGM
S. mohan kriket sellaŋkərai
Mohan cricket play.FUT

'Mohan will play cricket.'

3.2.6 Third person zero of verbal person marking

This feature (WALS 103A, Siewierska 2013c) is about 'the distribution of zeroes among verbal third person markers of the sole argument of an intransitive clause (i.e., only of the S argument).' In effect, this feature is about whether or not there is person marking of S. The six values of this feature are 1. no person marking of the S; 2. no zero realization of third person S forms; 3. zero realization of some third person singular S forms; 4. zero realization of all third person singular S forms; 5. zero realization only of all third person S forms/no third person S forms; and 6. zero realization only of third person non-singular S forms. The feature values of Tamil and Sinhala are different in that Tamil has no zero realization of third person S forms (value 2), that is all the third person S forms that the language distinguishes are overtly marked on the verb, while—as a logical consequence of what was shown in 3.2.5—Sinhala has no person marking of the S (value 1), as in (6):

(6) T. kamal alu-th-aan
Kamal cry-PST-3SGM
S. kamal ænduwa
Kamal cry.PST

‘Kamal cried.’

3.2.7 Order of person markers on the verb

This feature (WALS 104A, Siewierska 2013d) is concerned with the order of A and P person markers on the verb relative to each other. Of the five values of this feature—1. A and P do not, or do not both, occur on the verb; 2. A precedes P; 3. P precedes A; 4. both orders of A and P occur; and 5. A and P are fused—the value of Tamil and Sinhala is ‘A and P do not, or do not both, occur on the verb’ (value 1). Note that value 1 covers languages that exhibit no verbal person marking (Sinhala), those which have marking of only one of the transitive arguments (Tamil) and those which allow both arguments to be marked but not at the same time. Thus, there is a difference between the two languages: Tamil has marking of one of the transitive arguments, namely A, as in (7T), whereas Sinhala exhibits no verbal person marking at all, as in (7S) below:

(7) T. sitha piramb-ai muri-th-aal
Sitha cane-ACC break-PST-3SGF
S. sitha weewælə kæduwa
Sitha cane break.PST

‘Sitha broke the cane.’

3.2.8 Ditransitive constructions: the verb ‘give’

This feature (WALS 105A, Haspelmath 2013a) is about the coding properties of the two objects in a ditransitive construction involving the verb ‘give’. There are four values for this feature: 1. indirect-object construction; 2. double-object construction; 3. secondary-object construction; and 4. mixed. In 1, the theme of a ditransitive verb, that is, the argument expressing what is given, is coded like a monotransitive patient—referred to as direct object—and the recipient is coded differently, referred to as indirect object. In 2, both the theme and the recipient of the ditransitive verb are coded like a monotransitive patient. In 3, the recipient of the ditransitive verb is coded like a monotransitive patient—referred to as primary object—whereas the ditransitive theme is coded differently; it is referred to as secondary object. The value of Tamil and Sinhala is ‘indirect-object construction’ (value 1), as in (8):

(8) T. kala avan-itta thirapp-ai kudu-th-aal
Kala he-LOC key-ACC give-PST-3SGF
S. kala eyaa-tə yathurə dunna
Kala he-DAT key give.PST

‘Kala gave the key to him.’

Note that the theme in (8S) lacks overt case marking—but in this respect, it is completely similar to a monotransitive patient; as mentioned in 2.3.1.2, in Sinhala, the accusative case is marked only on animate nouns, and not on inanimate nouns.

3.3 Features related to valence and voice: a comparison

3.3.1 Reciprocal constructions

This feature (WALS 106A, Maslova and Nedjalkov 2013) is concerned with ‘the different ways in which languages encode reciprocal situations, i.e., situations like ‘They love each other’. The reciprocal situation comprises at least two simple situations (e.g., *She loves him* and *He loves her*) […]’. The four values of this feature are 1. there are no non-iconic reciprocal constructions; 2. all reciprocal constructions are formally distinct from reflexive constructions; 3. there are both reflexive and non-reflexive reciprocal constructions; and 4. the reciprocal and reflexive constructions are formally identical. The value of this feature in Tamil and Sinhala is: all reciprocal constructions, like (9a), are formally distinct from reflexive constructions, like (9b) (value 2):

(9) a. T. sarojav-um kumar-um oruvar-ai-oruvar
Saroja-INCL Kumar-INCL one(HON)-ACC-one(HON) (RECP)
asi-(k)ko-nd-aarkal
scold.PTCP-get-PST-3PL
S. saroja-i kumar-ui denna-tə denna bænə gaththa
Saroja-INCL Kumar-INCL two-DAT two (RECP) scold.PTCP get.PST

‘Saroja and Kumar scolded each other.’

(9) b. T. ravi thann-ai(-ye) kuththi-ko-nd-aan
Ravi self-ACC-FOC (REFL) stab.PTCP-get-PST-3SGM
S. ravi thaman-tə(-mə) ænə gaththa
Ravi self-DAT-FOC (REFL) stab.PTCP get.PST

‘Ravi stabbed himself.’

3.3.2 Passive constructions

This feature (WALS 107A, Siewierska 2013e) is about passive constructions. Of the two values of this feature—1. there is a passive construction; and 2. there is no passive construction—Tamil and Sinhala have ‘a passive construction’ (value 1), as in (10). Note, however, that the passive is rarely used in speech, though it is used in writing in both Tamil and Sinhala:

(10) T. intha veedu-kal pirathama manthiriy-aal kodukk-a-pat-t-athu
These house-PL prime minister-INS give-INF-PASS-PST-3SGN
S. mee geval agramaathyə visin denu læbuwa
These house.PL prime minister INS give.PTCP get.PST

‘These houses were given by the prime minister.’

3.3.3 Antipassive constructions

This feature (WALS 108A, Polinsky 2013a) is concerned with the antipassive construction, which is a derived detransitivized construction with a two-place predicate, related to a corresponding transitive construction. The three values of this feature are 1. antipassive with patient-like argument left implicit; 2. antipassive with patient-like argument expressed as oblique complement; and 3. no antipassive. Tamil and Sinhala have no antipassive (value 3) constructions.

3.3.4 Applicative constructions

This feature (WALS 109A, Polinsky 2013b) is about constructions in which the number of object arguments selected by the predicate is increased by one with respect to the basic construction. The term applicative is restricted to those cases where the addition of the object is overtly marked on the predicate. By using an applicative marker, a predicate is made to take an (additional) object. The object added in the applicative construction is called the applied object, while the verb from which the applicative is formed is called the base which can be transitive or intransitive. There are eight values for this feature: 1. benefactive object only; both bases; 2. benefactive object only; transitive base only; 3. benefactive and other; both bases; 4. benefactive and other; transitive base only; 5. non-benefactive object only; both bases; 6. non-benefactive object only; transitive base only; 7. non-benefactive object only; intransitive base only; and 8. no applicative construction. Tamil and Sinhala applicative constructions have a benefactive object only and transitive base only (value 2), as in (11) (ex.76b, p. 175, Chandralal 2010 translated into Tamil (11T)):

(11) T. Ranjit chitra-kku poothal-ai thiranthu kudu-th-aan
Ranjit Chitra-DAT bottle-ACC open.PTCP give-PST-3SGM
S. Ranjit chitra-tə boothəlee ærəla dunna
Ranjit Chitra-DAT bottle open.PTCP give.PST

‘Ranjit opened the bottle for Chitra.’

3.3.5 Periphrastic causative constructions

This feature (WALS 110A, Song 2013a) is concerned with biclausal (a.k.a. syntactic or analytic) causative constructions in which causation is expressed by independent words instead of morphemes. Periphrastic causative constructions can be of two kinds: the sequential type and the purposive type. In the sequential type, the clause expressing the cause and that expressing the effect are juxtaposed strictly in that order, with or without a linking element between them. The purposive type also involves two clauses, one representing an event carried out for the purpose of realizing another event denoted by the other clause (hence the term purposive). As such, the clause of cause and that of effect do not need to occur in that order. The three values of this feature are 1. sequential type but no purposive type; 2. purposive type but no sequential type; and 3. both sequential type and purposive type. Since in Tamil and Sinhala the clause of cause does not follow the clause of effect, as in (12), they have value 2:

(12) T. teecher [pillaihal-ai vahupparai-ai koot-a] pan-nin-aar
Teacher students-ACC classroom-ACC sweep-INF make-PST-3SGM(HON)
‘The teacher made the pupils sweep the classroom.’

S. tiichər [lamay-lauwa panthikaaməre athugaa-wa] gaththa
Teacher students-INS classroom sweep-CAUS.PTCP get.PST
‘The teacher got the pupils to sweep the classroom.’

Song (2013b) notes that although the verb of effect and that of cause are adjacent to each other in a Tamil sentence like (12T), the causative construction it involves is not regarded as the compound type (see 3.3.6 below) because the verb of effect (the infinitive form) actually contains purposive marking *-a*. The presence of the purposive marking on the verb, according to him, indicates that the two verbs do not form a unit, and that the verb of effect is part of a subordinate clause of purpose. Chandralal (2010: 170) notes that *ə* in the causative morpheme *wə* (see 3.3.6 below) changes to *a* in the formation of the perfect participle base which indicates that the two verbs do not form a unit and the verb of effect is part of the subordinate clause.

3.3.6 Nonperiphrastic causative constructions

This feature (WALS 111A, Song 2013b) is about mono-clausal (a.k.a. morphological) causative constructions in which causation is expressed by morphemes. There are two types of nonperiphrastic causative construction: the morphological type and the compound type. The morphological type involves a morphological process which applies directly to a basic verb, whereas in the compound type, the causer's action is expressed by a separate verb instead of a morphological element, but that verb must be adjacent to the basic verb (cf. the periphrastic construction in (12) above). These two kinds give rise to four values of this feature: 1. no morphological type or compound type; 2. morphological type but no compound type; 3. compound type but no morphological type; and 4. both morphological type and compound type. Tamil and Sinhala have the 'morphological type but no compound type' (value 2), as in (13):

(13) T. appa mohan-ai kondu antha marath-ai vettu-**vi**-kkir-aar
Father Mohan-ACC INS that tree-ACC cut-CAUS-PRS-3SGM(HON)
S. thaaththa mohan-lauwa ee gaha kappə-**wə**-nəwa
Father Mohan- INS that tree cut-CAUS-PRS

'Father is getting Mohan to cut the tree.'

The two languages use a morphological process in the nonperiphrastic in that the causative nature of the event is expressed by a morphological unit, a suffix *-vi* in Tamil and *-wə* in Sinhala followed by tense and agreement suffixes in Tamil and tense suffix in Sinhala.

3.4 Features related to negation and questions: a comparison

3.4.1 Negative morphemes

This feature (WALS 112A, Dryer 2013r) is concerned with the nature of morphemes signalling clausal negation in declarative sentences. Of the six values of this feature—1. negative affix; 2. negative particle; 3. negative auxiliary verb; 4. negative word, unclear if verb or particle; 5. variation between negative word and affix; and 6. double negation—Tamil and Sinhala both have a negative particle (value 2), as in (14), based on the affirmative (12a) in 2.1.12:

(14) T. nandini poona kilamai pillaikal-ukku inglish padippi-kka illai
Nandini last week students-DAT English teach-INF NEG
S. nandini giyə sumaane lamay-tə ingriisi igænnuw-e nææ
Nandini last week students-DAT English teach.PST-NMLZ NEG

‘Nandini didn’t teach English to the students last week.’

In addition to this rather simple feature, there is of course more to say about negation in Tamil and Sinhala. In 3.4.2 and 3.4.3 below, two further features related to negation, WALS 143A and 144A are analyzed. These two features, the last two of the 144 features, which were later included in Dryer and Haspelmath (2013) are classified under ‘word order’. Since they are related to further aspects of negation, they are analyzed in the next two subsections for convenience.[37]

3.4.2 Order of negative morpheme and verb

Dryer (2013s) divides WALS 143 into six features, 143A–G, of which only two features, WALS 143A and E, are relevant to the two languages under study.[38] WALS 143A, ‘order of negative morpheme and verb’, is concerned with ‘the order of negative morpheme and lexical verb for negative morphemes that code simple clausal negation’. Of the 17 (not stated owing to constraints of space) values of WALS 143A, Tamil and Sinhala have VNeg (value 2), as can be seen in (14) in 3.4.1.

WALS 143E, ‘preverbal negative morphemes’ (Dryer 2013), is about whether languages allow preverbal negative morphemes, either separate words or prefixes. The four values of this feature are 1. preverbal negative word; 2. negative prefix; 3. both preverbal negative word and negative prefix; and 4. no preverbal negative morpheme. Tamil has ‘no preverbal negative morpheme’ (value 4), whereas Sinhala frequently uses a ‘negative prefix’ (value 2) (see also 2.3.1.1), as in (15aS) and the default negative form in the two languages is given in (15b); see also (14):

(15) a. S. leela enə sumaane koləmbə-tə no-yai
Leela coming week Colombo-DAT NEG-go.FUT
‘Leela will not go to Colombo next week.’

---

[37] Since these two features were introduced later, they are found only in Dryer and Haspelmath (2013), and not in Haspelmath et al. (2005b), the print version of the former.

[38] Only the features relevant to the two languages are analyzed here and 3.4.3 below (for details about other features and their values, see Dryer 2013s, 2003t).

(15) b. T.	leela	vaara	kilamai	kolumbu-kku	poo-ha	illai
	Leela	coming	week	Colombo-DAT	go-INF	NEG
S.	leela	enə	sumaane	koləmbə-tə	yann-e	nææ
	Leela	coming	week	Colombo-DAT	go.FUT-NMLZ	NEG

‘Leela will not go to Colombo next week.’

3.4.3 Position of negative morpheme with respect to subject, object, and verb

This feature (WALS 144A, Dryer 2013t) which expands WALS 143A is concerned with the order of negative morpheme with respect to the subject, object and verb. Dryer (2013t) divides this feature into 25 features, WALS 144A–144Y, of which only WALS 144A is analyzed in this chapter. According to (Dryer 2013t), for languages which employ a single negative word and a single position for the negative word in negative clauses, there are 24 logically possible types—four for each order of subject, object and verb. Of these 24 logically possible types, fifteen are attested. WALS 144A has 21 Types of which Types 1–15 are the ‘attested word orders involving a single negative word, the subject, the object, and the verb, while Types 16–21 are those which don’t involve just a single negative word and a single word order in negative clauses (see Dryer 2013t for the twenty one values). Of the 21 Types of WALS 144A, Tamil and Sinhala are of SOVNeg, as in (14/15b) above. In summary, negative clauses in Tamil and Sinhala, by default, are of the single word, single position type, ‘SOVNeg’ (see (14/15b) above).

3.4.4 Symmetric and asymmetric standard negation

This feature (WALS 113A, Miestamo 2013a) concerns symmetric (Sym) and asymmetric (Asy) standard negation. Standard negation can be defined as the basic way (or ways) a language has for negating declarative verbal main clauses. It, therefore, excludes the negation of existential, copular or non-verbal clauses, the negation of subordinate clauses, and the negation of non-declarative clauses like imperatives. In Sym standard negation, the structure of the negative is identical to the structure of the affirmative, except for the presence of the negative marker(s), while in Asy standard negation, apart from the negative marker(s), the structure of the negative differs from the structure of the affirmative in various other ways too. Affirmative and negative structures can be symmetric or asymmetric in two ways: there can be (a)symmetry either between the affirmative and negative constructions, or between the paradigms that the affirmative and negative constructions form.

There are three values for this feature: 1. Type Sym—languages where standard negation is always symmetric; 2. Type Asy—languages where standard negation is always asymmetric; and 3. Type SymAsy—languages where both symmetric and asymmetric standard negation is found. Tamil has Type Asy (value 2), whereas Sinhala has Type SymAsy (value 3). The default negative form (14/15b) of an affirmative declarative clause in the two languages is of the TypeAsy kind. In addition, the preverbal negative construction (15aS) in Sinhala is of the TypeSym, as (16bS) below; hence, the value is TypeSym/Asy:

(16) a. S. eyaa   yai
He   go.FUT
'He will go.'

(16) b. S. eyaa   no-yai
He   NEG-go.FUT
'He won't go.'

(16bS) is symmetric in that the structure of the affirmative (16aS) and the negative (16bS) is the same, except that the negative prefix *no-* is added to the negative construction. Tamil has a similar kind of negative construction, as (17bT):

(17) a. T. avan   poo-v-aan
He   go-FUT-3SGM
'He will go.'

(17) b. T. avan   poo-h-aan
He   go-FUT.NEG-3SGM
'He won't go.'

However, (17bT) is asymmetric in that the affirmative future suffix -*v* is replaced by -*h* in the negative construction. The default negative form of the affirmative declarative clause (16a/17a) in the two languages is of TypeAsy, as (18):

(18) T. avan   poo-(h)a   illai
He   go-INF   NEG
S. eyaa   yann-e   nææ
He   go.FUT-NMLZ   NEG

'He won't go.'

When the negative particle is attached to the lexical verb in the affirmative sentence (17a), it loses its finiteness and the agreement features in Tamil (18T), while the finite form of the verb in (16a) is changed into a tensed nominalized form in Sinhala (18S).

3.4.5 Subtypes of asymmetric standard negation

This feature (WALS 114A, Miestamo 2013b) is about asymmetric negation. It, therefore, applies only to languages where asymmetric negation is found (Types Asy and Type SymAsy in 3.4.4). Three subtypes of asymmetry are distinguished according to whether the asymmetry is connected to:

(i) the finiteness of verbal elements (A/Fin), that is, the negative construction adds a new finite element (finite verb) to the clause, and the lexical verb becomes non-finite and/or subordinate to the added finite element;
(ii) the marking of reality status of events (A/NonReal)—the negative clause isobligatorily marked by a category expressing non-realization;
(iii) the marking of verbal categories in some other ways (A/Cat),especially marking of grammatical categories (such as tense, aspect, mood, person, number, etc.) under negation.

Since a language can have different negative structures in different contexts, different subtypes of asymmetric negation may be found within one and the same language. Therefore, this feature involves three further types of language where two of the three subtypes of asymmetric negation are combined. The seven values of this feature are 1. in finiteness: Subtype A/Fin; 2. in reality status: Subtype A/NonReal; 3. in other grammatical categories: Subtype A/Cat; 4. in finiteness and reality status: Subtypes A/Fin and A/NonReal; 5. in finiteness and other grammatical categories: Subtypes A/Fin and A/Cat; 6. in reality status and other grammatical categories: Subtypes A/NonReal and A/Cat; and 7. non-assignable (no asymmetry found, that is, languages which have Type Sym). Tamil has Subtypes A/Fin and A/Cat (value 5) whereas Sinhala has only Subtype A/Fin (value 1).

The default negative form (18) in Tamil and Sinhala is of Subtype A/Fin (value 1), in which negation affects the finiteness of verbal elements. Typically, the negative construction adds a new finite element (finite verb) to the clause, and the lexical verb becomes non-finite and/or subordinate to the added finite element.

Tamil has another negative construction, which is of Subtype A/Cat. In this construction, the negative form of the future modal is used, as in (19T) (the affirmative form is (17aT):

(19) T. avan poo-(h)a maatt-aan
He go-INF will.NEG-3SGM
‘He won’t go.’

In (19T) the future modal is not only marked for negation, but also for person, gender and number.

3.4.6 Negative indefinite pronouns and predicate negation

This feature (WALS 115A, Haspelmath 2013c) is concerned with whether negative indefinite pronouns like ‘nobody’, ‘nothing’, ‘nowhere’ may or may not co-occur with the ordinary marker of predicate negation. Of the four values of this feature—1. negative indefinites co-occur with predicate negation; 2. negative indefinites preclude predicate negation; 3. negative indefinites show mixed behaviour; and 4. negative existential construction—Tamil and Sinhala have negative indefinites which show mixed behaviour (value 3): i) the indefinites co-occur with predicate negation, as in (20a); and ii) in instances like *if*-clauses, polar questions or sentences which express possibility etc., they preclude predicate negation, as in (20b):[39]

(20) a. T. yaar-um vahuppu-kku var-a illai
Who-INCL class-DAT come-INF NEG
S. kauru-th panthiyə-tə aaw-e nææ
Who-INCL class-DAT come.PST-NMLZ NEG

‘No one came to class.’

(20) b. T. yaar-um vahuppu-kku var-a-laam
Who-INCL class-DAT come-INF-may
S. kauru-th panthiyə-tə endə puluwaŋ
Who-INCL class-DAT come.INF may

‘Anyone may come to class.’

3.4.7 Polar questions

This feature (WALS 116A, Dryer 2013r) is about the method a language uses to indicate that an utterance is a polar question (see also 2.2.12 above). There are seven values for this feature: 1. question particle; 2. interrogative verb morphology; 3.

---

[39] The clitics *-um* in Tamil and *-th* in Sinhala are glossed as INCL(usive) following Lehman (1989: 152) who notes that *-um* is added to a wh-word/phrase to transform the latter ‘into a referential expression with general or all-inclusive reference.’ In Sinhala too, the clitic *-th* added to the wh-word/phrase expresses general or all-inclusive reference. The clitics, *-um* in Tamil and *-th* in Sinhala, also function as conjunctive coordinators in the two languages (4.3.5).

question particle and interrogative verb morphology; 4. interrogative word order; 5. absence of declarative morphemes; 6. interrogative intonation only; and 7. no interrogative-declarative distinction. Tamil employs interrogative verb morphology, the clitic *-aa* (value 2), whereas Sinhala employs the question particle *-də* (value 1), as in (21):

(21) T. avan marath-ai vetti-n-aan-**aa**
He tree-ACC cut-PST-3SGM-Q
S. eyaa gaha kæpuwa **də**
He tree-ACC cut.PST Q

‘Did he cut the tree?’

3.5 Features related to predication: a comparison

3.5.1 Predicative possession

This feature (WALS 117A, Stassen 2013a) is about the various encoding options for predicative possession (alienable), that is, sentences in which ownership of a certain object (the possessed item) is predicated of a possessor, in a way that is illustrated by the English sentence *John has a motorcycle*. There are five values for this feature: 1. locational possessive; 2. genitive possessive; 3. topic possessive; 4.conjunctional possessive; and 5. have-possessive. Since the oblique marking on the possessor NP in the two languages has as its meaning the specification of a locational relation with the marker ‘to’, also classified as dative possessive (see Stassen 2013a for the classification of possessives), Tamil and Sinhala have a locational possessive (value 1), as in (22):

(22) T. mohan-itta oru kaar iru-kkir-athu
Mohan-LOC one car exist-PRS-3SGN
S. mohan gaavə kaar ekak thiyenəwa
Mohan LOC car one exist.PRS

‘Mohan has a car.’

3.5.2 Predicative adjectives

This feature (WALS 118A, Stassen 2013b**)** concerns the various options in the encoding of predicative adjectives, that is, items which predicate a property of the subject. This feature distinguishes between two kinds of languages: i) languages in which predicative adjectives are encoded in a way that is parallel to predicative verbs; and ii) languages in which the encoding of predicative adjectives and verbs is different. There are, therefore, three values for this feature: 1. predicative adjectives have verbal encoding; 2.

predicative adjectives have nonverbal encoding; and 3. predicative adjectives have mixed encoding. Tamil and Sinhala have nonverbal encoding (value 2), as in (23a); cf. (23b) in which predicative verbs occur:

(23) a. T. rani nalla vadivu
Rani very pretty
S. rani huŋgak lassəna-i
Rani very pretty-ASS

‘Rani is very pretty.’

(23) b. T. rani aadi-n-aal
Rani dance-PST-3SGF
S. rani nætuwa
Rani dance.PST

‘Rani danced.’

3.5.3 Nominal and locational predication

This feature (WALS 119A, Stassen 2013c) is about the relationship between the encoding of nominal and locational predicates, especially whether nominal predications (such as *John is a tailor*) and locational predications (such as *John is in Paris*) can or cannot be encoded by the same strategy. If the encoding strategy for locational predications is (or can be) used for nominal predications in a language, that language is called a ‘share-language’ and if the encoding strategies for the two constructions must be different, it is called a ‘split-language’. Of the two values for this feature, namely 1. split (different) and 2. shared (identical), Tamil and Sinhala are of the split type (value 1), as in (24a) (nominal predication) and (24b) (locational predication):

(24) a. T. ratha oru teecher
Ratha one(INDF) teacher
S. ratha guruvəriy-ak
Ratha teacher(F)-one(INDF)

‘Ratha is a teacher.’

(24) b. T. ratha ippa london-il irru-kki-raa
Ratha now London-LOC exist-PRS-3SGF(HON)
S. ratha thæŋ london-wələ innəwa
Ratha now London-LOC exist.PRS

‘Ratha is in London now.’

3.5.4 Zero copula for predicate nominals

This feature (WALS 120A, Stassen 2013d) is about zero copula encoding for predicate nominals. The question here is whether predicate nominals in a language always require an overt copula, or whether omission of the copula is allowed. There are, therefore, only two values for this feature: 1. zero copula is impossible; and 2. zero copula is possible. In Tamil and Sinhala, a zero copula is possible (value 2), as in (25) (see also (24a) above:

(25) T. kili oru paravai
Parrot one(INDF) bird
S. girəvaa kurull-ek
Parrot bird-INDF

‘A parrot is a bird.’

3.5.5 Comparative constructions

This feature (WALS 121A, Stassen 2013e) is concerned with comparative constructions. It involves three things: i) a predicative scale encoded as a gradable predicate, and two objects encoded as noun phrases (NPs), ii) an object of comparison (the comparee NP) and iii) the "yard-stick" of the comparison (the standard NP). A typical instance of a comparative construction in the languages of the world is the one similar to *John is taller than Lucy*, in which the NP following *than* is the standard NP. Based on the encoding of the standard NP, two types of comparative constructions are distinguished: (i) fixed-case comparatives and (ii) derived-case comparatives. In the former, the standard NP is always in the same case, regardless of the case of the comparee NP, whereas in the latter type, the standard NP derives its case assignment from the case of the comparee NP. The comparative constructions in Tamil and Sinhala are of the fixed type in that the standard NP is always in the accusative case in Tamil and in the dative case in Sinhala, as in (26):

(26) T. antha kinatt-ai vida intha kinaru aalam
That well-ACC than this well deep
S. ee lində-tə vadaa mee lində gæmburui
That well-DAT more.than this well deep

‘This well is deeper than that well.’

Each of the two types (i) and (ii) are further divided into two subtypes: the two subtypes of the fixed-case comparatives are the locational comparative and the exceed comparative; while the two subtypes of the derived-case comparatives are the conjoined comparative and the particle comparative. There are four values, based on these four kinds: 1. locational comparative; 2. exceed comparative; 3. conjoined comparative; and 4. particle comparative. In the locational comparative, the standard NP is invariably construed in a case form which also has a locational/adverbial function, whereas in the exceed comparative, the standard NP is constructed as the direct object of a transitive verb with the meaning 'to exceed' or 'to surpass'.[40] Tamil and Sinhala belong to the locational kind because the standard NP is not constructed as the direct object of a transitive verb in them. By contrast, the standard NP has a locational/adverbial function in that the standard NP is marked as the source of a movement, with a marker meaning 'from'; thus, the comparative constructions in the two languages can be classified as *From*-comparatives.

3.6 Discussion

In this chapter, 27 features of Tamil and Sinhala are compared in four sections (3.2–3.5) above. The implications for the possible contact-induced changes that Sinhala may have undergone which arise from the comparative analysis are discussed below.

3.6.1 Features related to case and agreement: implications of the comparison

The typological profiles of Tamil and Sinhala based on the comparative analysis of eight features related to case and agreement in 3.2 are given in table 3.1:

---

[40] Values (kinds) 3 and 4, namely conjoined comparative and particle comparative, are not discussed as they are not relevant to the comparative constructions in the two languages (see Stassen 2013e for details).

	WALS features	**Tamil(T)**	**Sinhala (S)**	**Similar/ Different**
3.2.1	Alignment of case marking of full noun phrases (98A)	nominative-accusative (standard)	nominative-accusative (standard)	similar
3.2.2	Alignment of case marking of pronouns (99A)	nominative-accusative (standard)	nominative-accusative (standard)	similar
3.2.3	Alignment of verbal person marking (100A)	accusative alignment	accusative alignment	similar
3.2.4	Expression of pronominal subjects (101A)	pronominal subjects are expressed by affixes on verbs	optional pronouns in subject position	different
3.2.5	Verbal person marking (102A)	only the A[/S] argument [is marked]	no person marking of any sort	different
3.2.6	Third person zero of verbal person marking (103A)	no zero realisation of third person	no person marking of any sort	different
3.2.7	Order of person markers on the verb (104A)	A and P do not, or do not both, occur on the verb	A and P do not, or do not both, occur on the verb	similar
3.2.8	Ditransitive constructions: the verb ‘give’ (105A)	indirect-object construction	indirect-object construction	similar

Table 3.1: Typological profiles of features related to case and person marking

Of the eight features that are compared in this section, the values of five are similar and those of three are different. With regard to alignment of case marking, Tamil and Sinhala are of the nominative-accusative (standard) kind. This value for the first feature (3.2.1) determines the values for the second (3.2.2) and third (3.2.3) features, which are also similar. With respect to case alignment, there is no difference between full NPs and pronouns: the two languages, by default, have the same case alignment, that is A(gent) and P(atient) in transitive clauses take nominative and accusative case respectively, while S(ubject) in intransitive clauses takes nominative case, as examples (1–3) in 3.2.1–3.2.3 show. The values of the next three features (3.2.4–3.2.6) are different in the two languages. These three values differ by virtue of a difference in one single phenomenon in the two languages, namely agreement. There are distinct agreement features marked on the verb in Tamil but not in Sinhala, as shown in examples (4–8) in 3.2.4–3.2.8. Special mention must be made of the value of the seventh feature (3.2.7):

Tamil and Sinhala share the same value because this value includes languages which exhibit no verbal person marking (like Sinhala) and those which have marking of only one of the transitive arguments (like Tamil). However, with respect to person marking on the verb, the two languages are different (see 3.2.4 – 3.2.6). The value of the last feature (3.2.8) in this subsection is the same in that the default order of the two objects is indirect object followed by the direct object in both languages.

Described below in 3.6.1.1 and 3.6.1.2 respectively are the case system and the agreement system which are important for the explication of the features discussed in 3.2.

3.6.1.1 Case system

The Structural case relates the semantic roles of the arguments of the verb to the syntactic structure of the clause. As shown in 3.2.3, in both Tamil and Sinhala, the nominative case is assigned to A, one of the two core arguments of a transitive verb and S, the only core argument of an intransitive verb, while the accusative case is assigned to P, the other core argument of the transitive verb. In the two languages, A and S do not have overt marking for nominative case. In Tamil, all NPs functioning as P have overt marking for accusative case, but in Sinhala only animate NPs that perform the role of P have overt accusative marking. Some scholars have claimed that Sinhala is of the ergative-absolutive kind.[41] However, Henadeerage (2002) and Chandralal (2010) dispel this claim and convincingly establish that Sinhala is a nominative-accusative language.

A brief introduction to morphological cases in the two languages is in order here. Tamil and Sinhala have a fairly distinct morphological case system which is richer in the former than in the latter. Morphological case is expressed via case suffixes or postpositions in the two languages. Schiffman (2004: fn. 1, p, 301) points out that ‘all Dravidian literary languages are described by native grammarians as having eight cases…’ (see also Sarma 1999 and Lehmann 1989). Sinhala too has eight cases, but two pairs, namely ablative/instrumental and genitive/locative are syncretic, that is, the same suffix is used to mark the two cases in each pair (Chandralal 2010). In Sinhala, case suffixes are added to the indefinite suffix if the noun is indefinite. In DPs with numerals, case suffixes are added to the numeral. In Tamil, the case morpheme is suffixed to the DP; it is added to the plural marker if the DP is plural.

In addition to these cases, postpositions are used in the two languages to express specific semantic functions. Krishnamurti (2003: 208) observes that ‘Case relations in

---

[41] For references, see Henadeerage (2002) and Chandralal (2010).

Dravidian are expressed either by bound morphemes or by grammaticalized nouns or verbs, called postpositions.' Noting that there is no clear cut distinction between case and postpositions in Tamil, Schiffman (2004: fn. 2, p. 302)) adds:

> There seems to be a somewhat universal notion that case is to be understood as consisting of those bound morphemes that do not occur elsewhere in the language, whereas postpositions are independent, non-bound free forms that cannot be attached directly to stems of nouns or pronouns but must follow some case marker. They supposedly can (in most instances in the Dravidian languages at least) be easily shown to be derived from nouns or verbs…

In Tamil the postposition -*aaka* grammaticalized from *aaku* 'become', is added to the dative case -*kku* to form the benefactive marker, while *irunthu* grammaticalized from the past participle form of *iru* 'be' or 'sit' is added to the locative case -*il* to form the ablative marker. Table 3.2 gives the main case suffixes and postpositions in Tamil and Sinhala, based on Lehmann (1989) and Sarma (1999) for Tamil, and Chandralal (2010) and Gair (2007) for Sinhala.[42]

---

[42] For convenience, ablative/instrumental and genitive/locative in Sinhala are shown separately, unlike in Chandralal (2010: 81). Also, only the postpositions that perform the function of cases are given; for others, see Lehmann (1989) for those in Tamil and Chandralal (2010) for those in Sinhala.

Case	Tamil		Sinhala	
	Cases	**Postpositions**	**Cases**	**Postpositions**
Nominative	morphologically unmarked			
Accusative	*-ai*		*-wə*	
Dative	*-kku*		*-tə*	
Locative	*-il*	*itta, idam* ‘to’ or ‘near’	*-ge/* *-ee*	*gaavə/langə* ‘to’ or ‘near’ with the dative marker *-tə*
Instrumental	*-aal*	*kondu*	*-gen* *-en/-in*	*atin/lauwa/visin* ‘by’
Sociative/ Comitative	*-oodu, -ooda*	*udan* ‘with’		*ekkə* ‘with’
Ablative		*(il)irunthu* ‘from’	*-gen/* *-en/-in*	*indan/indəla* ‘from’
Allative		*varai(kkum)/* *mattum* ‘till’		*turu/kaŋ* ‘till’
Genitive	*-udaiya,* *-(in)da, -ra* *-athu*		*-ge/* *-ee*	*atee* ‘at’
Benefactive		*-(kk)aaka*	*-tə*	
Vocative	*-ee*		*-oo*	

Table 3.2: Case suffixes and postpositions

Cases and postpositions in Tamil and Sinhala play the same syntactic and semantic roles, some of which are discussed below. The fact that these cases and postpositions in Sinhala have similar roles as the corresponding ones in Tamil indicates that the cases/postpositions in Sinhala may have been modelled on those in Tamil.

Paramasivam (1983: 151 cited in Lehmann 1989:31–35) notes that the Dative case marker in Tamil performs a wide range of functions. Sinhala too uses the dative case to express these meanings. Some of the functions (i–v) are illustrated below:

(i) Indirect object

The dative case marks the noun as the indirect object of a ditransitive verb, as in (27):

(27) T. nimal appa-(vu)kku vaahana-(th)ai kaat(i)-n-aan
Nimal father-DAT vehicle-ACC show-PST-3SGM
S. nimal thaatha-tə vaahənəyə pennuwa
Nimal father-DAT vehicle show.PST

‘Nimal showed the vehicle to his father.’

(ii) Goal of motion

The dative case is used when inanimate nouns are the goal of motion, as in (28):

(28) T. rani aaspathiri-kku poo-n-aal
Rani hospital-DAT go-PST-3SGF
S. rani ispirithaale-tə giyaa
Rani hospital-DAT go.PST

‘Rani went to the hospital.’

However, when animate objects are the goal of motion, the locative marker is used, as in (29):

(29) T. rani doctor-itta poo-n-aal
Rani doctor-LOC go-PST-3SGF
S. rani doctor langə-tə/ gaavə-tə giyaa
Rani doctor LOC-DAT LOC-DAT go.PST

‘Rani went to the doctor.’

Note that the locative marker in the two languages is grammaticalized from the postposition *kitta* in Tamil and *langə/gaavə* in Sinhala ‘near’. The fact that the locative marker -*itta* in Tamil is homophonous with the dative case -*tə* in Sinhala suggests that the latter may have been adapted from Tamil.

(iii) Purpose

The dative case marked on a noun can convey the purpose of an action, as in (30):

(30) T. vimala pakidi-kku pillai-ai kadi-th-aal
Vimala joke-DAT child-ACC bite-PST.3SGF
S. vimala vihiluwə-tə laməya-wə hæppuwa
Vimala joke-DAT child-ACC bite.PST

‘Vimala bit the child for fun.’

It is also used to express such semantic relations as cause, reason or purpose to a noun, as in (31) (exs. (7)–(9), p. 130: Chandralal 2010 translated into Tamil):

(31) a. T. saavu-kku kaaranam
Death-DAT reason
S. marənəyə-tə heetuwə
Death-DAT reason

‘The reason for the death’

(31) b. T. varutha-(thu)kku marunthu
Illness-DAT medicine
S. ledee-tə behethə
Illness-DAT medicine

‘The medicine for the disease’

(31) c. T. pirachanai-kku theervu
Problem-DAT solution
S. prashne-tə visəndumə
Problem-DAT solution

‘Solution for the problem’

The dative case suffix is also used as a subordinator for conveying causal/ purpose relation between the subordinate clause and the main clause in complex sentences, as in (32):

(32) T. nimal computer-ai thiruthi kudu-th-athu-kku
Nimal computer-ACC repair.PTCP give-PST-NMLZ-DAT
avan kaasu kudukk-a venum
he money give-INF should/need
S. nimal computer ekə hadəla dunna-tə
Nimal computer one repair.PTCP give.PST-DAT
eyaa gevandə oone
he pay.INF should/need

‘He needs to/must pay Nimal for repairing the computer.’

(iv) Point in/duration of time

The dative case marked on nouns related to time expresses either a point in time or duration of time, as in (33):

(33) T. kootam naalu mani-kku thodangi irandu manithiyala-(th)ukku
Meeting four hour-DAT start.PTCP two hour-DAT
nada-nth-athu
happen-PST-3SGN
S rasweema hathərə-tə pataŋ araŋ pæyə dekək-(ə)tə pævæththuna
Meeting four-DAT start.PTCP take.PTCP hour two-DAT happen.PST

‘The meeting started at two o’clock and lasted for two hours.’

(v) Proportion

The dative case marked on temporal nouns expresses “the concept of proportion[/frequency] as ‘per’ or ‘a’ in English do” (Lehmann 1989: 34), as in (34):

(34) T. intha marunth-ai oru naalai-kku rendu neram
This medicine-ACC one(INDF) day-DAT two time
edu-ngo
take-2SG(HON)(IMP)
S. mee behethə davəsəkə-tə de særəy-ak ganndə
This medicine one day-DAT two time-INDF take.2(IMP)

‘Take this medicine two times a day.’

In addition to these functions, the dative case is marked on the subject to convey the experiencer role in dative subject constructions (4.3.11). These are the functions the dative case marker performs in the two languages.

In addition to their genitive markers, Tamil and Sinhala use the locative markers -(*k*)*itta* and *-langə/-gaavə* respectively to express possession, as in (35) (see also (22) in 3.5.1):

(35) T. avan-itta niraya kaasu iru-kk-uthu
He-LOC a lot money be-PRS-3SGN
S. eyaa langə godak salli thienəwa
He LOC a lot money be.PRS

‘He has a lot of money.’

In Sinhala, the dative case suffix *-tə* is also used to express possession.

The ablative marker in the two languages is used to express the starting point of the place and time of an action, and the allative marker is used to express the end point of only the time of an action, as in (36a), whereas the dative suffix is used to express the end point of the place of the action, as in (36b):

(36) a. T. avarkal naalu mani ilirunthu aaru mani varai cricket
They four o’clock ABL six o’clock ALL cricket
vilaiaadu-kir-aarkal
play-PRS-3PL
S. eyaala hathəre indəla hayə wenə kaŋ cricket
They four ABL six (o’clock) become.PTCP ALL cricket
selaŋkərənəwa
play.PRS

‘They play cricket from four till six.’

(36) b. T. naangal town ilirunthu kovil-(u)kku oodi-n-oom
We town ABL temple-DAT run-PST-1PL
S. api taume indəla kovile-tə diwwa
We town ABL temple-DAT run.PST

‘We ran from town to the temple.’

Note that in (36a) the allative marker indicates the end point of the time of action, whereas (36b) the dative marker indicates the end point of the place of action. It is also important to note that the dative marker in Sinhala is extended beyond the contexts in which the dative marker in Tamil is used, e.g., some adverbs are formed by adding the dative suffixe -*tə* to adjectives, as in -*tə* added to the adjective *hondə* ‘good’ to form *hondətə* ‘well’, and others like *lassənətə* ‘beautifully’ *rasətə* ‘tastily’ (Chandralal 2010: 54). This kind of generalization is another process observed in intense language contact situations.

The fact that similar cases and postpositions express the same semantic relations in the two languages, as the above discussion shows, clearly indicates that Sinhala has modelled its cases/postpositions on the corresponding ones in Tamil. Note also that some of the postpositions are derived in the same way in the two languages, that is, they grammaticalized from the same lexical items which demonstrates replica grammaticalization, one of the two kinds of contact-induced grammaticalization.

What is crucial for the features under discussion is that surface characteristics, such as morphological case marking on NPs do not provide sufficient evidence to identify grammatical relations like subject and object in Sinhala (Chandralal 2010). Failure to appreciate this, together with the presence of a pervasive dichotomy between volitive and involitive verbs (see 4.3.9), may be behind the claim that Sinhala has ergative-absolutive case alignment. Thus, Sinhala subjects can be marked not only for nominative but also for dative, accusative and instrumental case, as in (37a-c) (adapted from exs. (1a–c.), pp. 121–122, Chandralal 2010):

(37) a. S. kumar-tə nætəvuna
Kumar-DAT dance.INVL.PST
'Kumar danced involuntarily.'

(37) b. S. miniha-wə vaahənə-yen eliyə-tə visiuna
Man-ACC vehicle-ABL out-DAT throw.INVL.PST
'The man got thrown off from the vehicle.'

(37) c. S. rani-athin veeduruwə binduna
Rani-INS glass break.INVL.PST
'Rani accidentally broke the glass.'

Note that (37a–cS) resemble constructions belonging to the split ergative/absolutive system found in NIA languages because the subjects are of the non-nominative kind.[43] It is important to note, however, that accusative subjects in Sinhala occur only in intransitive sentences like (37bS). As Chandralal (2010:122) notes, in clauses describing such involuntary events, the subjects are 'non-nominative subjects bearing the semantic roles of Experiencer, Undergoer, and Accidental agent.' He adds that the internal arguments of these verbs remain unaffected. This, according to him, shows that 'morphological case on DPs/NPs cannot be taken as reliable proof of syntactic subjecthood in Sinhala.' Chandralal (2010:122) is of the view that it is these types of facts relating to the appearance of accusative marking on subjects in intransitive sentences, like (37bS), as well as objects in transitive sentences, that have made scholars claim that Sinhala, at least with regard to its colloquial variety, is an ergative-absolutive language.

The above discussion shows that Sinhala is a nominative-accusative language. In this regard it is similar to Tamil, which is of the nominative-accusative kind, but diverges from many of the other languages of the NIA family, which are of the tripartite/ergative-absolutive kind, e.g., Hindi (see Appendix). Note that instances like those in (37) are different from those in some NIA languages, e.g., Hindi, in which the tripartite/ergative-absolutive system occurs only in the perfective aspect, while the nominative-accusative system occurs in the imperfective aspect (Pashto too has a similar system, see Huang (1984)). Those in (37) in Sinhala can occur both in the perfective and imperfective aspects. The non-nominative subjects triggered by involuntary events like those in (37) may be assumed to be the remnants of an ergative-absolutive system that Sinhala may have had before it came into contact with Tamil. If

---

[43] The dative subject in (37a) is different from the dative subjects the two languages allow with distinct kinds of verbs (see 4.3.11).

these are remnant NIA features, then, Sinhala may have changed from ergative-absolutive or tripartite case alignment to nominative-accusative case alignment due to its contact with Tamil. Masica (1991: 343–344) argues that ‘The total loss of the ergative construction, that is, its replacement by a nominative-accusative construction, […] has occurred, e.g., in Standard Bengali, Oriya, and Sinhalese […].’

3.6.1.2 Agreement system

Agreement between subject and verb (SV), that is, such features of the subject as Person Number-Gender (PNG) agreeing with the finite verb of the sentence is the most important difference between Tamil and Sinhala. The former has agreement, whereas the latter has no agreement; thus, Sinhala is less like Tamil and more like Malayalam, another Dravidian language, which too does not have SV agreement. In Tamil, the verb of an affirmative finite clause ‘agrees with the nominative subject for person and number in all three persons, and for gender in the third person [singular]’ (Sarma 1999:10). In Sinhala, the situation is somewhat more complicated: ‘Spoken Sinhala verbs do not show inflectional agreement for any category in contradiction to the literary verb, which shows agreement for person, number and gender’ (Gair 2007: 866; see also 2005). In the two languages, the finite verb typically marks the end of a sentence: in Tamil, the PNG features of a proper noun or pronoun are marked on the finite verb, whereas in Sinhala, these features of a proper noun or pronoun are not marked on the finite verb. Tables 3.3 and 3.4 show agreement between subject pronouns and finite verbs in Tamil and Sinhala (see Lehmann (1989) and Sarma (1999) for Tamil, and Chandralal (2010) and Gair (2007) for Sinhala).[44]

---

[44] In addition to the verb forms given in tables 3.3 and 3.4, some varieties of Tamil have developed verb forms which the respective pronominals are suffixed to, e.g., *va-nth-*(*a*)*naan* ‘come-PST-1SG’; *va-nth-*(*a*)*naangal* ‘come-PST-1PL’; *va-nth-*(*a*)*nee* ‘come-PST-2SG; *va-nth-aval* ‘come-PST-3SGF’; *va-nth-athukal* ‘come-PST-3PL (NON-HUM).

Tamil			**Sinhala**		
Pronouns		**verb form: 'came'**	**Pronouns**		**verb form: 'came'**
naan	I	*va-nth-een* come-PST-1SG	*mamə*	I	*aawa* come. PST
naangal *naam*	We	*va-nth-oom*	*api*	We	
nee	You (SG informal)	*va-nth-aai*	*oyaa*	You (SG)	
neer	You (SG)	*va-nth-eer*	*oyaa-la*	You (PL)	
neengal	You (PL; also SG, HON)	*va-nth-eerkal*			

Table 3.3: Subject verb agreement between first and second person pronouns and finite verbs

Tamil			**Sinhala**		
Pronouns		**verb form: ‘came’**	**pronouns**		**verb form: ‘came’**
avan *ivan*	‘He’ (HUM)	*va-nth-aan* come-PST-3SM	*eyaa* *arəyaa* *meyaa*	‘S/he’ (HUM)	*aawa* come. PST
aval *ival*	‘She’ (HUM)	*va-nth-aal*			
avar *ivar*	‘S/he’ (HUM; HON)	*va-nth-aar*			
avaa *ivaa*	‘She’ (HUM; HON)	*va-nth-aa*			
avarkal *ivarkal*	‘They’ (HUM)	*va-nth-aarkal*	*eyaala* *areyəla* *meyaala*	‘They’ (HUM)	
athu/ *ithu*	‘It’ (NON-HUM)	*va-nth-athu*	*eekaa* *arəkaa* *ookaa* *uu/aru* *meekaa* *muu*	‘It’ (NON-HUM male; also HUM male derogatory)	
			eeki *arəki* *ooki* *meeki*	‘It’ (NON-HUM female; also HUM female derogatory)	
			eekə *arəkə* *ookə* *meekə*	‘It’ (INAN)	
avai- (*kal*) *ivai-* (*kal*)	‘They’ (NON-HUM, INAN; also HUM)	*va-nth-ana*	*ekung/eung* *arəkung* *ookung* *ung/arung* *meekung* *mung/meung*	‘They’ (NON-HUM; also HUM derogatory)	
athukal *ithukal*	‘They’ (NON-HUM)	*va-nth-athukal*	*eewa* *arəwa* *oowa* *meewa*	‘They’ (INAN)	

Table 3.4: Subject verb agreement between third person pronouns and finite verbs

Sinhala verbs inflect for tense, past and non-past, which is identified by the final bound morphemes: the non-past suffix -*nə* and the suffix for the indicative mood -*wa* (or -*a*), as in *duwə-nə-wa* ‘run’ (Chandralal 2010: 52). Modern spoken Sinhala verbs, according to (Chandralal 2010: 53); also Gair (1998b, 2007), are classified into three kinds according to their derivational morphological composition, especially the vowel with which the stems of verbs end: Class I (*ə*-ending stems), e.g., *kapə-nə-wa* ‘cut’ which includes the largest number of verbs, both transitive and intransitive; Class II (*i-ending* stems), e.g., *adi-nə-wa* ‘draw/pull’ which includes verbs which are either transitive or intransitive; and III (*e*-ending stems) e.g., *ide-nə-wa* ‘ripen’ which includes verbs which are intransitive. The formation of the past tense verb is more complex in that it involves different changes which vary according to the verb class or the conjugation type (see Chandralal 2010: 67–70). Further, Chandralal (2010: 77) distinguishes two derivational suffixes—the passive (inactive)/intransitive suffix *-e* and the causative suffix *-wə* (see (13) in 3.3.6)—whose addition to the verbal root makes it a different verb stem.

Cutting across the three kinds of verbs classified based on stem vowels in Sinhala is a function-based taxonomy of three types, namely Active, Passive and Causative to which every verb is considered to belong (Chandralal 2010).[45] For example, the three forms of a verb like ‘cry’ are: *andə-nə-wa* ‘cry’ active; *ænde-nə-wa* ‘get to cry’ passive; and *andə-wə-nə-wa* ‘cause to cry’. Note that in the passive form the stem ends with suffix *-e*, while *-wə* is suffixed to the stem in the causative form. However, every verb does not belong to all three types, while there may be verbs belonging to one or two types; hence there can be ‘three-type’, ‘two-type’ and ‘single-type’ verbs, e.g. a verb like ‘fall’ is a two-type verb which belongs to the Passive type *wæte-nə-wa* ‘fall’ and Causative type *wattə-nə-wa* ‘cause to fall’.

Semantically, the verbs in Sinhala are of two types, namely dynamic/active-type verbs and stative/processive-type verbs. The Active and Causative types belong to the dynamic/active category, whereas the Passive type (see paragraph above) belongs to the stative/processive category. With regard to the three kinds of verbs with the three stem vowel endings, the dynamic/active category includes *ə*-ending stems and *i*-ending stems, while the stative/processive category incorporates *e*-ending stems. Chandralal (2010: 53) notes that this categorization ‘has a direct influence on the event and aspectual interpretation of clauses and makes explicit the role of the agent’s control, intentionality and volition in event presentation’ (see 4.3.9, for volitive and involitive verbs in Sinhala).

In contrast to cases in Sinhala (3.6.1.1) which are suffixed to nouns just as in agglutinative languages, the formation of different verb forms in Sinhala involves stem variation or suppletion as in languages with fusional morphology. As the discussion

---

[45] Chandralal (2010: 77) points out that there is yet another independent form known as Causative-Passive which is obtained by passivizing the causative form.

shows, the differences between the verb forms are morphophonemic. In Tamil, which is consistently agglutinative, distinct tense, PNG suffixes are added to the verb stems to form finite verbs. The finite verb forms and their negative forms in Tamil and Sinhala are given in table 3.5 (see also tables 3.3 and 3.4):

Form	**Tamil**		**Sinhala**		
Base 'do'	***sei***		***kərə***		
	Positive	**Negative**	**Positive**	**Negative**	
Past	*sei-th-aan*	*seiy-a-(v)illai*	*kəraa/* *kəlaa*	*nokəraa/* *nokəlaa*	*kerae nææ/* *kəlae nææ*
Present	*sei-yi-raan*		*kərənəwa*	*no-kərənəwa*	*kəranne nææ*
Future	*sei-v-aan*		*kərai*	*nokərai*	
Imperative	*sei*	*seiyaathe/* *seiyaamal-iru/* *seiya veendaam*	*kərandə/* *kərannə*	*nokərandə/* *nokərannə*	*kəran næthuwə ində/* *kərandə epa*

Table 3.5: Finite verb forms

Observing that in Dravidian languages auxiliaries occur after an inflected main verb (infinitive or past/non-past participle), Krishnamurti (2003: 373–374) distinguishes two kinds of auxiliaries: (i) those that change the argument structure (valency) of the main verb such as passive, transitive-causative, reflexive/reciprocal and benefactive (applicative) auxiliaries; and (ii) those that preserve the valency but express other grammatical relations like aspect, intensity (effective involvement), modality (expressing necessity, possibility, probability, ability etc.) and mood (attitudinal). Note that the transitive and causative auxiliaries, as Krishnamurti (2003: 274) points out, 'have co-occurrence restrictions with the main verb and modify its lexical structure and meaning' (hence, these features are indicated by being given within brackets in (27)). In cases where all the features of a finite verb occur, the internal structure of a finite verb in Tamil is (38) (adapted from the structure of the Dravidian finite verb posited by Krishnamurti 2003: 29):

(38) Stem [(Root + (in-/transitive) + (non-/causative)] + Aspect + Tense + GNP [PNG] marker

As for NIA languages, Masica (1991: 258) posits the following formula (39) for the finite verb:

(39) VERB + Aspect Marker + $^{\text{(CONCORD)}}$ + Tense/Mood Marker + $^{\text{(CONCORD)}}$

He notes that the concord markers are bracketed 'because one or the other may be absent in one paradigmatic form or another in one NIA language or other.' Colloquial Sinhala, he observes, is the only major language in which both concord markers are absent. The other auxiliaries (the ones mentioned for Tamil above) involved in the construction of a complex VP occur in the same order as in a complex Tamil VP. However, SV agreement, as the above discussion shows, is a significant difference between the two languages.

It is necessary to provide a brief survey of the equally important non-finite verb forms which have already made their appearance above as components of verb phrases, e.g., participles. Non-finite verbs, which occur in subordinate clauses in the two languages, are marked for tense in some instances. In Tamil and Sinhala, they are not generally marked for agreement with the subject, though they take subjects. [46] Steever (1988) distinguishes two kinds of non-finite verbs in Dravidian on the basis of what kind of word, namely verb or noun they combine with: (i) those, e.g., infinitives and conjunctive participles, which combine with a verb to form compound verbs and adverbial clauses (see table 2.3); and (ii) those, e.g., adjectival or relative participles, which combine with a noun to form relative clauses. In addition, Steever (1988) distinguishes another kind of non-finite verbs, namely the nominalized verb forms in Dravidian. Sinhala too has all three non-finite forms (see Masica 1991 for a similar kind of classification of non-finite verbs in NIA languages). The non-finite verb forms and their negative forms in Tamil and Sinhala are given in table 3.6 (for verb forms of aspect (progressive and perfective) and modals in Tamil and Sinhala, see 4.3.3 and 4.3.4 respectively).

---

[46] Note, however, that the verbs are marked for agreement in utterance complement clauses (4.2.7) and headless relative clauses in Tamil (6.3).

Form		**Tamil**		**Sinhala**		
Base ‘do’		***sei***		***kərə***		
		Positive	**Negative**	**Positive**	**Negative**	
Infinitive		*seiya*	*seiyaamal-irruka*	*kərandə*	*nokərandə*	*kəran-næthuwə ində*
Conjunctive/ perfective participle		*seithu*	*seiyaamal*	*kərəla*	*nokərəla*	*kərəla næthuwə*
Adjectival/ relative form	Past	*seitha*	*seiyaatha*	*kələ/ kərəpu*	*nokələ/ nokərəpu*	*kelaə-næthi/ kərəpu næthi*
	Present	*seiyira*		*kərənə*	*nokərənə*	*kəran næthi*
	Future	*seiya-poora*	*seiya-poohaatha*	*kərandə-yanə*	*nokərandə-yanə*	*kərandə-yannæthi*
(Tensed) Nominalized/ focused form	Past	*seithathu*	*seiyaath-athu/ seiyaamal-poonathu*	*kərae/ kəlae*	*nokərae/ nokəlae*	*kərae/ kəlae næththe*
	Present	*seiyirathu*	*seiyaamal-poorathu*	*kəranne*	*nokəranne*	*kəranne næththe*
	Future	*seiya-poorathu*	*seiyaamal-poovathu/ seiya-poohathathu*	*kərannə yanne*	*nokərannə yanne*	*kərannə yan næththe*

Table 3.6: Non-finite verb forms

Note that the negative forms of imperatives (table 3.5) and infinitives use the verb ‘be’, *iru* in Tamil and *ində* in Sinhala, in addition to the main verb, e.g., *seiyaamal iru, kəran næthuwə ində* respectively, whereas the future adjectival participle and nominalized forms use the verb ‘go’, *poo* in Tamil and *yandə* in Sinhala in addition to the main verb, e.g., *seiyappoorathu, kərannə yanne* respectively. These again show replica grammaticalization because the structure of these forms is replicated from Tamil. As is obvious from the above survey of the non-finite forms, the two languages share the same non-finite system in that the three kinds of non finite forms and their negative forms in Tamil have almost similar corresponding forms in Sinhala.

Given in (40) are the full structures of a finite verb/verb phrase in Tamil and Sinhala:

(40) T. kamala than-da mahal-inda nihan-(g)al-ai
Kamala self-GEN daughter-GEN nail-PL-ACC
vettu-vi-thu-kondu iru-nth- aa
cut-CAUS-PTCP-get.PTCP be-PST-3SGF(HON)
S. kamala thaman-ge duwə-ge niyəpothu kappə-wə-genə
Kamala self-GEN daughter-GEN nails cut-CAUS-get.PTCP
hitiya
be.PST

‘Kamala was having her daughter’s nails cut.’

Note that the structure of the complex VP given in (40T,S) is the same except for the agreement features marked on the verb in (40T). This difference in agreement accounts for the difference in the values of three features in 3.2.4–3.2.6, since all these features are concerned with the expression of subjects/pronominal subjects and the marking of agreement features of A and S on the verb.[47]

Given Masica’s (1991) observation above that colloquial Sinhala is the only major NIA language in which both concord markers are absent, it might be hypothesized that Sinhala may not have had concord (agreement) markers at any time or it may have had these features but lost them subsequently. The latter option is in fact the correct one. Firstly, literary (written) Sinhala still has an agreement system (Gair 1998e), as in (41):

(41) T. avan antha puthakath-ai vaasi-th-aan
He that book-ACC read-PST-3SGM
S. ohu ee pothə kiyewweeyə
He that book read.PST.3SGM

‘He read that book.’

Note that in (41S) the agreement features are marked on the verb. Secondly, classical Sinhala had an agreement system which it lost after the 14th century (Paolillo 1994). Thirdly, besides the majority of finite verbs in spoken Sinhala which are unmarked for agreement features, there is a class of verbs which are marked for agreement (see 6.4), which may be taken to be a vestige of the agreement system that Sinhala once had. These facts taken together provide solid evidence for the conclusion that modern spoken Sinhala has lost its agreement system. This loss of agreement features is what Chandralal (2010: 66) refers to as a long process of simplification from the old Indian

---

[47] Note that the value of the feature in 3.2.7 is the same in the two languages, but this feature is also concerned with SV agreement.

period to the present day that the verbal system of Sinhala has undergone. The continuous transformations of the Sinhala verbal system, he notes, ‘have failed to make significant inroads into the literary or written language because they are not socially sanctioned among scholars in particular and within the writing tradition in general.’

What is significant to note with respect to the absence or lack of SV agreement in Sinhala is that Sinhala, as Paolillo (1994) argues, has lost the major part of its agreement system because of the contact-induced restructuring that its verbal system has undergone, triggered crucially by the cleft construction which Sinhala has replicated from Tamil (see 6.4 for details).

3.6.2 Features related to valence and voice: implications of the comparison

The typological profiles of Tamil and Sinhala established based on the comparative analysis of the six features related to valence and voice in 3.3 are given in table 3.7:

	WALS features	**Tamil(T)**	**Sinhala(S)**	**Similar/ Different**
3.3.1	Reciprocal constructions (106A)	distinct from reflexive	distinct from reflexive	similar
3.3.2	Passive constructions (107A)	present	present	similar
3.3.3	Antipassive constructions (108A)	no antipassive	no antipassive	similar
3.3.4	Applicative Constructions (109A)	benefactive object only and transitive base only	benefactive object only and transitive base only	similar
3.3.5	Periphrastic causative constructions (110A)	purposive but no sequential	purposive but no sequential	similar
3.3.6	Nonperiphrastic causative constructions (111A)	morphological but no compound	morphological but no compound	similar

Table 3.7: Typological profiles of features related to valence and voice

The values of all six features in this section are similar in the two languages. It is useful to analyze these features from the perspective of valency, especially as operations increasing/reducing valency, to find out whether the structure of any of these features may have been replicated in Sinhala modelled on the structure of the corresponding Tamil features. Of these features, reciprocal/reflexive pronouns (3.3.1), passives (3.3.2) and antipassives (3.3.3) are valency reducing operations, while applicative constructions

(3.3.4), periphrastic causative constructions (3.3.5) and nonperiphrastic causative constructions (3.3.6) are valency increasing operations.

The structure of the reciprocal/reflexive constructions ((9a) and (9b) in 3.3.1) in Tamil and Sinhala are the same. The reciprocal pronouns (9a) in the two languages are generated by reduplication (see 4.3.12). In Tamil, the indefinite (generic) pronoun ‘one’ is repeated, whereas in Sinhala the indefinite pronoun referring to ‘two’—without the indefinite suffix *-ek*—is repeated, expressing the meaning ‘each other’. In the two languages, it is the first pronoun that is marked for case. Note, however, that there is a difference between the case marked on the pronouns in (9a): in Tamil the first pronoun is marked for the accusative case (9aT), whereas in Sinhala the first pronoun is marked for the dative case (9bS). In Sinhala, when more than two (three or above) are involved, the reciprocal pronoun *ek-ekkena* ‘one to another/one to one’ derived from the indefinite (generic) pronoun *ekkena* ‘one’ is used. In this kind, the numeral *ekə* ‘one’ is prefixed to the generic pronoun *ekkena* ‘one’ to which the case is suffixed, e.g., in *ek-ekkena-tə* [one-one DAT] (cf. *denna-tə denna* ‘two-DAT two’ (9aS)). In Tamil, whether two or more are involved, the same reciprocal *oruvar-ai-oruvar* ‘one(HON)-ACC-one(HON)’ (9aT) is used. The case is suffixed to the generic pronoun at the end, unlike the other reciprocal in (9aS).

Despite these differences in the structure of the reciprocal pronoun, it seems likely that Sinhala has adopted the reciprocal from Tamil because the reduplicated kind of reciprocal is characteristic of the Dravidian languages (Yadurajan 1988) and the structure of the reciprocal pronouns in the two languages is the same. Of the NIA languages, Bengali, e.g., *paraspar* ‘each other’ and Hindi have distinct reciprocals meaning ‘each other’, while others like Marathi and Gujarati among others use reciprocals derived from numeral ‘one’, but do not involve reduplication. It is, therefore, plausible to assume that Sinhala has modelled its reciprocal on the reciprocal pronoun in Tamil. Adopting the distinction introduced in section (1.3), this modelling may be classified as replica grammaticalization in that Sinhala has replicated the structure of the Tamil reciprocal using Sinhala lexical equivalents. Only Sinhala among the NIA and Dravidian languages has reciprocal pronoun derived from numeral ‘two’, which may have been an endogenous change occurred in the Sinhala reciprocal pronoun.

Somewhat similar is the modelling of the reflexive pronoun in Sinhala on the model of Tamil, seen in (9b) in 3.3.1. The reflexive pronoun *thaman* in Sinhala is an instance of MAT(ter) replication of the Tamil reflexive pronoun *thaan* in that Sinhala uses the form-meaning unit of the Tamil reflexive pronoun. In the two languages the reflexive is exclusively third person. It is unspecified for number and gender in Sinhala (Henadeerage 1998), whereas in Tamil, it is unspecified only for gender. The plural form of *thaan* in Tamil is *thaagal*; the difference between the two reflexives is due to agreement features being marked on both subject proforms and verbs in Tamil.

A further point of interest in the reciprocal and reflexive constructions is that the two languages use the reflexive auxiliary or 'verbal reflexive' grammaticalized from the lexical verb *kol* in Tamil and *gandə* in Sinhala 'take/get/hold/contain' (see 4.3.3 for grammaticalized auxiliaries), in addition to the main verb; the two verbs form a compound verb, as shown in (9a,b).[48] The auxiliary verb facilitates the binding of the anaphor, that is, reciprocal and reflexive pronouns, in its local domain, conveying mutual action between two or more participants involved, or as (Schiffman 1999) notes, 'self affective or self benefactive action […] that affects the subject of the sentence in some way, usually to his/her benefit, but sometimes not in any clearly beneficial way.' Again, the major NIA languages, like Hindi and Bengali, do not use verbal reflexives. All major Dravidian languages, however, use the verb 'take/hold' as an auxiliary in reciprocal and reflexive constructions (Yadurajan 1988). Therefore, the almost identical elements and structures involved in the realization of these two features ((9a) and (9b)) suggest that these features of Sinhala have been replicated from the corresponding features of Tamil, resulting from their contact. Both kinds of replication, MAT and PAT, are attested in these features in Sinhala.

Though the structure of the passive constructions in the two languages is similar (see examples in (10) in 3.3.2), there are no noticeable features in Sinhala which need to be attributed to its contact with Tamil. The features involved in the passive constructions are cross-linguistically common. In Tamil, the PASS(ive) auxiliary is a grammaticalized form of the lexical verb, *padu*, 'experience/suffer/undergo' (10T), while in Sinhala the most common passive auxiliary is the grammaticalized form of the lexical verb, *labandə*, 'get/receive' (10S).

Tamil and Sinhala do not have the antipassive construction (3.3.3). However, they have constructions like (42a,b) which resemble respectively the antipassive with patient-like argument left implicit; and the antipassive with patient-like argument expressed as oblique complement (exs. (14, 15), p. 155, Chandralal 2010 translated into Tamil).

---

[48] Here and below, in instances of grammaticalization of lexical verbs (see 4.3.3) , the base (imperative) form of the lexical verb in Tamil and Sinhala is given; the lexical verb takes either the participle or finite form depending on its function in the VP.

(42) a. T. nulambu Ø kadi-kku-thu
Mosquito(es) bite-PRS-3SGN
S. maduruwo Ø kanəwa
Mosquito(es) eat.PRS

‘We are/I am bitten by mosquitoes.’

(42) b. T. kai-la nulambu kadith-ittu-thu
Hand-LOC mosquito(es) bite.PTCP-leave.PST-3SGN
S. at-ee maduruw-ek kaala
Hand-LOC mosquito-INDF eat.PFV

‘A mosquito has bitten me on the hand.’

(42a,b) are not passive constructions (cf. (10); also Chandralal 2007). Neither are they antipassive constructions because the verbs in the antipassive constructions are derived from the corresponding transitive verb ‘often with the help of overt morphology’, that is, an affix (Polinsky 2013a). Note that in (42a,b) the present active (default) form of the verb is used without any antipassive affixes. Nevertheless, in (42a) the patient-like argument is left implicit, whereas in (42b) it is expressed by an oblique (locative) case. What is important to note is that the two constructions in the two languages are almost identical.

In the applicative construction in the two languages (3.3.4), the beneficiary/benefactee (indirect object) is introduced to the construction by the APPL(icative) verb/auxiliary which occurs as the finite verb. The APPL verb in both languages is the grammaticalized form of the verb ‘give’ *kudu* in Tamil and *dendə* in Sinhala and the semantic role of the applied object is that of benefactee. Therefore, in (11) in 3.3.4 ‘Ranjith’ is the benefactor; ‘bottle’ is the basic object; ‘Chithra’ is the benefactee; and the past participle form of the verb ‘open’ denotes the action, while the APPL verb introduces the benefactee, increasing thereby, the valency of the verb. Several languages of South Asia, both Dravidian and NIA languages, have this kind of *give*-applicative constructions. Thus, replication of this feature by Sinhala from Tamil is a possible but not necessary assumption.

Finally, Sinhala shares with Tamil the same types of causative—periphrastic (3.3.5) and non-periphrastic or morphological (3.3.6)—constructions. In the periphrastic type, a grammaticalized form of the lexical verb *pannu* ‘make/do’ in Tamil and *gandə* ‘take/get’ in Sinhala are used as the causative auxiliary verb in the matrix clause representing ‘the causing event’ (cf. ‘the caused event’ (Lehmann 1989: 219)). This difference between the verbs used results in the difference in meaning between (12T) and (12S) in 3.3.5. There is also difference between the main verb in the embedded clause, which represents in the two languages: in Tamil, the infinitive form is used, whereas in Sinhala the participle form is used. The cases marked on the causees are also different: in Tamil, the causee is marked for the accusative case, whereas in Sinhala, it is marked for the instrumental case. Another difference between the periphrastic constructions in the two languages is that in Sinhala, as Chandralal (2010: 170, fn. 3) observes, the causative morpheme -*wə* which is also used in the non-periphrastic construction is suffixed to the participle form of the main verb after changing the final morpheme -*ə* into -*a* (see 3.3.5). In this construction in Sinhala, causation is, therefore, double marked by the causative auxiliary and the causative morepheme. This kind of coexistence of two categories in the same construction, as shown above, is a structural effect of contact-induced grammaticalization (Heine and Kuteva 2005).

In the morphological causative type, the causative morpheme -*vi* in Tamil and -*wə* in Sinhala, is added to the main verb to convey causation (13) in 3.3.6. The morphological causative is a well-attested areal feature of South Asia, e.g. Ramchand (2011) analyzes an almost similar kind of construction in Hindi/Urdu which she refers to as ‘indirect causative’. Therefore, no Sinhala-Tamil contact-induced change needs to be assumed. However, the near homophony of the suffixes -*vi* in Tamil and -*wə* in Sinhala strongly suggests that Sinhala has modelled its causative suffix on the Tamil causative suffix. This feature would exemplify MAT replication. Note also that Sinhala has extended the suffix -*wə* to its periphrastic causative construction (see above). The almost identical constituents and structures of the two causative constructions support the idea that there has been convergence between Sinhala and Tamil on these morphosyntactic constructions.

3.6.3 Features related to negation and question: implications of the comparison

The typological profiles of Tamil and Sinhala established on the basis of the comparative analysis of the seven features related to negation and questions in 3.4 are given in table 3.8:

	WALS features	**Tamil(T)**	**Sinhala(S)**	**Similar/ Different**
3.4.1	Negative morphemes (112A)	negative particle	negative particle	similar
3.4.2	Order of negative morpheme and verb (143A)	VNeg	VNeg	similar
	Preverbal negative morphemes (143E)	no preverbal negative morpheme	negative prefix	different
3.4.3	Position of negative morpheme with respect to subject, object, and verb (144A)	SOVNeg	SOVNeg	similar
3.4.4	Symmetric and asymmetric standard negation (113a)	Type Asy	Type SymAsy	different
3.4.5	Subtypes of asymmetric standard negation (114A)	A/Fin and A/Cat	A/Fin	different
3.4.6	Negative indefinite pronouns and predicate negation (115A)	Predicate negation also present	Predicate negation also present	similar
3.4.7	Polar questions (116A)	Interrogative verb morphology	Question particle	different

Table 3.8: Typological profiles of features related to negation and questions

Of the eight features compared, the values of four are similar, whereas the values of four are different. The first seven features are related to negation, while the last feature is related to polar questions. As for negation, the default Tamil/Sinhala negative form is of the single particle/word, single position type, with the clause-final negative particle *illai* in Tamil and *nææ* in Sinhala, following the verb (14 in 3.4.1). There is a difference of opinion among scholars with regard to whether the negative marker in Tamil is a suffix or a particle.[49] Lehmann (1989) calls it a negative marker. Amritavalli and Jayaseelan (2005) note that Malayalam, Kannada and Tamil have a common negative marker, namely *illai/illa*. It also functions in Dravidian as an independent utterance, ‘no’ often referred to as anaphoric negator. The fact that the negative element functioning as a single response to queries demonstrates that it is not a suffix, but a particle (unlike the polar question clitic in Tamil). Further evidence for the idea that the negative marker is a negative particle comes from the fact that it occurs alone as the

---

[49] The feature value of negative morphemes (3.4.1), for Tamil based on Asher (1982) given in Dryer and Haspelmath (2013) is ‘negative morpheme’ which is incorrect. The feature value for the two languages is ‘negative particle’. A possible reason for this assumption is that in Tamil, the particle *illai* is often articulated together with the verb, sounding as though it is a suffix.

negative form of the verb ‘be’ (Lehmann 1989: 230), as in (43b); ((43a) is the affirmative form; ex. (522a,b), p. 230, ibid. translated into Sinhala):[50]

(43) a. T. peey   iru-kk-(u)thu
    Ghost/s   be/exist-PRS-3SGN
 S. yakku   innəwa
    Ghosts   be/exist.PRS

    ‘Ghosts exist.’

(43) b. T. peey   illai
    Ghosts   be.NEG
 S. yakku   *nææ*
    Ghosts   NEG

    ‘Ghosts do not exist.’

The most important difference between the default negative construction in the two languages is that in Tamil, the negative particle occurs after the infinitive form of the verb (14T, 15bT), whereas in Sinhala, the negative particle occurs after the nominalized form of the verb (14S, 15bS). Note that the default Tamil negative form is neither marked for tense nor agreement; in other words, this negative form is tense, person, number and gender invariant, unlike its positive counterpart. The Sinhala negative form, on the other hand, is marked for tense, but unmarked for agreement. The nominalized form of the verb which occurs in Sinhala default negatives—just as in Sinhala default wh-questions (2.3.1.6)—may have been extended from the Sinhala cleft construction (see 6.1). This is also a kind of contact-induced change because the cleft construction in which this kind of verb occurs in the two languages was replicated from Tamil. In Sinhala, this form of the verb was overgeneralized into other contexts such as wh-questions and negatives, unlike in Tamil in which this form is not used in the default of these two constructions.

As in the constituent polar questions (2.3.1.6), the negative particle which occurs adjacent to a distinct constituent negates that constituent, lending focus reading, as in a constituent negative construction (44) (based on (14)):

---

[50] Lehmann (1989: 229) analyzes the Tamil negative particle *illai* as consisting of “*il* ‘be not’ (locative, existential and copula function)” and “a historical third person-plural-neuter suffix -*ai*” (ibid. 70). The structure of the Sinhala negative particle is not known.

(44) T. nandini **(illai)** poona kilamai **(illai)** pillaikal-ukku **(illai)** inglish **(illai)**
Nandini NEG last week NEG students-DAT NEG English NEG
padippi-th-athu
teach-PST-NMLZ
S. nandini **(nevei)** giyə sumaane **(nevei)** lamay-tə **(nevei)** ingriisi **(nevei)**
Nandini NEG last week NEG students-DAT NEG English NEG
igænnuw-e
teach.PST-NMLZ

‘It was not Nandini that taught English to students.’
‘It was not last week that Nandini taught English to students.’
‘It was not to students that Nandini taught English last week.’
‘It was not English that Nandini taught to students last week.’

In instances like (44T), in addition to the default negative particle *illai*, a negative particle *alla* can also be used.[51] Gair (1980) notes that in instances like (44S), the negative form *nevei* (or its dialectal variants, *nemee*, *novey*) serves as a ‘narrow focus negator’. The structure of (44) is similar to the constituent polar constructions ((31b) in 2.3.1.6) and, the constituent negative constructions in the two languages have also been modelled on the focus particle constructions (6.1).

The values of features—negative morphemes (3.4.1), order of negative morpheme and verb (143A in 3.4.2), position of negative morpheme with respect to subject, object, and verb (3.4.3)—for the two languages are the same. The values of the preverbal negative morphemes (143E in 3.4.2) are different because Sinhala also has a negative construction in which the negative prefix is attached to the verb ((15aS) in 3.4.2), while Tamil does not have this construction. Therefore, of the four features compared, three are similar, while one is different. The default clause-final negative particle in the two languages is of the TypeAsy kind because the presence of the negative particle in the negative clause is not the only difference between the affirmative and negative clauses in the two languages; the affirmative differs from the negative in other ways too (see 3.4.4). However, the preverbal negative construction in Sinhala is of the TypeSym because the presence of the negative particle is the only difference between the negative clause and its affirmative counterpart. The values of symmetric and asymmetric standard negation (3.4.4) are, therefore, different: Tamil has TypeAsy, whereas Sinhala has TypeSymAsy. The values of subtypes of asymmetric standard negation (3.4.5) are also different because Tamil has subtypes A/Fin ((18T) in

---

[51] The former and the latter can be used interchangeably. Annamalai (2004) observes that Tamil is fast losing *alla* as the negator of focused constituents in instances like these, and uses *illai* instead. In the literary/written form of Tamil, *alla* is used.

3.4.4) and A/Cat ((19T) in 3.4.5), whereas Sinhala has only subtype A/Fin ((18S) in 3.4.4).

The values of negative indefinite pronouns and predicate negation in the two languages are similar in that both show mixed behaviour with regard to indefinite pronouns and predicate negation: indefinite pronouns can co-occur with or preclude predicate negation, as in (20a) and (20b) respectively. The feature values of polar questions for the two languages are different because, as already mentioned, Tamil uses a question clitic, while Sinhala uses a question particle. The position of the question clitic/particle in polar and constituent polar questions (2.3.1.6), however, is the same. The two features, the position of polar question particles (2.2.12) and polar questions (3.4.7), actually involve the same feature and the different values for them in the two languages need to be treated as a single difference between them.

Although Tamil and Sinhala share four of the seven features related to negation, it is difficult to determine whether these features of Sinhala have undergone any changes owing to its contact with Tamil (but see Gair's (1998d) view in 2.3.1.1). However, as discussed above, the use of the nominalized form of the verb in default negatives in Sinhala (14S in 3.4.1) is a contact-induced change. Another contact-induced change is that the constituent negative construction in the two languages (44) has been modelled on the focus particle construction (see 2.3.1.6; and also 6.5).

The values of negative indefinite pronouns and predicate negation from the two languages are the same (see (20a,b)). What also seems reasonable to claim is that the two Sinhala constructions (20aS and 20bS) in (3.4.6) have been modelled on the corresponding Tamil constructions (20aT and 20bT) for the following reasons: (i) the structures of the two constructions are almost identical; and (ii) the constituents involved in the two constructions are the same; and (iii) in the two languages these indefinite pronouns are either negative (20a) or positive (20b) (see also universal quantifiers in 4.3.7). With respect to (ii), it appears that Sinhala has replicated the pattern of the indefinite pronouns from that of Tamil indefinite pronouns in that both are interrogative-based, consisting of wh-words/phrases and the inclusive clitics *-um* in Tamil and *-th* in Sinhala.[52] The interrogative-based indefinite pronouns of Tamil and Sinhala are given in table 3.9: [53]

---

[52] For the five kinds of indefinite pronouns, see 'Indefinite Pronouns', WALS 46, Haspelmath 2013c.

[53] In Tamil and Sinhala, the clitic is added to the indefinite wh-phrase after the case (if it is marked for case), as in *yaar-ukku-m, kaa-tə-*(*wa*)*th* who-DAT-INCL 'to no one' respectively.

Negative indefinites	**Interrogative base**	**Tamil *-um***	**Sinhala *-th***
No one	Who	*yaar-um/evar-um*	*kauru-th*
To no one	who(m)	*yaarukk-um/ evarukk-um*	*kaatəwa-th*
Nothing	What	*ethuv-um*	*monəwa-th*
No where	Where	*engae-um*	*kohewa-th*
Never	When	*eppav-um*	*kawədaawa-th*

Table 3.9: Interrogative based (negative) indefinites

The composition of these interrogative based indefinitives is strikingly similar. In addition to this kind of indefinitives, the two languages also use another kind in which the inclusive clitics, *-um* in Tamil and *-th* in Sinhala, are added to the indefinite (generic) pronouns, which is derived from numeral ‘one’ in the two languages, denoting a ‘person’ *oruvar/(e)kenek* or ‘animal/thing’ *ondu/ekak* respectively (see structure in (45T/S)). The interrogative-based indefinite negative pronouns can be replaced by the indefinite (generic) pronoun, as in (45), based on (20):

(45) T. oru-(v)ar-um vahuppu-kku var-a illai
One(INDF)-3SG(HON)-INCL class-DAT come-INF NEG
S. (ek)-ken-ek-(wa)th panthiyə-tə aaw-e nææ
One-3SG(HUM)-INDF-INCL class-DAT come.PST-NMLZ NEG

‘No one came to class.’

It is useful to note that this kind of indefinite in the two languages can co-occur only with negative predication. The structure of the two constructions (20a,b) and (45) and the two kinds of indefinite pronouns (interrogative based and generic-pronoun based) show a remarkable degree of similarity in the two languages, supporting the idea that Sinhala has modelled these structures and constituents on the corresponding Tamil structures and constituents.

3.6.4 Predication features: implications of the comparison

The typological profiles of Tamil and Sinhala established based on the comparative analysis of the five features related to predication in 3.5 are given in table 3.10:

	WALS features	**Tamil(T)**	**Sinhala (S)**	**Similar/ Different**
3.5.1	Predicative possession (117A)	locational/ dative	locational/ dative	similar
3.5.2	Predicative adjectives (118A)	nonverbal encoding	nonverbal encoding	similar
3.5.3	Nominal and locational predication (119A)	different	different	similar
3.5.4	Zero copula for predicate nominals (120A)	possible	possible	similar
3.5.5	Comparative constructions (121A)	locational	locational	similar

Table 3.10: Typological profiles of features related to predication

The values of all five features compared in this section are similar.[54] Again, it is likely that Tamil and Sinhala have converged on these features because the structures which encode them are identical in the two languages. A cursory glance at the data compared above shows that in the two languages all five features involve a COP(ula). However, constructions encoding predicative adjectives ((23a) in 3.5.2) and nominal predication ((24a), (25) in 3.5.3, 3.5.4 respectively); and comparative constructions ((26) in 3.5.5) by default have a null COP, whereas constructions encoding predicative possession ((22) in 3.5.1) and locational predication ((24b) in 3.5.3) have an overt COP. To express a past state, however, all these constructions need to have the past form of the overt COP irrespective of whether they have overt or covert COP in the present, as in the past form of (24) given in (46):

(46) T. ratha   oru   teecher-aaka   iru-nth-aa
Ratha   one(INDF)   teacher-ADV   be-PST-3SGF(HON)
S. ratha   guruvəriy-ak-wə   hitiya
Ratha   teacher-INDF-ADV   be.PST

‘Ratha was a teacher.’

---

[54] The feature value of nominal and locational predication (3.5.3) for Tamil based on Asher (1982) given in Dryer and Haspelmath (2013) is ‘identical’, which is incorrect. The reason for this conclusion is not known. The value should be ‘different’, that is, the two constructions use different strategies: null copula is used in nominal predication, while overt copula is used in locational predication. The value of this feature for Sinhala based on Gair (1970) given in Dryer and Haspelmath (2013) is ‘different’ which is correct.

There are two copulas in Tamil (Lehmann 1989) and Sinhala (Gair and Paolillo 1988). One of them which expresses a state of being is used in (46): *iru* ‘be/sit/exist/’ in Tamil and *ində* (for animate) and *thiyandə* (for inanimate) ‘be/exist/stay’ in Sinhala.[55] In (46S) *hitiya* is the past form of *ində*.

The other copula is *aaku* in Tamil and *wendə* in Sinhala used in the written/literary form of the construction, encoding nominal predication in the two languages, as in (47) in which it is overt (adapted from ex. 7, p. 242, Gair 1998e):

(47) T. avar oru kamakaarar aav-aar
He-HON one(INDF) farmer be.PRS-3SGM(HON)
S. hetema goviy-ek weyi
He-NOM farmer-INDF be.PRS.3SG

‘He is a farmer.’

Note that in Tamil, *aavaar* ‘be /become/to come into existence’ (Lehmann 1989: 160) and in Sinhala, *weyi* ‘be/become’ (for structure, see 2.3.1.1) are used as the copula in (47). In Sinhala, *-yi* is an ASS(ertive) marker, and predicative adjectives are obligatorily marked as assertive in instances like (23a) in 3.5.2.

The two languages also share the same comparative constructions. In these constructions the standard NP in Tamil is marked for accusative case, while in Sinhala, it is marked for dative case. The near homophonous comparative particles *-vida* ‘than’ in Tamil (26T) and *vadaa or vædiyə* ‘more than’ can be considered another instance of MAT replication. Again, the two almost similar constructions indicate that the Sinhala comparative construction is a calque of the Tamil comparative construction. Except for nominal and locational predication, the Dravidian and NIA languages share the same values for most of these features (see Appendix). However, the different morphosyntactic features related to these five features and the almost isomorphic constructions will not in any way preclude the possibility of Sinhala restructuring these constructions on the model of the corresponding Tamil ones.

3.7 Conclusion

Of the 27 features related to simple clause compared in this chapter—eight related to case and agreement; six related to valence and voice; eight related to negation and questions; and five related to predication—the values of 20 are similar, while those of seven are different. They, therefore, show significant convergence on these features

---

[55] See fn. 48.

though less than their convergence on word order related features. Although most of the NIA and Dravidian languages share the same feature values, it can safely be concluded that the similarity between the structures and constituents of these features in the two languages has resulted from the morphosyntactic changes that Sinhala has undergone induced by its contact with Tamil.

Of the eight features related to case and agreement (3.2), the first three features related to case alignment and the last one about ditransitive constructions are similar. Sinhala may have changed its case alignment from the tripartite/split ergative-absolutive system characteristic of the NIA languages to the nominative-accusative (standard) system which it has owing to its contact with Tamil. The four features related to agreement (including WALS 104A in 3.2.7) are different between the two languages because the default finite verbs in spoken Sinhala lack agreement, whereas those in Tamil have agreement.

Tamil and Sinhala share similar values for all the features related to valency and voice (3.4). The structures and constituents involved in these features are remarkably similar in the two languages, especially given their different origins. The similarity between all structures—e.g., the reciprocal/reflexive constructions with their verbal reflexives, periphrastic and morphological constructions—except for the passive, show that Sinhala may have replicated the pattern of the corresponding Tamil constructions. The replication of form-meaning units (matter) or replica grammaticalization e.g., the reduplicated reciprocals, the near homophonous reflexives, the verbal reflexive grammaticalized from the same lexical verb in reciprocal/reflexive constructions and the causative morpheme in morphological constructions provide suggestive evidence that Sinhala has replicated the majority of structures and elements discussed here from Tamil.

The values of four of the eight features related to negation and questions (3.4) are similar between the two languages, whereas those of four are different. The structure of the default negatives, which are of the asymmetric type, is the same in the two languages in that the negative marker occurs clause-finally following the verb. The preverbal negative construction in Sinhala which is of the symmetric type, as discussed in 2.3.1.1 and 3.4.2 may have been an NIA feature in Sinhala. The other negatives in Tamil which are of the asymmetric type do not have corresponding negatives in Sinhala. As a consequence, there are differences between the negatives in the two languages. One difference between the default negative constructions in the two languages, namely the occurrence of the nominalized form of the verb as the default form in the Sinhala negative, is due to the fact that Sinhala has extended the nominalized verb form used in the cleft construction to this kind of negative. Owing to the use of the nominalized verb form in the Sinhala negative, the structure of the default negatives in the two languages is different. Note, however, that the difference between the default negatives of the two languages also results from a contact-induced change. It is significant to note that the differences between features of languages in contact in situations like that of Sinhala and Tamil provide interesting insights into the nature of changes in languages (see below for other instances). The reason for the disparity between the polar questions, as has already been discussed in 2.3.1.5 and 2.3.1.6, is that Sinhala uses a question particle, whereas Tamil uses a question clitic. However, it can safely be concluded that Sinhala has modelled its polar questions on Tamil polar questions for the following two reasons: (i) the structure of the polar and constituent polar questions are the same; and (ii) as shown in 2.3.1.5 the Sinhala question particle was adapted from the Tamil/Dravidian correlative particle.

Finally, for the features related to predication (3.5), all the values are the same for Tamil and Sinhala. Though the other languages of the region too share the same feature values, it is likely that Sinhala has modelled these features on the corresponding features of Tamil. Note the clear instance of matter replication, e.g., the near homophonous comparative markers and the almost isomorphic constructions—in some of which the copula is null, while in others it is overt—which the two languages share. Altogether, the features discussed in this section seem to be compatible with the idea that contact between Sinhala and Tamil has induced changes in almost all the features in Sinhala: some led to similarities between the features of the two languages, while others led to differences, but these changes have made these features of Sinhala evolve the way they are today.

Chapter 4

Features related to complex sentences and miscellaneous features

4.1 Preliminaries

The typological profiles of Tamil and Sinhala established thus far are based on various features related to simple clauses (Chapter 3), including almost all the word order features compared in Chapter 2. The present chapter compares 21 features: seven are related to complex sentences and 14 others which in many cases have seen changes in Sinhala induced by its contact with Tamil.

The phenomena considered in this chapter involve the combination of clauses into complex sentences, consisting either of two independent clauses with similar syntactic status, or of one independent (matrix) clause and a dependent (subordinate) clause, where they have a different syntactic status. The mechanism used to produce the former is called coordination and that which is used to produce the latter is called subordination. Thus, complex sentences are sentences with a matrix clause and one or more subordinate clauses combined by means of subordination or coordination. For the latter type, the more specific term compound sentence is often used.

Section 4.2 compares seven typological features of Tamil and Sinhala related to complex sentences. These seven features (WALS 122A–128A) belong to the section titled ‘Complex Sentences’ in Dryer and Haspelmath (2013; see section ‘H’ in Haspelmath et al. 2005). These seven features are all to do with different kinds of subordination (see also 2.2.14, 2.3.1.7). WALS 122A and 123A are concerned with relative clauses. The other features are about ‘Want’ Complement Subjects (WALS 124A); Purpose Clauses (WALS 125A); ‘When’ Clauses (WALS 126A); Reason Clauses (WALS 127A); and Utterance Complement Clauses (WALS 128A).

Compared in Section 4.3 are 13 diverse morphosyntactic and discourse-pragmatic features. These features are third person pronouns; topicalization; two kinds of auxiliaries, general auxiliaries mostly grammaticalized from lexical verbs in the two languages, and those conveying modality; coordination; alternative questions; quantifiers; verbal nominals; volitive/involitive verbs in Sinhala and effective/affective verbs in Tamil; final complementizer and final polar particle in embedded clauses; dative subject constructions; reduplication; serial verb constructions; and converbs. The values of the 13 features are determined by direct comparison of these features in the two languages.

In Section 4.4, the main findings arising from the compartive analysis of the features included in Sections 4.2 and 4.3 are used to establish the typological profiles of these features, and the implications arising from the findings are then discussed with the aim of studying the nature of morphosyntactic changes that Sinhala may possibly have undergone owing to its contact with Tamil.

4.2 Features related to complex sentences: a comparison

4.2.1 Relativization on subjects

This feature (WALS 122A, Comrie and Kuteva 2013b see also 2013a) is about the different strategies that the languages of the world use to relativize on the subject in a construction formed by the relative clause and its head noun (for the different kinds of order of relative clause and head noun, see 2.2.10). Comrie and Kuteva (2013a) distinguish between relativization on subjects (WALS 122A) and relativization on obliques (WALS 123A is compared in 4.2.2 below).

There are four values for this feature: 1. relative pronoun strategy; 2. non-reduction strategy; 3. pronoun-retention strategy; and 4. gap strategy (Comrie and Kuteva 2013b). Tamil and Sinhala use the gap strategy (value 4) to relativize on the subject, as in (1b) (relative clauses here and 4.2.2 below are enclosed in brackets); ((1a) is the sentence from which (1b) is formed):

(1) a. T. manushan naai-kku saapaadu kudu-th-aan
Man dog-DAT food give-PST-3SGM
S. miniha ballaa-tə kææmə dunna
Man dog-DAT food give.PST

‘The man gave the dog food/the man fed the dog.’

(1) b. T. [_______ naai-kku saapaadu kuduth-a] manushan
dog-DAT food give.PTCP-RELAT man
S. [_______ ballaa-tə kææmə diip-u] miniha
dog-DAT food give.PTCP-RELAT man

‘The man who gave food to the dog/the man who fed the dog’

Note that the gap within the relative clause known as the relativized position is said to be co-referential with the head noun (Tallerman 2015: 276). The gap is understood to refer to the noun, ‘the man’, which is moved from the subject position of the relative clause to be the head noun, creating thereby the gap; hence, the strategy is called gap strategy. The relativizers *-a* in Tamil and *-u* and *-ə* in Sinhala are suffixed to the participle form of the verb, which thereby becomes a relative participle (non-finite) form of the verb. Note that typologically, prenominal relatives tend to be non finite (Jayaseelan 2014; also Kayne 1994). Dravidian and Sinhala relatives are non finite.

4.2.2 Relativization on obliques

This feature (WALS 123A, Comrie and Kuteva 2013c) is about the different strategies the languages of the world use to relativize on obliques. Of the five values for this feature, values 1–4 are the same strategies as used in the relativization on Subjects (4.2.1), and value 5 is 'not possible', that is, it is not possible to relativize on obliques. Tamil and Sinhala employ the gap strategy in relativization on obliques too (value 4), as in (2b); ((2a) is the sentence from which (2b) is formed):

(2) a. T. [manushan naaiy-ai thadi-yaal adi-th-aan
Man dog-ACC stick-INS beat-PST-3SGM
S. [miniha ballaa-tə kootu-(w)ən gæhuwa
Man dog-DAT stick-INS beat.PST

'The man beat the dog with the stick.'

(2) b. T. [manushan naaiy-ai _______ adith-a] thadi
Man dog-ACC beat.PTCP-RELAT stick
S. [miniha ballaa-tə _______ gahap-u] kootuwə
Man dog-DAT beat.PTCP- RELAT stick

'The stick which the man beat the dog with'

4.2.3 'Want' complement subjects

This feature (WALS 124A, Haspelmath 2013d) is concerned with the syntax of 'want'. In sentences which encode the notion of 'want', the notional subject of its complement predication, that is, the dependent (embedded) clause, can, when it is coreferential with the wanter, be left implicit or be expressed overtly in the complement clause by means of a pronominal element. There are five values for this feature: 1. the complement subject is left implicit; 2. the complement subject is expressed overtly; 3. both construction types exist; 4. 'want' is expressed as a desiderative verbal affix; and 5. 'want' is expressed as an uninflected desiderative particle. In Tamil and Sinhala, the complement subject is left implicit (value 1), as in (3) below ('want' complement clauses are enclosed in square brackets):

(3) T. ena-kku [PRO oru kaahitham eluth-a] venum
I-DAT one(INDF) letter write-INF want.PRS.3SGN
S. ma-tə [PRO lium-ak liyandə] oone
I-DAT letter-INDF write-INF want

‘I want to write a letter.’

The two kinds of notional predications in these constructions are the predication of wanting and the predication that expresses the desideratum, that is, what needs to be done. The latter is expressed in these examples by a clause that lacks an overt subject. The constructional meaning makes it clear that the notional complement subject is to be understood as being the same as the experiencer argument of ‘want’ in the main clause. As illustrated in (3), it is represented as (big) PRO in Chomskyan syntactic theory.

4.2.4 Purpose clauses

This feature (WALS 125A, Cristofaro 2013a) is concerned with the form of the verb in purpose clauses which encode a particular relation between the events that the clauses describe. In this biclausal construction, the event coded by the main clause (the main event) is performed with the goal of obtaining the realization of the one coded by the purpose clause (the dependent event). Following the distinction introduced by Stassen (1985), verb forms in purpose clauses are classified as either balanced or deranked, and the clauses containing them are, therefore, called balanced and deranked clauses (Cristofaro 2013a). The balanced forms are those which can occur in an independent declarative clause (finite verbs), while deranked forms are those which cannot occur in an independent declarative clause (infinitival or participle forms of the verb). On the basis of the two classes of verb forms, three values are distinguished for this feature: 1. by balanced verb forms only; 2. by either deranked or balanced verb forms; and 3. by deranked verb forms only. In Tamil and Sinhala, purpose clauses can only be coded using deranked verb forms, that is, the infinitival verb form (value 3), as in (4) below (purpose clauses are enclosed in square brackets):

(4) T. avan$_i$ kovil-(u)kku [PRO$_i$ kumbid-a] poo-n-aan
He temple-DAT worship-INF go-PST-3SGM
S. eyaa$_i$ koovilə-tə [PRO$_i$ vandində] giya
He temple-DAT worship.INF go.PST

‘He went to the temple to worship.’

The above examples illustrate the typical kind of purpose relation referred to as motion predicates in which the main and dependent events share the same participant. However, there may be non-motion predicates in which the main and dependent events need not share the same participant, as in (5) below. In such cases too, Sinhala and Tamil both use the infinitive, a deranked verb form.

(5) T. naan$_i$ en-da puthakathth-ai avan-(u)kku$_j$ [PRO$_j$ vaasikk-a] kudu-th-een
I I-GEN book-ACC he-DAT read-INF give-PST-1SG
S. mamə$_i$ mage pothə eyaa-tə$_j$ [PRO$_j$ kiyəwandə] dunna
I I-GEN book he-DAT read.INF give.PST

'I gave him my book in order for him to read.'

4.2.5 'When' clauses

This feature (WALS 126A, Cristofaro 2013b**)** deals with the form of the verb in 'when' clauses which encode a temporal relation between two events, such that there is a temporal overlap between the two. The same deranked/balanced distinction used in WALS 125A is used to classify the verb forms in 'when' clauses. There are three values for this feature: 1. by balanced verb forms only; 2. by either deranked or balanced verb forms; and 3. by deranked verb forms only. In Tamil and Sinhala, 'when' clauses can only be coded using deranked, that is, participle verb forms (value 3), as in (6) below ('when' clauses are enclosed in square brackets):

(6) T. [teecher vakuppu-kku poona-pothu] pillai-kal
Teacher classroom-DAT go.PTCP-when student-PL
oru-(v)ar-um irukk-a-villai
One(INDF)-3SG(HON)-INCL be-INF-NEG
S. [tiichər panthiyə-tə yanə-kotə] lamay
Teacher classroom-DAT go.PTCP-when student.PL
(ek)-ken-ek-(wa)th hitiy-e nææ
One-3SG(HUM)-INDF-INCL be.PST-NMLZ NEG

'When the teacher went to the class, there were no students.'

These examples encode linked events which are simultaneous. 'When' clauses can encode events that are not simultaneous, that is, one event occurs before/after the other. There might be an interval of some days, or even months or years between them, as in (7) below:

(7) T. [avar-kal thiyater-(u)kku poona-poothu] padam
They theater-DAT go.PTCP-when film
thodangi-vittu-thu
start.PTCP-leave.PST-3SGN
S. [eyaala sinəmaavə-tə yanə-kotə] chithrəpatəyə pataŋ araŋ
They theater-DAT go.PTCP-when film start.PTCP get.PTCP
thibuna
be.PST

‘When they went to the theatre, the film had started.’

In (7), the event coded in the main clause occurred before that which is coded in the subordinate clause. However, the same participle form of the verb is used.

4.2.6 Reason clauses

This feature (WALS 127A, Cristofaro 2013c) is about the form of the verb in reason clauses which encode a causal relation between two events, such that one of the two (the event coded by the reason clause, the dependent event) represents the reason for the other event (the main event) to take place. The same deranked/balanced distinction of WALS 125A and 126A is used to classify the verb forms in reason clauses too. The three values for this feature are: 1. by balanced verb forms only; 2. by either deranked or balanced verb forms; and 3. by deranked verb forms only. In Tamil and Sinhala, reason clauses can only be encoded using deranked, that is, participle verb forms (value 3), as in (8) below (reason clauses are enclosed in square brackets):

(8) T. [avan netu uur-(u)kku poonath-aala] Ø pallikoodath-ukku
He yesterday village-DAT go.PTCP-because school-DAT
var-a-(v)illai
come-INF-NEG
S. [eyaa iiye gamə-tə giyə nisa] Ø iskoole-tə
He yesterday village-DAT go.PTCP because school-DAT
aaw-e nææ
come.PST-NMLZ NEG

‘Since he went to the village yesterday, he didn’t come to school.’

Note that in (8) Ø stands for the null subject in the main (independent) clause, known as (little) pro in Chomskyan syntactic theory.

4.2.7 Utterance complement clauses

This feature (WALS 128A, Cristofaro 2013d) is about the form of the verb in utterance complements. These are complement constructions defined ‘as [those] expressing a particular relation between events, such that one event (the one coded by the main clause, introduced by utterance predicates such as ‘say’ or ‘tell’) entails that another event (the one coded by the complement clause) is referred to’ (Cristofaro 2013d). The three values of this feature are those used in WALS 125A–127A above: 1. by balanced verb forms only; 2. by either deranked or balanced verb forms; and 3. deranked verb forms only. The verb forms in utterance complements in Tamil and Sinhala are balanced (value 1), as in (9) below (utterance complement clauses are enclosed in square brackets):

(9) T. rohan [kumar avan-da veet-ai vithth-itt-aan endu]
Rohan Kumar he-GEN house-ACC sell.PTCP-leave.PST-3SGM COMP
son-n-aan
say-PST-3SGM
S. rohan [kumar eyaa-ge gee vikkaa kiyəla] kiwwa
Rohan Kumar he-GEN house sell.PST COMP say.PST

‘Rohan said that Kumar had sold his house.’

In (9), Rohan’s statement refers to the occurrence of the event that Kumar had sold his house.

4.3 Miscellaneous features: a comparison

4.3.1 Demonstrative-based third person pronouns

As in many other languages, the third person pronouns are related to the demonstratives in that they are formed from the demonstrative stem in the two languages. In Tamil, the stems of the two demonstratives—*a*-(expressing remoteness) and *i*-(expressing proximity)—combines with a bound pronominal base with the feature content (person, number and gender) to form the third person pronouns. As for gender, Tamil singular pronouns distinguish among masculine, feminine, honorific and neuter (non-human, both animate and inanimate), while plural pronouns distinguish between human and non-human (both animate and inanimate) only. For example, *a-/i-* + (*v*)*an* = *avan* ‘that one’/*ivan* ‘this one’ = ‘he’ (Lehmann 1989: 94–97).

The same is true for Sinhala in that the demonstrative stem (see below) and a bound pronominal base with the feature content (person, number, gender) combine to form the third person pronouns.[56] However, unlike the two demonstratives in Tamil, there are four demonstratives in Sinhala ‘showing the proximity-anaphor distinction’: i. Speaker proximal: *mee* ‘this/these’; ii. Addressee proximal *oyə* ‘that/those’; iii. Distal from both speaker and addressee *arə* ‘that/those’ (over there); and iv. Distal from both speaker and addressee and anaphoric *ee* ‘that/those’ (Chandralal 2010: 47; also Gair and Karunatillake 2000). For example, *mee-/ee-/arə- + yaa = meyaa* ‘this one’*/eyaa/arəyaa* ‘that one’ = ‘s/he’. [57] As for gender, the default Sinhala singular and plural pronouns distinguish between human and nonhuman (animate and inanimate) only, while the informal forms distinguish among masculine, feminine and neuter (non-human, both animate and inanimate). (10) shows the third person pronoun in a finite clause:

(10) T. avan kai-ai uyar-thin-aan
He hand-ACC raise-PST-3SGM
S. eyaa athə issuwa
He hand raise.PST

‘He raised his hand.’

The third person plural pronouns in the two languages are also demonstrative-based and the plural suffix is added to the singular pronoun (see table 3.4 for third person plural pronouns in the two languages).

4.3.2 Topicalization

Topicalization is the syntactic process whereby one constituent is made the topic of a sentence or a clause. In Tamil and Sinhala, as in English, the constituent that is topicalized is moved to the topic position (left peripheral) at the beginning of the sentence. It is marked with the topic marker *-endaal* in Tamil and *-naŋ* in Sinhala: both expressing the meaning ‘as for X’ or ‘talking about X’, as in (11b) (the default form is (11a) (the topic particles are in boldface):

---

[56] In Tamil, the feature content of the pronoun, as has been mentioned before, triggers corresponding features to be marked on the finite form of the verb, whereas in Sinhala the feature content of the pronoun does not trigger any changes in the verb. The same finite form of the verb marked only for tense is used with a subject of any feature content in Sinhala (see tables 3.3 and 3.4).

[57] Note that in Sinhala, the stem of the second person singular pronoun *oyaa* ‘you’ is *o-* which may probably be the reason why there is no third person pronoun with the demonstrative stem *o-* because the third person pronouns with stem *o-* would have created ambiguity between the second and third person pronouns which have the same base *-yaa*. There are, however, informal third person pronouns with the stem *o-* (see tables 3.4) because these pronouns have bases which are different from the bases of second person pronouns.

(11) a. T. mary-kku maampalam nalla viruppam
Mary-DAT mango very like
S. mary ambəvələ-tə huŋgak kæməthi
Mary mango-DAT very like

‘Mary likes mangoes very much.’

(11) b. T. mampalam-**endaal** mary-kku nalla viruppam
Mangoes-TOP mary-DAT very like
S. ambəvələ-tə **naŋ** mary huŋgak kæməthi
Mango-DAT TOP Mary very like

‘Mangoes, Mary likes very much.’

Also used as topic markers in the two languages are *vanthu* in Tamil and *æwilla* in Sinhala, both of which are grammaticalized from the lexical verb ‘come’, as in (12):

(12) T. naangal **vanthu** kalavu edukkir-a manisar illai
We TOP steal take.PTCP-RELAT people NEG
S. api **æwilla** horəkam kərən-ə minissu nevei
We TOP steal do.PTCP-RELAT people NEG

‘We are not the kind of people who steal.’

Since the two languages have scrambling, constituents can also be scrambled to the topic position without a topic particle ((18) in 2.3). Sarma (2003) calls this kind of movement in Tamil L(eft)-ward extraction.[58]

Further, it is also possible to topicalize a complete clause in the two languages, as in (13):

(13) T. ithai thiruthira-**enda** kashtam
This repair.PTCP-TOP difficult
S. meekə hadandə **naŋ** amaaru-i
This repair.INF TOP difficult-ASS

‘It is difficult repairing it (something).’

---

[58] In a direction diametrically opposite to topicalization and L-extraction are two other movements in the two languages, namely clefting—one of the two ways in which exhaustive focus is assigned to a constituent in a sentence—and R(ightward)-extraction, a movement like L-extraction which moves a constituent to post-verbal position (Sarma 2003, 1999; for more details, see Chapter 6).

4.3.3 Auxiliaries grammaticalized from lexical verbs

Tamil and Sinhala clauses have compound verbs constructed of the main lexical verb that denotes the action in the clause and one or more auxiliaries which are either participles or finite verbs. Herring (1989) points out that compound verb constructions or verb + verb sequences (see below) are a well-known phenomenon in South-Asian languages, where one of the members has attained a grammaticalized or semi-grammaticalized status. The auxiliaries found in these compound VPs below in the two languages are lexical verbs which have undergone grammaticalization. Moreover, they are used in the same syntactic and semantic contexts. In effect, these verbs have undergone semantic extension from their original context to new contexts. Two examples of these lexical-verbs-turned-auxiliaries that have surfaced in the discussion so far are the auxiliary used in reciprocal/reflexive constructions, *kol* in Tamil and *gandə* in Sinhala ‘get/take/hold’ (3.3.1) and the auxiliary used in the applicative construction, *kudu* in Tamil and *dendə* in Sinhala ‘give’ (see 3.3.4; for the structure of the finite verb in the two languages, see 3.6.2.2).

These grammaticalized verbs are used either as the final finite verb in a verb group (compound verb)—marked for tense and agreement in Tamil and tense in Sinhala—or as a non-finite verb, especially as a participle depending on their function in the VP. There are some other grammaticalized auxiliaries which are different in the two languages, i.e., they have not been derived from the same lexical verbs; these forms are not discussed in this dissertation (for Tamil auxiliaries, see Lehmann 1989; for Sinhala auxiliaries, see Chandralal 2010). The examples below show the most important grammaticalized auxiliaries used in compound verbs and the contexts in which they are used in the two languages.[59]

(i) *kol* in Tamil and *gandə* in Sinhala ‘get/take/hold’

In addition to its use as a verbal reflexive, as in (9a) and (9b) in 3.3.1, repeated here as (14a), this auxiliary is also used ‘as the first component [a participle] in a compound continuative aspectual marker’ (Herring 1994: 178) in the two languages, as in (14b) (auxiliaries here and below are in boldface):

---

[59] In the titles of (i)–(viii) below, the base (imperative) form of the verbs are given for Tamil and Sinhala (see fn. 48).

(14) a. T. ravi thann-ai(-ye) kuththi-**kon-**d-aan
Ravi self-ACC-FOC stab.PTCP-get-PST-3SGM
S. ravi thaman-tə(-mə) ænə-**gath**tha
Ravi self-DAT-FOC stab.PTCP-get.PST

'Ravi stabbed himself.'

(14) b. T. sri thirapp-ai thedi-**kond-**iru-nth-aan
Sri key-ACC search.PTCP-get.PTCP-be-PST-3SGM
S. sri yathurə hoya-**genə** hitiya
Sri key search.PTCP-get.PTCP be.PST

'Sri was searching for the key.'

Note that (14a) implies that the action has taken place (perfective), while (14b) implies that the action was in progress (continuous).

(ii) *iru* in Tamil and *ində/thiyandə* in Sinhala 'be'

This auxiliary which is used as one of the two copulas in the two languages (3.6.4) functions as the finite verb in a compound verb phrase to express progressive aspect, often with the auxiliary *kol* or *gandə* 'get' (see 14b above), as in (15) (adapted from ex. (45), p. 248, in Chandralal (2010) translated into Tamil):

(15) T. oru naal singam ondu than-da kuha-il
One(INDF) day lion one(INDF) self-GEN den-LOC
thungi-kondu-**iru-ntha-thu**
sleep.PTCP-get.PTCP be-PST-3SGN
S. dawas-ak sinha-ek taman-ge guhaa-we nidaa-genə **hitiya**
day-INDF lion-INDF self-GEN den-LOC sleep.PTCP-get.PTCP be.PST

'One day a lion was sleeping in his den.'

Note that in (15) the auxiliary *kondu* in Tamil/*genə* in Sinhala which occurs as a participle, as in (14b), preceding the finite verb, conveys continuity of action.

In the two languages the same auxiliary among others is also used to express perfective aspect, as in (16):

(16) T. teacher vahuppu-kku vanthu **irru-kkir-aa**
Teacher class-DAT come.PTCP be-PRS-3SGF(HON)
S. tiichər panthiyə-tə æwilla **innəwa**
Teacher class-DAT come.PTCP be.PRS

'Teacher has come to the class.'

Note that *iru* in Tamil and *ində/thiyandə* in Sinhala 'be' is also used as an auxiliary in negative imperatives and infinitives (tables 3.5 and 3.6 respectively in 3.6.1.2).

(iii) *vaa/poo* in Tamil and *endə/yandə* in Sinhala 'come/go'

This auxiliary is used as the finite verb after the participle forms of the main verb and the auxiliary *kol/gandə* to express continuity, specifically that the action gradually progresses, as in (17):

(17) T. naalukkunaal saaman villai eeri-kondu-**poo-/varu-kir-athu**
Day by day things price increase.PTCP-get.PTCP-go/come-PRS-3SGN
S. davəsin davəsə badu milə nægə-genə **yanəwa/enəwa**
Day by day things price increase.PTCP-get.PTCP go/come.PRS

'Things go up in price day by day.'

Note that *poo* in Tamil and *yandə* in Sinhala 'go' is also used in the negative future adjectival participles and nominalized forms as an auxiliary (3.6.1.2).

(iv) *paar* in Tamil and *balandə* in Sinhala 'see'

This auxiliary is used as the finite verb after the participle form of the main verb to express the notion of 'assess[ing]/check[ing] the nature/quality of a person or a thing' (Lehmann 1989), as in (18):

(18) T. avan muthalali-yoda kathaithu **paar-th-aan**
He manager-COM speak.PTCP see-PST-3SGM
S. eyaa mudəlaali-ekkə kathaakərəla **bæluwa**
He manager-COM speak.PTCP see.PST

'He spoke to the manager (to find what kind of person s/he is).'

(v) *vai* in Tamil and *thiyandə* in Sinhala 'keep'

This auxiliary is also used as the finite verb after the participle form of the main verb to express that the action of the main verb is performed for some purpose, benefit etc. (Lehmann 1989), as in (19):

(19) T. kavitha pencil-ai theeti **vai-th-aal**
Kavitha pencil-ACC sharpen.PTCP keep-PST-3SGF
S. kavitha pænsələ ulkərəla **thibba**
Kavitha pencil sharpen.PTCP keep.PST

'Kavitha sharpened the pencil (in order that she could draw the diagram).'

(vi) *poodu* in Tamil and *daandə* in Sinhala 'put/put away'

This auxiliary is also used as the finite verb after the participle form of the main verb to express the speaker's sorrow over/dissatisfaction with something disturbing/unpleasant, as in (20) adapted from ex. (108), p 147, in Chandralal 2010 translated into Tamil):

(20) T. Ø ennath-ai natt-aalum kaatu yaanai-kal alithu
What-ACC grow.PTCP-CONC wild elephant-PL destroy.PTCP
poodu-thu-kal
put-PRS-3PL
S. Ø monəwa hæduwa-th wal ali vinaasəkərəla **daanəwa**
What grow.PTCP-CONC wild elephants destroy.PTCP put.PRS

'The wild elephants destroy whatever we plant.'

(vii) *padu-* in Tamil 'feel/experience/' and *wendə* in Sinhala 'become/happen'

This auxiliary is also used as the finite verb with adjectives expressing states of mind such as happiness, sadness, anger, as in (21):

(21) T. appa kavalai-**pat-t-aar**
Father sad-feel-PST-3SGM(HON)
S. thaaththaa duk-**una**
Father sad-feel.PST

'Father felt sad.'

In (21S) *una* is the past form of *wendə.*

(viii) *sei-* in Tamil and *kərandə* in Sinhala 'do'

This auxiliary is combined with nouns in the two languages to produce verbs of these nouns/adjectives given in table 4.1.

	Tamil	**Sinhala**
1.	*uthavi-sei* ‘help-do’ = ‘help’	*udau-kərandə*
2.	*velai-sei* ‘work-do’ = ‘work’	*wædə-kərandə*
3.	*nanmei-sei* ‘good-do’ = ‘do good’	*hondə-kərandə*
4.	*muetchi-sei* ‘try-do’ = ‘try’	*uthsaha-kərandə*
5.	*theemei-sei* ‘bad/evil-do’ = ‘do evil’	*narəkə-kərandə*
6.	*neetu* ‘lengthen’	*dik-kərandə* ‘lengthen-do’ = ‘lengthen’
7.	*kurai* ‘shorten/reduce’	*kotə/adu-kərandə* ‘shorten/reduce-do’=‘shorten/reduce’
8.	*padi* ‘study’	*paadam-kərandə* ‘study-do’ = ‘study’
9.	*vilaiyaadu* ‘play’	*sellam-kərandə* ‘play-do’ = ‘play’
10.	*pakidi-vidu* ‘joke-leave’ = ‘joke’	*vihiilu-kərandə* ‘joke-do’ = ‘joke’
11.	*sellam-kudu* ‘pet-give’ = ‘pet’	*hurəthal-kərandə* ‘pet-do’ = ‘pet’
12.	*kalavu-edu* ‘steal-take’ = ‘steal’	*horəkam-kərandə* ‘steal do’ = ‘steal’

Table 4.1: ‘Do’ verbs

The verbs containing ‘do’ in Tamil are relatively fewer in number than in Sinhala in that only verbs (1–5) in Tamil have ‘do’ as an auxiliary, whereas all the verbs in Sinhala have ‘do’.

As (i)–(viii) above make clear, there are several noticeable cases of auxiliaries which have been grammaticalized from the same lexical verbs in the two languages.

4.3.4 Auxiliaries conveying modality

In addition to the auxiliaries discussed above, there are other constituents which make up complex VPs in the two languages. This section analyzes some of these further constituents, especially evidentials and modals involved in conveying different kinds of modality—the speaker’s attitude towards or assessment of the event or state, expressing meanings such as necessity, certainty, possibility, etc. The literature abounds with subclassifications of the general concept of modality (see Palmer 2001, Portner 2009 and Saeed 2003). Here, the classification given in Velupillai (2012) is adopted as it is linguistic-typology based. Modality is divided into two kinds: propositional modality and event modality. Propositional modality concerns the speaker's attitudes towards the truth value of the information given in the proposition, while event modality deals with potential actions which have not yet been realized, but are possible or probable (Velupillai 2012).[60]

Propositional modality can be further divided into two subcategories: evidential and epistemic modality. Evidentials encode the type of evidence that the speaker has for his/her statement. This evidence can be either direct or indirect. Direct evidentials are used when the speaker has some sort of sensory evidence, mostly visual or auditory, for the action or event s/he is describing. Indirect evidentials are used 'when the speaker was not a witness to the event but learned of it after the fact' (de Haan 2013a). They are of two kinds, namely inferential evidential and quotative evidential. Tamil and Sinhala do not have direct evidentials, but they have both kinds of indirect evidentials.

The inferential evidential is used when the speaker draws an inference on the basis of available physical evidence. If the available physical evidence for the event/state is less than solid, inferential evidentiality is marked in Tamil and Sinhala using the inferential clitic *poola* and *wage* respectively—grammaticalized forms of the postposition meaning 'like' as in (22) (auxiliaries here and below are in boldface):

(22) T. kootam   mudi-nj-(u)thu-**poola**
        Meeting   finish-PST-3SGN-INFER
      S. rasviimə   ivərai   **wage**
        Meeting   finish.PST   INFER

'It seems that the meeting has finished.'

Note that the finite form of the verb is used in (22). If the available physical evidence for the state/event is strong, Tamil uses the form of 'be/exist' *irrukum* or the infinitive form of the verb 'be' together with the modal auxiliary *venum* 'should/need' to form *irukka venum*, whereas Sinhala uses *æthi* 'might be' (Gair 1998b, 2007) or the infinitive form of the verb *wendə* 'happen' together with either *æthi* or the modal auxiliary *oone* 'should/need' to form *wendə æthi* and *wendə oone* respctively, as in (23):

---

[60] For a succinct account of the different kinds of modality, see Velupillai (2012) which is the source of much of the information discussed in this sub-section.

(23) T. kootam mudinju-**iru-kkum**/ **-irrukk-a** **venum**
Meeting finish.PTCP-exist-FUT.3SGN exist-INF should
S. rasviimə ivərə **æthi/** **wendə** **æthi/** **wendə** **oone**
Meeting finish.PTCP might be happen.INF might be happen.INF should

‘The meeting must be/have finished.’

Note that Chandralal (2010) notes that *æthi* in Sinhala (23) “substitutes for a ‘be’ verb”. Gair (1998b, 2007) classifies it as a quasi verb along with its negative counterpart *næthi*, among others because these verbs do not inflect for tense, but have verbal properties, including inflection for some of the categories pertinent to verbs such as focusing, relativizing and conjunctive participle forms. He (2007: 869) adds that ‘They characteristically occur alone as predicators of sentences or with a dependent form, usually an infinitive in a semantically modal function’ (23).

The quotative evidential (a.k.a. reportative, hearsay, or second-hand evidential) is used when the speaker has been told about the action or event by another person. Quotative evidentiality is marked in the two languages via a verbal clitic, as in (24):[61]

(24) T. ungal-udaia kaar-ai avan vaangu-v-aan-**aam**
You(HON)-GEN car-ACC he buy-FUT-3SGM-QUOT
S. oyaa-ge kaar-ekə eyaa gannəwaa-**lu**
You-GEN car-one he buy.FUT-QUOT

‘(It is said that) he will/would buy your car.’

Epistemic modality, the second subcategory of propositional modality, codes the speaker’s qualitative judgment of the proposition, that is, to what extent the speaker is certain about a given proposition. Thus, epistemic modality expresses the notions of necessity and possibility. These are conveyed in the same manner in the two languages, as can be seen in (25) and (26) respectively:[62]

---

[61] This is different from the complementizer *endu* in Tamil and *kiyəla* in Sinhala ‘that’, also referred to as ‘quotative complementizer’ used in utterance complement clauses (4.2.7).

[62] These and others below are the most common ways in which desiderative modality and possibility (supposition) are conveyed in the two languages.

Necessity:

(25) T. shanthi nalla padikk-a **venum**
Shanthi well study-INF should
S. shanthi hondətə igenəgandə **oone**
Shanthi well study.INF should

‘Shanthi should study hard.’

Possibility:

(26) T. avarkal naalai-kku engal-ai raa saapaatu-kku koopid-a-**laam**
They tomorrow-DAT we-ACC night meal-DAT call-INF-may
S. eyaala hetə api-wə rææ kææmə-tə andəgahandə **puluwaŋ**
They tomorrow we-ACC night meal-DAT call.INF may

‘They might invite us for dinner tomorrow.’

Modal auxiliaries in the two languages co-occur with the infinitive form of the verb (25,26).

Event modality, the second kind of modality mentioned above, is about events that have not taken place, but are imminent. Event modality is of two types: deontic modality and dynamic modality. Only the concepts relevant to these two types are discussed below; the different sub types of deontic and dynamic modalities are not discussed for want of space. Deontic modality is concerned with events that are initiated or conditioned by external factors such as obligation, permission and order, that is, the subject is obliged/permitted/ordered to perform the action; these events may also include those that are possible. (27)–(29) below expresses notions of obligation, permission being granted, possibility and ability respectively:

Obligation:

(27) T. nee intha veelai-ai seithu mudikk-a **venum**
You this work-ACC do.PTCP finish-INF should
S. oyaa mee wæde kərəla ivərəkərannə **oone**
You this work do.PTCP finish.INF should

‘You should finish this work.’

Permission:

(28) T. neeng-al ippa veelai-ai nippaat-a-**laam**
You-PL now work-ACC stop-INF-may
S. oyaala-tə thæŋ wædə nathərəkərannə **puluwaŋ**
You-PL-DAT now work stop.INF may/can

‘You may stop work now.’

Possibility:

(29) T. indai-kku pinneram malai peiy-a-**laam**
Today-DAT evening rain fall-INF-may
S. adə hawəsə-tə vahində **puluwaŋ**
Today evening-DAT rain.INF may/can

‘It may rain this evening.’

The same auxiliary that is used to convey necessity in (25) (*venum* in Tamil and *oone* in Sinhala ‘should’) is used to convey obligation in (27). Both Tamil and Sinhala use the auxiliary (*-laam* in Tamil and *puluwaŋ* ‘may/can’ in Sinhala) used in (26) to convey the notions of permission (being granted) in (28) and possibility in (29). The most common modal auxiliary in Tamil used to convey possibility is *-laam* ‘may’, while in Sinhala, *puluwaŋ* ‘may/can’ is used.

Dynamic modality, the other kind of event modality, is concerned with events that are initiated or conditioned by internal factors such as ability, that is, the subject is able or willing to perform the action.

Physical ability:

(30) T. ena-kku antha petti-ai thuukk-a **eelum/mudium**
I-DAT that box-ACC lift-INF can
S. matə ee pettiyə ussandə **puluwaŋ**
I-DAT that box lift.INF can

‘I can lift that box.’

In (30) the auxiliaries *eelum/mudium* in Tamil and *puluwaŋ* in Sinhala ‘can’, are used to express the notion of ability.[63]

---

[63] The modals which express the different kinds of epistemic modality have distinct negative forms in the two languages: the negative forms of the modals expressing possibility, physical ability etc. are *eelaathu, mudiaathu* in Tamil and *bææ* in Sinhala ‘cannot/may not’; and the negative forms of those expressing necessity are *veendaam* ‘need not, *thevaiillai* don’t need’ in Tamil and *oone nææ* ‘need not’, *epa* ‘don’t need’ in Sinhala.

By way of summary, the different kinds of modality are shown in figure 4.1:

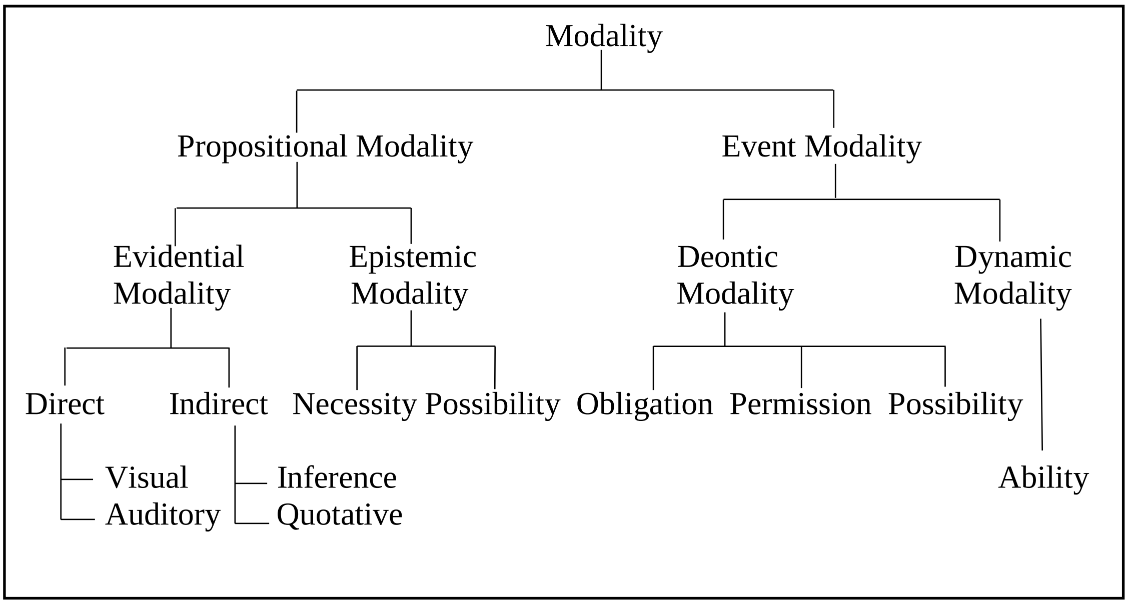


Figure 4.1: Classification of modality

Although the two languages employ different lexical items, the morphosyntactic features used to convey different kinds of modalities, namely indirect evidentials, epistemic and event modalities are in many ways similar.

4.3.5 Coordination

Coordination links two constituents that are at the same syntactic level, that is, they are not dependent on each other. The grammatical element that links constituents are coordinators. According to Velupillai (2012), there are three kinds of coordinators: conjunctive coordinator, ‘and’, disjunctive coordinator, ‘or’ and adversative coordinator, ‘but’. Coordination is of two kinds: asyndetic coordination and syndetic coordination. The former uses no coordinator to link the two units (so they are simply juxtaposed), while the latter uses coordinators to link the units. Syndetic coordination is either monosyndetic, which uses one coordinator, or bisyndetic, which uses two or more coordinators—the coordinator being marked on each CONJ(unct)/DISJ(unct). The conjoining of two or more constituents in Tamil and Sinhala using the conjunctive coordinator involves bisyndetic coordination in that a clitic (*-um* in Tamil and *-(u)i* or *-th* in Sinhala) is added to the conjuncts, as in (31) (ex. (13), p 183 Chandralal 2010 translated into Tamil) in which three nominal phrases are conjoined) (coordinators here and below are in boldface):

(31) T. aasiriyar-kal-**um** pettoor-**um** pillaikal-**um** paathai-ai
Teacher-PL-CONJ parent-PL-CONJ children-CONJ road-ACC
suththam sei-th-aarkal
clean do-PST-3PL
S. guruwər-**ui** demaupiyo-**i** laməin-**ui** paarə
Teacher-CONJ parent-CONJ children-CONJ road
suddə kəlaa
clean do.PST

‘The teachers, parents and children cleaned the road.’

All kinds of constituents can be conjoined in this way but not finite clauses. Should it be necessary to conjoin two finite clauses with a temporal relation, they are juxtaposed in the chronological order in which the events take place, as in (32):

(32) T. nimal veed-ai puuti-n-aan Ø thirap-ai baag-il poo-t-aan
Nimal house-ACC lock-PST-3SGM key-ACC bag-LOC put-PST-3SGM
S. nimal gee væhuwwa Ø yathura bææg-ek-e dæmma
Nimal house lock.PST key bag-INDF-LOC put.PST

‘Nimal locked the house and put the key in the bag.’

In (32) the event encoded in the clause to the left took place before the one encoded in the clause to the right.

However, modal verb constructions can be conjoined using the normal conjunctive clitics *-um* in Tamil and *-th* in Sinhala, as in (33) (ex. (15), p. 184, Chandralal 2010 translated into Tamil):

(33) T. engalu-kku veedu-kal katt-a-(v)**um** eelum idikk-a-(v)**um** eelum
We-DAT house-PL build-INF-CONJ can break-INF-CONJ can
S. apə-tə gewal hadanna-**th** puluwan kadanna-**th** puluwaŋ
We-DAT house.PL make.INF-CONJ can break.INF-CONJ can

‘We can build houses and also destroy houses.’

Chandralal (2010: 184) is of the view that in Sinhala, the conjunctive coordinator *-th* can be distinguished from *-ui* ‘in that the former has an emphatic overtone.’ This view seems to be acceptable because *-ui* cannot be used in instances, like (33S), which are articulated with an emphasis on the two conjuncts.

Tamil and Sinhala have a number of disjunctive coordinators to conjoin nominal or verbal constituents. The disjunctive clitic *-oo* in Tamil and *-ho* in Sinhala is a bisyndetic disjunctive coordinator, as in (34a). In addition, Tamil and Sinhala use *sari* and *hari* respectively as disjunctive coordinators, as in (34b) (here, both coordinate verbal constituents).

(34) a. T. rupa sapid-a-(v)**oo** kulikk-a-(v)**oo** pooi irukk-a
Rupa eat-INF-DISJ bathe-INF- DISJ go.PTCP be-INF
venum
should (INFER)
S. rupa kandə **ho** naandə **ho** gihilla æthi
Rupa eat.INF DISJ bathe.INF DISJ go.PTCP be (INFER)

'Rupa must have gone to eat or bathe.'

(34) b. T. rupa saapid-a **sari** kulikk-a **sari** pooi irukk-a
Rupa eat-INF DISJ bathe-INF DISJ go.PTCP be-INF
venum
should (INFER)
S. rupa kandə **hari** naandə **hari** gihilla æthi
Rupa eat.INF DISJ bathe.INF DISJ go.PTCP be (INFER)

'Rupa must have gone to eat or bathe.'

The Tamil disjunctive coordinator *-oo* is frequent; Sinhala *ho* is somewhat rarer, especially in spoken Sinhala. The disjunctive coordinator *hari* in Sinhala is more frequently used than *sari* in Tamil.

The two languages also have monosyndectic disjunctive coordination, which involves the use of disjunctive particle *illaiendaal* in Tamil and *næthnaŋ in* Sinhala, as in (35), in which two nominal constituents are coordinated:

(35) T. naan ravi-oda **illaiendaal** mohan-oda opis-ukku poo-v-een
I Ravi-COM DISJ Mohan-COM office-DAT go-FUT-1SG
S. maŋ ravi-ekkə **næthnaŋ** mohan-ekkə kanthooruwə-tə yanəwa
I Ravi-COM DISJ Mohan-COM office-DAT go.PRS/FUT

'I will go to the office with Ravi or Mohan.'

Finally, the same disjunctive coordinator seen in (35T,S) is used together with *onril* in Tamil and *ekkoo* in Sinhala at the beginning of the sentence to exclusively refer to one entity or event, akin to the *either…or* construction in English, as in (36) (ex. (25) p. 186, Chandralal 2010 translated into Tamil):

(36) T. **ondil** veet-il vittitu vanth-itt-een **illaiendal** kaaru-kulla
Either home-LOC leave.PTCP come.PTCP-PST-1SG DISJ car-LOC
irukk-a venum
be-INF should (INFER)
S. **ekkoo** gedərə daala æwilla **næthnaŋ** kaarek-ee æthi
Either home put.PTCP come.PST DISJ car-LOC be (INFER)

‘Either I have left it at home or it might be in the car.’

4.3.6 Alternative questions

An alternative question is a question that presents two or more answer options and presupposes that only one is true. In Tamil and Sinhala, the question clitic *-aa* and the question particle *-də* respectively are added to each of the two constituents that encode the possible answers, as in (37) (the question clitic and particle are in boldface):

(37) T. una-kku tee-**aa** coopi-**aa** venum
You-DAT tea-Q coffee-Q want.PRS.3SGN
S. oyaa-tə thee **də** coopi **də** oone
You-DAT tea Q coffee Q want.PRS

‘Do you want to have tea or coffee?’

4.3.7 Quantifiers

In both Tamil and Sinhala, quantifiers are formed by adding the conjunctive clitic and the disjunctive clitic to wh-words/phrases: adding the CONJ clitic, *-um* in Tamil and *-th* in Sinhala to wh-words/phrases yields universal quantifiers, while adding a slight variant of the DISJ clitic *-oo* to wh-words/phrases yields existential quantifiers (see 4.3.5).[64] To illustrate, the commonest quantifiers in Tamil and Sinhala are given in table 4.2:

---

[64] Note that *-ho* is the disjunctive particle in Sinhala (4.3.5).

Universal Quantifiers		**Existential Quantifiers**	
Tamil	**Sinhala**	**Tamil**	**Sinhala**
yaar-um who-CONJ	*kauruwa-th* who-CONJ	*yaar-oo* who- DISJ	*kaud-oo* who- DISJ
‘anyone’		‘someone’	
ethu-(v)um what-CONJ	*mokak(wa)-th* what-CONJ	*eth-oo* what- DISJ	*mokakd-oo* what- DISJ
‘anything’		‘something’	
enge-um where-CONJ	*kohe(wa)-th* where-CONJ	*enge-oo* where- DISJ	*kohed-oo* where- DISJ
‘anywhere’		‘somewhere’	
eppa-(v)um *eppooth-um* when-CONJ	*kawada(wa)-th* when-CONJ	*eppai-oo* *eppooth-oo* when- DISJ	*kawədad-oo* when- DISJ
‘ever/at any time’		‘at some time’	

Table 4.2: Universal and existential quantifiers

4.3.8 Verbal nominals

Tamil and Sinhala both use verbal nominals, i.e., nouns which are formed by adding distinct nominalizer suffixes to verbs or adjectives. The verbal nominals formed thus can be divided into tensed and untensed verbal nouns (Lehmann 1989). The commonest nominalizer suffixes used in the formation of verbal nouns are given in table 4.3 (verbal nominals is a broad area in the two languages, see Lehmann (1989) for Tamil and Chandralal (2010) for Sinhala):

Tamil	**Sinhala**
(1) Tensed: adding the nominalizer suffix to adjectival participle:	
-athu	***-ekə***
thedu-(k)ir- + athu = thedu-(k)ir-athu	*hoyənə + ekə = hoyənə ekə*
search-PRS NMLZ searching	search.PRS NMLZ searching
sei-(k)ir- + athu = sei-(k)ir-athu	*kərənə + ekə = kərənə ekə*
do-PRS NMLZ doing	do.PRS NMLZ doing
saapi-tt- + athu = saapi-tt-athu	*kaapu + ekə = kaapu ekə*
eat-PST NMLZ eating	eat.PST NMLZ eating
(2) Untensed: adding the nominalizer suffix to the verb stem:	
(i) *-al* or *-thal*	***-iimə* or *-illə***
vettu + thal = vettuthal	*kapə + -iimə/illə = kæpiimə/kæpillə*
cut NMLZ cutting	cut NMLZ cutting
sei + thal = seithal	*natə + iimə/illə = nætiimə/nætillə*
do NMLZ running	dance NMLZ dancing
illu + thal = illuthal	*ædə + iimə/illə = ædiimə/ædillə*
pull NMLZ pulling	pull NMLZ pulling
(ii) *-kai*	***-um***
vaal + kai = vaalkai	*rawə + um = ræwum*
live NMLZ life	frown NMLZ frown
sei + kai = seikai	*natə + um = nætum*
do NMLZ action	dance NMLZ dance
vaa(ru) + kai = varukai	*bani+ um = bænum*
come NMLZ arrival	scold NMLZ scolding

Table 4.3: Verbal nominals

The tensed forms (1) are formed by adding the suffix to verbal adjectives, e.g., ‘searching (in the present)’ and ‘eating (in the past)’. Of the untensed verbal nominals (2), those which belong to type (i) are ‘verbal nouns denoting actions’, e.g.,‘cutting’, ‘running’ and ‘pulling’, whereas those belonging to type (ii) are ‘more consolidated as nouns’ (Chandralal 2010: 49), e.g., ‘action’, ‘life’,‘arrival’, ‘frown’ and ‘dance’.

4.3.9 Volitive/involitive verbs in Sinhala and effective/affective verbs in Tamil

One of the distinctive features of Sinhala well-attested in the literature is the distinction between volitive and involitive verbs. Chou and Hettiarachchi (2013) define volitive verbs as those denoting volitional or intentional action; and involitive verbs as those denoting non-volitional or unintentional states of affairs, also referred to as anticausative verbs, as illustrated respectively in (38aS,bS) (exs. (3) a. and b., p 153, Chandralal 2010):

(38) a. S. ranjit puusa(-wə) pææguwa
Ranjit cat(-ACC) trample.VOL.PST
'Ranjit trampled the cat.'

(38) b. S. ranjit-tə puusa(-wə) pææguna
Ranjit-DAT cat(-ACC) step on.INVL.PST
'Ranjit accidentally stepped on the cat.'

Note that the involitive verb in (38bS) requires a dative subject. The closest Tamil equivalents of (38aS,bS) are (39aT,bT) respectively. In order to show that the action is non-volitional, that is, inadvertent, the perfective aspect together with the adverbial 'accidentally' is used, as in (39bT), whereas the past tense is used to show that the action is volitional in (39aT).

(39) a. T. ranjit puunai-ai mithi-th-aan
Ranjit cat-ACC trample-PST-3SGM
'Ranjit trampled the cat.'

(39) b. T. ranjit puunai-ai (thatseyalaha) mithith-itt-aan
Ranjit cat-ACC accidentally step on.PTCP-leave.PST-3SGM
'Ranjit has accidentally stepped on the cat.'

It is useful to note that the verb in (39bT) does not require a dative subject, unlike the involitive verb in (38bS).

'The distinction between the two kinds of verbs in Sinhala,' as Jany (2006: 69) points out, 'is for the most part encoded in the verbal derivational morphology.' Some involitive verbs and their volitive counterparts are given in table 4.4 below (for details, see Jany 2006, Henadeerage 2002, Gair 1998b, Chandralal 2010 and Chou and Hettiarachchi 2013).

Involitive verbs	**Volitive verbs**
wætenəwa 'to fall'	*wattənəwa* 'to drop'
mærenəwa 'to die'	*marənəwa* 'to kill'
æhenəwa 'to hear'	*ahanəwa* 'to listen'
issenəwa 'to rise'	*ussənəwa* 'to lift'
ærenəwa 'to open automatically'	*arinəwa* 'to open'
kædenəwa 'to break naturally '	*kadənəwa* 'to break'
pælenəwa 'to split or break'	*palənəwa* 'to split'
kærəkenəwa 'to turn automatically'	*karəkənəwa* 'to turn'
gæləwenəwa 'to get removed naturally'	*galəwənəwa* 'to remove or dismantle something'

Table 4.4: Involitive and volitive verbs in Sinhala

A similar kind of distinction, though not identical to that in Sinhala, is found in some verbs in Tamil. Paramasivam (1979; cited in Lehmann 1989:50), distinguishes between affective verbs and effective verbs in Tamil: an affective verb is a kind of verb, the subject of which undergoes the action (or state or change of state) described by the verb stem; an effective verb, on the other hand, is a kind of verb which represents an action of the subject.' (40aT,bT) (exs. (105) and (106), p. 50, Lehmann 1989) illustrate the affective and effective verbs respectively in Tamil:

(40) a. T. avan-udaiya talai tirumb-iy-athu
He-GEN head turn-PST-3SGN
'His head turned.'

(40) b. T. avan talaiy-ai tirupp-in-aan
He head-ACC turn-PST-3SGM
'He turned his head.'

The affective verbs in Tamil seem to correspond to the involitive verbs in Sinhala, while the effective verbs, to volitive verbs, as illustrated respectively in (41aS,bS) (the equivalents of (40aT,bT) show:

(41) a. S. eyaa-ge oluwə hæruna
He-GEN head turn.INVL.PST
'His head turned.'

(41) b. S. eyaa oluwə hærewwa
He head turn.VOL.PST
'He turned his head.'

Some affective verbs and their effective counterparts in Tamil are given in table 4.5 (see Lehmann 1989: 50–52):

Affective Verbs	Effective Verbs
peruku ‘something increases’	*perukku* ‘someone increases …’
thirunthu ‘someone something becomes better/correct’	*thiruththu* ‘someone repairs …’
nirambu ‘something becomes full’	*nirappu* ‘someone fills …’
oodu ‘someone/animal runs’	*oottu* ‘someone drives…’

Table 4.5: Affective and effective verbs in Tamil

4.3.10 Final complementizer and final/medial polar question particle in embedded clauses

In Tamil and Sinhala, the COMP(lementizer) which occurs in utterance complement clauses (4.2.7) is always clause final. This clause final complementizer is also referred to as a ‘quotative complementizer’. If the embedded clause is a polar question, the question clitic/particle which occurs clause-finally in the embedded clause precedes the COMP in both languages, as in (42) (question clitic, particle and complementizers are in boldface):

(42) T. anand enna-ta [lalith koota-(thu)kku varu-v-aan-**aa** **endu**]
Anand I-DAT Lalith meeting-DAT come-FUT-3SGM-Q COMP
kee-tt-aan
ask-PST-3SGM
S. anand magen [lalith ræsveemə-tə enəwa **də** **kiyəla**] æhuwa
Anand I.ABL Lalith meeting-DAT come.PRS Q COMP ask.PST

‘Anand asked me whether/if Lalith will come for the meeting.’

4.3.11 Dative subject constructions

Included in this chapter for completeness of the description of the morphosyntactic features of Tamil and Sinhala are several well-attested areal features (4.3.11–4.3.13). Most of them are thought to be Dravidian in origin, but diffused from Dravidian into NIA languages (Schiffman 2010, Emeneau 1956, Ebert 2010, and Thomason 2000).

The dative subject construction, as the term implies, is a construction in which the dative case is assigned to the subject with a specific subset of verbs. Lehmann

(1989: 184) notes that dative subjects in Tamil are licensed by ‘verbs of mental experience’ such as ‘know’, ‘understand’; ‘verbs of emotional experience’ such as ‘like’; and ‘verbs of physical and biological experience’ such as ‘hungry’, ‘ache’, ‘itch’ etc. The same set of verbs licenses dative subjects in Sinhala as well. The dative subjects in the two languages are illustrated in (43):

(43) T. ena-kku   leela-(v)ai   therium
          I-DAT   Leela-ACC   know.PRS.3SGN
       S. matə   leela-wə   dannəwa
          I-DAT   Leela-ACC   know.PRS

‘I know Leela.’

In addition, modal auxiliaries in the two languages such as *venum* in Tamil/*oone* in Sinhala ‘should’, *eelum* in Tamil/*puluwaŋ* in Sinhala ‘can’ etc. also license dative subjects, as in (25) repeated as (44):

(44) T. ravi-kku   rendu   kai-aal-um   eluth-a   eelum
          Ravi-DAT   two   hand-INS.INCL   write-INF   can
       S. ravi-tə   ath   dek-en-mə   liyandə   puluwaŋ
          Ravi-DAT   hand   two-INS-FOC   write.INF   can

‘Ravi can write with both hands.’

4.3.12 Reduplication

Reduplication is another feature which is used frequently in the two languages. It is a morphological process whereby a set amount of phonological material is copied from a base form and fused with it to form a stem onto which other morphemes may then be added. Velupillai (2012: 101) argues that the form of the reduplicant (the repeated element) is dependent on the form of the base because it is part of the latter that is being repeated. Given in table 4.6 are some common reduplicated forms in the two languages. The lexical items that are reduplicated, the syntactic properties of these items and the semantic functions that the reduplicated forms perform are the same in the two languages (see Lehmann 1989 for Tamil and Chandralal 2010 for Sinhala).[65]

---

[65] Excluded in table 4.6 is the reciprocal pronoun in the two languages discussed in 3.6.

Base	**Tamil**	**Sinhala**
1) Infinitive forms of the verb: to express continuity or iterativity		
study	*padikka-padikka*	*igenəgandə-igenəgandə*
	‘the more someone studies…’	
play	*vilaiaada-vilaiaada*	*sellaŋkərandə-sellaŋkərandə*
	‘the more someone plays …’	
cook	*samaikka-samaikka*	*uyandə-uyandə*
	‘the more someone cooks…’	
2) Past participle form of the verb: to express iterativity		
dance	*aadi-aadi*	*natəla-natəla*
	‘danced continuously’	
beat	*adithu-adithu*	*gahala-gahala*
	‘beaten repeatedly’	
see	*paarthu-paarthu*	*baləla-baləla*
	‘saw repeatedly’	
3) Adjective and adverbs: to function as intensifiers		
small	*sinna-sinna pillaikal*	*podi-podi lamay*
	‘small children’	
good	*nalla-nalla paatukal*	*hondə-hondə sindu*
	‘good songs’	
round	*vatta-vatta mesaikal*	*raum-raum meesə*
	‘round tables’	
slow	*mella-mella vanthaan*	*hemin-hemin aawa*
	‘(He) came slowly.’	
4) Wh-questions: to convey distributive meaning, referring to members of a group or objects of a set individually		
wh-phrases	*yaar yaar vanthaarkal*	*kaudə kaudə aawe*
	‘Who all (are the people who) came?’	
	enge enge poonai	*kohedə kohedə giye*
	‘Where did (you) go?/What are the places (you) went to?’	

Table 4.6: Reduplicated forms

4.3.13 Converbs

Converbs (a.k.a conjunctive participle etc.) are another distinct areal feature of the region, which Tamil and Sinhala share. Converbs are non-finite verb forms that denote subordinate adverbial, that is, they occur in subordinate clauses and express temporal, modifying or causal relations with the actions expressed in the main clause. By default, these subordinate clauses precede the main (finite) clauses in Tamil and Sinhala (cf. compound verbs in 4.3.3). The causal relation between the subordinate and main clauses (see also 4.2.6) in the two languages is conveyed by the perfect participle in the subordinate clause, as in (45):

(45) T. amma koopittu nimal padikk-a poo-n-aan
Mother call.PTCP Nimal study-INF go-PST-3SGM
S. amma andəgahala nimal paadaŋ kərandə giya
Mother call.PTCP Nimal study do.INF go.PST

‘Having been called by his mother/Since mother called, Nimal went to study.’

(46) below expresses a modifying relation:

(46) T. pillai appa-itta oodi va-nth-athu
Child father-LOC run.PTCP come-PST-3SGN
S. laməya thaaththaa-langə-tə duwəla aawa
Child father-LOC-DAT run.PTCP come.PST

‘The child came running to his/her father.’

In (46), the non-finite verb (participle) functions as the gerund. This kind of modifying relation is also expressed by REDUP(licated) converbs, as in (47):

(47) T. pillai appa-itta paainthu paainthu poo-n-athu
Child father-LOC jump.PTCP jump.PTCP (REDUP) go-PST-3SGN
S. laməya thaaththa langə-tə panə(la) panə(la) giya
Child father LOC-DAT jump-PTCP jump.PTCP (REDUP) go.PST

‘The child went hopping to his/her father.’

The use of reduplication in instances like this is another areal feature of the region.

Tamil and Sinhala also have converb constructions, like (48), which Jayaseelan (2004) distinguishes as serial verb constructions (adapted from ex. 1c. (Malayalam), p. 67, Jayaseelan 2004 translated into Tamil and Sinhala):

(48) T. naan oru maanga pudungi kaluvi vetti uppu pootu
I one mango pluck.PTCP wash.PTCP cut.PTCP salt put.PTCP
saap-itt-een
eat-PST-1SG
S. mamə ambə gediy-ak kadəla soothəla kapəla
I mango fruit-one pluck.PTCP wash.PTCP cut.PTCP
lunu daala kææwa
salt put.PTCP eat.PST

‘I plucked, washed, cut, put salt and ate a mango.’

Aikhenvald (2006:1) defines the serial verb construction as ‘a sequence of verbs which act together as a single predicate, without any overt marker of coordination, subordination, or syntactic dependency of any other sort.’ In light of this definition, constructions like (48)—though they have no overt marker of coordination—are not serial verb constructions, as Jayaseelan (2004) refers to them as. In (48), the function of the conjunctive (non-finite) participles which linearly precede the main clause is to combine a set of events related temporally; these events do not constitute a single event; neither are they independent. Therefore, (48T,S) are converb constructions.

4.4 Discussion

In this chapter, 20 features of Tamil and Sinhala are compared in 4.2 and 4.3. The implications for the possible contact-induced changes that Sinhala may have undergone which arise from the comparative analysis are discussed below.

4.4.1 Features related to complex sentences: implications of the comparison

The typological profiles established on the basis of the comparison of seven features related to complex sentences in Tamil and Sinhala in 4.2 are given in table 4.7:

	WALS features	**Tamil (T)**	**Sinhala (S)**	**Similar/ Different**
4.2.1	Relativization on subjects (122A)	gap strategy	gap strategy	Similar
4.2.2	Relativization on obliques (123A)	gap strategy	gap strategy	Similar
4.2.3	‘Want’ complement subjects (124A)	subject is left implicit	subject is left implicit	Similar
4.2.4	Purpose clauses (125A)	deranked	deranked	Similar
4.2.5	‘When’ clauses (126A)	deranked	deranked	Similar
4.2.6	Reason clauses (127A)	deranked	deranked	Similar
4.2.7	Utterance complement clauses (128A)	balanced	balanced	Similar

Table 4.7: Typological profiles of features related to complex sentences

The typological profiles of the two languages show that all the values of the features compared in 4.2 are the same and crucially, the structures of the complex sentences are almost identical. The fact that the two languages share the same structures adds support to the idea that there has been convergence between Sinhala and Tamil.

As noted in 2.3.1.3, the prenominal relative clause between Sinhala is likely to have been restructured on the model of Tamil/Dravidian prenominal relative clause. The two kinds of relativization on subjects (4.2.1) and obliques (4.2.2) use the same gap strategy to relativize the two constituents. Jayaseelan and Amritavalli (2005) point out that the gap strategy is employed in the relativization of all Dravidian languages (also Subbarao 2010). The strategy used for the relativization on subjects and obliques (WALS 122A and 123A) in Hindi[66] is non-reduction involved in the correlative type. The use of the gap strategy in Sinhala relative clauses is a consequence of its use of the prenominal type relative clause, which shows that the Sinhala prenominal relative clause together with the gap strategy used in it has been modelled on the Tamil/Dravidian relative clause.

As for the rest of the features, the values of ‘want’ complement clauses (WALS 124A in 4.2.3) in the two languages are the same in that the notional subject of the embedded clause, which is the complement of the verb ‘want’, is left implicit in these languages. WALS 125A–128A have only three values based on the distinction between balanced (finite) and deranked (non-finite) forms of verbs and, by extension, clauses, whether they are independent or dependent. The values of the first three features (125A–127A in 4.2.4–4.2.6), namely purpose, ‘when’ and reason clauses, are deranked in that the non-finite form of the verb is used in the subordinate clauses of these complex sentences. The value of the last feature (WALS 128A in 4.2.7) is balanced in that the finite form of the verb is used in the subordinate (embedded) clause.[67] Not much evidence could be obtained from the values of these four features in arguing for or against restructuring of these clauses in Sinhala. Note also that these features have almost the same values for the seven languages of South Asia included in the research (see Appendix).

There are other morphosyntactic features related to complex sentences which point to contact-induced restructuring of Sinhala. The use of a clause final complementizer which has been grammaticalized from the verb *say* is believed to be

---

[66] Hindi is the only NIA language for which the values of these two features are determined in WALS (See Appendix).

[67] The value of this feature for Tamil as identified in WALS based on Asher (1985) is ‘deranked’ which is incorrect. It is evident from ((9) in 4.2.7) that the verb form in utterance complement clauses in Tamil is ‘balanced’. Note that the value of this feature for Malayalam, a Dravidian language with very similar morphosyntax as that of Tamil, is also ‘balanced’.

another areal feature of the South Asian languages, the source of which is probably Dravidian (Biberauer et al. 2009). Subbarao (2010: 4) claims that the complementizer in some of the NIA languages is clause-initial, and the verb ‘say’ in the complementizer is ‘in its participial, infinitival or conditional form in South Asian languages’. He adds that in Dravidian, it invariably occurs to the right of the embedded clause (clause-finally) and the verb ‘say’ is always in the conjunctive participial form. This is true for Sinhala in that the complementizer is always clause final and the COMP *kiyəla* is always the participle form of the verb ‘say’ like the Tamil complementizer. The fact that all Dravidian languages have these two features, the clause-final COMP and the COMP grammaticalized from the verb ‘say’, whereas the NIA languages do not have either one or both of them suggests that Dravidian is the source of these features which may have found their way into the NIA languages which have them, e.g., Bengali and Sinhala. However, it is difficult to establish when Sinhala got them from Tamil: whether (i) in India, through the influence of other NIA languages which got them from Dravidian, or (ii) in Sri Lanka because of its direct contact with Tamil.

Subbarao also notes that the default position of finite complement clauses in Dravidian is the preverbal argument position, that is, the position at which the direct object occurs; it may stay in this default clause medial position, as in ((9) in 4.2.7) or be extraposed to the right of the matrix clause/to the left of the subject, as in (49):

(49) T. Rohan son-n-aan [kumar avan-ra veet-ai
Rohan say-PST-3SGM Kumar he-GEN house-ACC
vith-thi-tt-aan endu]
sell.PTCP-leave-PST-3SGM COMP
S. rohan kiwwa [kumar eyaa-ge gee vikkaa kiyəla]
Rohan say.PST Kumar he-GEN house sell.PST COMP

‘Rohan said that Kumar had sold his house.’

This possibility of extraposing finite complement clauses, Subbarao notes, ‘is not found in several Indo-Aryan languages’. Its presence in Sinhala, as in (49), indicates possible influence from Tamil.

It is worth noting that the COMP or *say*-complementizer has been extended to other contexts in Dravidian. It is used as a reason marker in reason clauses which occur in complex sentences to indicate a causal relation between the two clauses (Subbarao 2010). This holds true for Sinhala too, as in (50):

(50) T. naangal koota-(thu)kku poo-v-oom endu avarkal engal-ai
We meeting-DAT go-FUT-1PL COMP they we-ACC
town-(u)kku pooh-a koopid-a-illai
town-DAT go-INF call-INF-NEG
S. Api ræsviimə-tə yai kiyəla eyaala api-tə towmə-tə
We meeting-DAT go.FUT COMP they we-DAT town-DAT
yannə kathakər-e nææ
go-INF call.PST-NMLZ NEG

'They didn't ask us to go to town thinking that (= because they thought that) we would go to the meeting.'

The same complementizer can be used as a marker for naming and labelling in a relative clause-like construction in the Dravidian languages (Subbarao 2010) and Sinhala, as in (51):

(51) T. Tolstoy *War and Peace* end-a puthakathai eluthi-n-aar
Tolstoy *War and Peace* COMP-RELAT book-ACC write-PST-3SGM(HON)
S. Tolstoy *War and Peace* kiyən-ə pothə liwwa
Tolstoy *War and Peace* COMP-RELAT book write.PST

'Tolstoy wrote the book titled *War and Peace*.'

In (51) the relativizer *-a* in Tamil and *-ə* in Sinhala is suffixed to the complementizer. Similarly, the polar clitic *-aa* in Tamil and the polar particle *də* in Sinhala can also be added to the complementizer, as in (52) ((50) slightly modified) to verify the proposition stated in the embedded clause:

(52) T. naangal kootathu-kku poo-v-oom end-aa nee-(n)gal engal-ai
We meeting-DAT go-FUT-1PL COMP-Q you-PL we-ACC
town-(u)kku pooh-a koopid-a-illai
town-DAT go-INF call-INF-NEG
S. Api ræsviimə-tə yai kiyəla də oyaa-la api-tə towmə-tə
We meeting-DAT go.FUT COMP Q you-PL we.DAT town-DAT
yannə kathakər-e næththe
go-INF call.PST-NMLZ NEG.NMLZ

'Was it because (you thought) we would go to the meeting that you did not ask us to go to town?'

In this type of sentence, the polar clitic/particle is added to the complementizer. The above discussion shows that the complementizer has been adapted to be used in other contexts.

Although most of the values of the features compared in this section seem to be shared by many of the languages of South Asia, the subtle morphosyntactic features related to complex sentences that Tamil and Sinhala have in common suggest that in complex sentences too Sinhala has undergone considerable changes on the model of Tamil.

4.4.2 Miscellaneous features: implications of the comparison

The 13 diverse morphosyntactic features of Tamil and Sinhala compared in 4.3 are given in table 4.8:[68]

	Features	**Similar/Difference**
4.3.1	Demonstrative-based third person pronouns	similar
4.3.2	Topicalization	similar
4.3.3	Auxiliaries grammaticalized from lexical verbs	similar
4.3.4	Auxiliaries conveying modality	similar
4.3.5	Coordination	similar
4.3.6	Alternative questions	similar
4.3.7	Quantifiers	similar
4.3.8	Verbal nominals	similar
4.3.9	Volitive/involitive versus effective/affective distinction	different
4.3.10	Final complementizer and final/medial polar question particle in embedded clauses	similar
4.3.11	Dative subject construction	similar
4.3.12	Reduplication	similar
4.3.13	Converbs or conjunctive participles	similar

Table 4.8: Miscellaneous morphosyntactic features

12 of the 13 morphosyntactic features compared in this section too are strikingly similar in Tamil and Sinhala. Analysis of these features suggests that Sinhala may well have adopted most of these features from Tamil. Even the features classified as areal features

---

[68] The values of these features which do not belong to WALS have not been determined. As a result, the typological profiles of these features are not established, instead whether these features are similar or different in the two languages are stated in table 4.8.

show that Sinhala has converged more with the Dravidian than with other NIA languages.

The third person pronouns in Tamil and Sinhala (4.3.1) are formed by combining the demonstrative stem with the relevant bound pronominal base encoding the feature content. The Tamil demonstrative forms are cognate with those in other Dravidian languages. Moreover, the NIA languages, e.g., Bengali, Hindi, Marathi etc. too have demonstrative based third person pronouns. Since both NIA and Dravidian languages have demonstrative-based third person pronouns and that the source of this kind of pronouns, that is, in which of the two language families they originated, is not known, it is difficult to conclude that Sinhala has acquired this kind of pronouns from Tamil. Gair and Karunatillake (2000: 716) rightly point out that 'Dravidian influence may […] be discerned in the pronominal system, in which third-person pronouns are formed from a deictic plus a nominal element, though the Sinhala deictic system is different from the Dravidian one in having four terms plus and interrogative set…'

Topicalization in the two languages (4.3.2), shows that Sinhala has replicated several features of Tamil topicalization. In Tamil, the topic marker *-endaal* (11bT) is formed by combining the *say*-complementizer *endu* (4.2.7) and the conditional suffix *-aal*, while in Sinhala *-naŋ* (11bS) which is one of the two conditional suffixes (Gair 2007). The other topic markers, *vanthu* (12T) in Tamil and *æwilla* (12S) in Sinhala, are both grammaticalized from the lexical verb 'come'. These topic markers exemplify replica grammaticalization, that is, the grammaticalization of the category in the replica language follows the path way afforded by the grammaticalization of the corresponding category in the source language. What is important to note is that the two distinct topic markers grammticalized from almost the same items in the two languages point to paired gammaticalization (1.3) attested in language contact situations which have intense contact; here, the topic marking is the grammatical function and the formation of the two topic markers are the two grammaticalization processes. The presence of two kinds of topic constructions, topicalizing i) a particular constituent (11b,12) or ii) the entire clause (13), similarly suggests that Sinhala has replicated the pattern of the topic constructions from Tamil.

Auxiliaries grammaticalized from lexical verbs in compound verbs (4.3.3), another feature which Tamil and Sinhala share, demonstrates 'shared polysemy' (1.3) in that they share the same verb-auxiliary pairing. The compound verb is widespread cross-linguistically and common to the languages of South Asia. There is no consensus among scholars regarding the origin of compound verbs; some trace it to Dravidian, while others trace it to NIA languages (see Herring 1989 and the references therein). Herring (1989: 177) argues that ‘the continuative auxiliary—and quite possibly the system of aspectual auxiliaries in general […] was borrowed into languages which had regular contact with Tamil.’ Sinhala may have acquired some of these auxiliaries as an NIA language prior to reaching Sri Lanka, but it may not have acquired all eight ((i–viii) in 4.3.3). Therefore, it seems safe to conclude that Sinhala has modelled many of the grammaticalized auxiliaries on those in Tamil for two reasons: a) almost all the common auxiliaries in the two languages have been grammaticalized from the same lexical verbs, with the same syntactic, and semantic functions; and b) the Tamil auxiliaries have cognates in other Dravidian languages (see Krishnamurti 2003). It has not been verified whether Sinhala has cognates of this kind of auxiliaries in other NIA languages.

The morphosyntactic features used to convey the two kinds of modality (4.3.4), are almost the same in the two languages. Koptjevskaja (2011) claims that evidentiality is ‘a grammatical phenomenon known to diffuse via language contact.’ Citing Aikhenvald (2004), she notes that evidentiality is a property of several well-established linguistic areas. Sinhala shares the two kinds of indirect evidential constructions with Tamil. Sinhala may have replicated the two inferential evidential constructions (22,23) from Tamil in that both share almost the same constructions. It is important to note that both use the grammaticalized form of the postposition ‘like’ (*wage* in Sinhala and *poola* in Tamil) as the inferential marker (22)—another instance of replica grammaticalization.

The quotative evidential construction in Tamil and Sinhala (24) is distinctly of the Dravidian kind which indicates that Sinhala has replicated the pattern of the Tamil evidential construction. It may be noted that NIA languages like Hindi and Bengali have no distinct grammatical evidential. The quotative clitic is added to the finite form of the verb, while in negative quotatives, the negative particle precedes the quotative clitic. Moreover, as in the constituent polar (2.3.1.6) and constituent negative (3.6.3) constructions, the quotative clitic which occurs adjacent to a distinct constituent indicates that the constituent is being quoted, lending a focus reading, as in a constituent quotative construction (53) (based on (24)):

(53) T. ungal-udaia kaarr-ai(-**aam)** avan(-**aam)** vangu-v-aan
Your(HON)-GEN car-ACC-QUOT he-QUOT buy-FUT-3SGM
S. oyaa-ge kaar ekə(-**lu)** eyaa(-**lu)** gann-e
Your-GEN car-one-QUOT he-QUOT buy.FUT-NMLZ

‘(It is said that) it is your car that he would buy.’
‘(It is said that) it is he that would buy your car.’

The structure of the quotative evidential construction, especially the positions at which the quotative clitic occurs, lends further support to the idea that the Sinhala quotative evidential construction is a calque of the parallel Tamil construction.

The morphosyntactic features used to convey epistemic and event modalities are also the same in the two languages. The epistemic/event modal particles/clitics always follow the verb in the two languages. These modals also have distinct negative forms (see fn. 63). Mention must be made of the fact that in both languages, the word *venum* in Tamil and *oone* in Sinhala is used as a verb ‘want’ (4.2.3; see also 4.3.11) and as an auxiliary ‘must/should/need’ (4.3.4). The former requires a dative subject, as in (54), while the latter requires a nominative subject (25) in 4.3.4):

(54) T. shanthi-kku nalla padikk-a **venum**
Shanthi-DAT well study-INF should
S. shanthi-tə hondətə igenəgandə **oone**
Shanthi-DAT well study.INF should

‘Shanthi wants to study well.’

Note that in the two languages (53) means that ‘Shanthi has the desire to study well’, whereas (25) in 4.3.4) means that ‘Shanthi should study hard (because she has not done her exam well)’. Except for the dative case on the subject in (53) and the nominative case on the subject in ((25) in 4.3.4), both sentences share the same construction, but they convey different meanings.

Data for the two kinds of coordination, conjunctive and disjunctive (4.3.5) amply demonstrate further similarities between Sinhala and Tamil.[69] (i)–(iii) below show the features related to coordination which Sinhala may have replicated modelled on the corresponding Tamil features:

---

[69] Tamil and Sinhala use adversative coordination too and have distinct adversative coordinating particles *aanaal* and *namuth* respectively.

(i) the two languages use bisyndectic conjunctive coordination, that is, the conjunctive/disjunctive coordinator is added to both of the conjuncts (31). This bisyndectic coordination is used in Dravidian, whereas other NIA languages, like Bengali use monosyndectic coordination.

(ii) Sinhala has adopted the form-meaning unit of the disjunctive coordinators used in the two kinds of disjunctive coordination:

a) Sinhala may have adopted the disjunctive coordinator-*ho* from Tamil *-oo* (34a). In Dravidian, *-oo* is used both as the correlative clitic (Subbarao 2010) and the disjunctive clitic (Jayaseelan 2001b for Malayalam and Amritavalli 2003 for Kannada). The correlative clitic in Dravidian may have been extended to be used as the disjunctive clitic or vice versa.

b) the other bisyndectic disjunctive coordinator *sari* in Tamil, *hari* in Sinhala has been grammaticalized from the lexical adjective *sari/hari* 'correct'/'right' (34b); the latter, which is nearly homophonous with the former, believed to have been calqued from Tamil. The lexical adjectives *sari* and *hari* are also used to elicit confirmation, affirmation or approval from the interlocutor, as in (55):

(55) T. sari-aa
correct-Q
S. hari-də
correct-Q

Literal meaning 'Is it correct?'
Idiomatic meaning 'Do you accept/agree?

c) the monosyndectic disjunctive coordinator in Tamil *illai-endaal* consists of the negative particle *illai* and the complementizer prefix *endu* combined with the conditional suffix *-aal* (akin to the topic particle *endaal* in (11bT)), while *nœ-(th)naŋ* in Sinhala consists of the negative particle *nææ* and the conditional suffix *-naŋ* (akin to the topic particle *-naŋ* in (11bS)), conveying the meaning 'if not' (35), which again exemplifies replica grammaticalization.

d) the disjunctive coordinator used for exclusive identification of one constituent in Tamil, *ond-il*, consists of *ondu* 'one' and the locative case suffix *-il*, while *ekk-oo* in Sinhala consists of *ekə* 'one' and the clitic *-oo* (36). Sinhala may have extended the existential quantifier clitic

-oo (see 4.3.7) which is the Tamil/Dravidian correlative/disjunctive clitic *-oo* (see (ii.a) above) appropriated by Sinhala.

(iii) the structures of (31)–(36) which are almost mutually intertranslatable indicate that Sinhala may have replicated the pattern of these constructions.

Alternate questions in Sinhala too may have been modelled on those of Tamil (4.3.6) in that the question clitic/particle in Tamil/Sinhala occurs adjacent to the constituents encoding the options, just as in the bisyndectic disjunctive coordinating construction in the two languages. As shown in 2.3.1.5, Sinhala has modelled the question particle from the Tamil/Dravidian correlative clitic to be used in alternate questions first. It has been extended to polar and wh-questions later. What is more important is that (37T,S) are mutually intertranslatable morpheme by morpheme. These facts provide evidence for the replication of alternate questions by Sinhala from Tamil.

It is reasonably evident that Sinhala has also adopted the form-meaning unit of the two kinds of quantifiers from Tamil (see 4.3.7). Sinhala, like Tamil, forms the universal quantifiers by combining wh-phrases with the conjunctive clitic *-th* (see also indefinites given in table 3.9) and the existential quantifier by combining wh-phrases with clitic *-oo* which is a slight variant of the disjunctive clitic *-ho.* As mentioned above (ii.d), Sinhala appropriated the Tamil/Dravidian correlative/disjunctive clitic *-oo* as its existential quantifier clitic. In Malayalam too the two kinds of quantifiers are formed in exactly the same way as those in Tamil (see Jayaseelan (2001b).[70] Lahiri (1998) notes that in Hindi 'negative polarity items' are formed by adding the 'emphatic marker' *bhii* to the corresponding simple existentials. It could not be verified whether other NIA languages too have similar formation of the two kinds of quantifiers. However, in Tamil and Sinhala, the constituents and the structure of the two kinds of quantifiers are the same. Note also the almost identical constituents and the structure of quantifiers in Malayalam. Thus the two kinds of quantifiers also demonstrate replica grammaticalization in that these two kinds in Sinhala have been exactly modelled on those in Tamil, that is, the conjunctive and disjunctive clitics are added to wh-words/phrases to form universal and existential quantifiers respectively.

It is significant to note that both languages have a single clitic (*-um* in Tamil and *-th* in Sinhala) which is used in three different morphosyntactic functions: (i) as the concessional suffix (2.3.1.7); (ii) as the conjunctive coordinator (4.3.5); and (iii) as the

---

[70] Tamil and Malayalam use question words and the same conjunctive *-um*/disjunctive *-oo* clitics in the formation of quantifiers.

clitic used to form the wh-based indefinite pronouns and the generic pronoun based indefinite pronoun (3.6.3) and universal quantifiers (4.3.7). The clitics which the other Dravidian languages use are homophonous with the clitic in Tamil. It is, therefore, likely that the origin of the clitic is Tamil/Dravidian, and Sinhala may have initially formed the clitic to perform one of the functions (i–iii) and subsequently extended the same clitic to perform the other functions, as has Tamil. It is difficult to determine in which of the three morphosyntactic phenomena of Tamil, the clitic originated, but what is possible to assume is that once the clitic became consolidated in one phenomenon, it may have been extended to the other two phenomena. Here too, as in many instances above, the replication is at the level of clitic, as expected in intense language contact situations.

A further similarity between the two languages is the existence of the two kinds of verbal nominals (4.3.8), namely tensed and untensed ones, together with the almost similar ways in which they are formed in the two languages. With respect to the tensed verbal noun forms in (1) in table 4.3, it is useful to note that the verbal adjectival to which the nominalizer suffix is added is in effect the non-finite (participle) form of the verb used in the relative clauses of the two languages. Both languages use the tensed nominalized form of the verb in their cleft constructions (6.1). Tamil uses the same tensed verbal nominals given in (1) in table 4.3 in its cleft constructions, whereas Sinhala uses a somewhat reduced form of its tensed verbal nominals as its nominalized verb form, e.g., *dunne* ‘give’, *kææwe* ‘eat’ etc. (see 6.3 for a detailed account of the formation of these Sinhala nominalized verb form used in the cleft construction). As for the untensed forms, there is no basis for the distribution of the nominalizer suffixes -(*th*)*al* or *kai* in Tamil and *-iimə or -illə/-um* in Sinhala. In other words, there are no rules (see Lehmann 1989 and Chandralal 2010) stipulating to which class of verbs—especially the three kinds of verbs in Sinhala classified on the basis of stem vowel (3.6.1.2)—these suffixes are added. Sometimes both suffixes can be added to the same verb to produce the two kinds of nouns (4.3.8), e.g., *sei-thal* ‘doing’ and *sei-kai* ‘action’ in Tamil and *næt-iimə/illə* ‘dancing’ *and næt-um* ‘dance’ in Sinhala. The similarities between the formation of the tensed and untensed verbal nominals also show that the pattern of the two kinds of verbal nominals have been adopted by Sinhala from Tamil.

Of the 13 features in 4.3, the only feature that is different between the two languages is the volitive/involitive distinction between Sinhala verbs as compared with the affective/effective distinction in Tamil verbs (4.3.9). Though different terms are used to refer to these two systems in the literature, the analysis of these two kinds of verbs suggests that they perform similar semantic functions. However, the affective/effective distinction in Tamil is found only in a limited number of Tamil verbs given in table 4.5, unlike the volitive/involitive distinction which is found in the majority of Sinhala verbs. How this distinction evolved in Sinhala verbs and whether such distinction exists in the verbs of other NIA languages cannot be verified. It is most likely that Sinhala may have inherited this feature as an NIA language. It may have also been an endogenous development taken place in Sinhala. Note that in the two kinds, one kind may have evolved from the other because of the subtle morphophonological differences between the two: if the involitive forms are assumed to be underlying, then the front vowel in the first two syllables is retracted in the volitive forms, e.g., [i] to [u] in *issenəwa* ‘to rise’ and *ussənəwa* ‘to lift’, in which the vowel height is maintained, whereas if the volitive verb forms are assumed to be underlying, a kind of vowel fronting occurs to yield the involitives, e.g., the vowel fronting from [ɑ] to [æ] in *arinəwa* ‘to open’ and *ærenəwa* ‘to open automatically’ (pc with Dr. S. J. Hannahs). It is also significant to note that Tamil does not have constructions, like (38bS), in which the dative subject co-occurs with involitive verbs. This construction in Sinhala, like (37aS) in 3.6.1.1—which is different from the dative subject constructions in 4.3.11—may have been a feature remnant of the split ergative-absolutive system characteristic of NIA languages which Sinhala is believed to have had before it acquired the nominative-accusative system.

As discussed above, the clause final complementizer (4.3.10) in utterance complement clauses (4.2.7) is widely accepted as an areal feature of the South-Asian languages that has been borrowed by Indo-Aryan from Dravidian (Biberauer et al. 2009 and the references therein). Biberauer et al. (2009) further claim that all modern Indic languages (with the exception of Sinhala) also have an initial COMP, the origins of which are moot. Thus, like the Sinhala question particle which was changed from the clause-initial position (Sanskrit/Pali) to the clause-final position (Dravidian) (2.3.1.5), the Sinhala *say*-complementizer may also have been changed from clause-initial to clause-final position under the influence of the parallel construction in Tamil.

A more structural approach to the facts is proposed by Biberauer et al. (2009), who discuss the effect of the Final-Over-Final Constraint (FOFC) on borrowing in

South Asian languages which have long been in contact.[71] On the basis of their finding that COMP is uniformly higher than the Polar head (Q), they claim that if Q co-occurs with COMP in an embedded clause, there are only three options with regard to the position of Q and COMP in the languages of South Asia: (i) final Q precedes final COMP; (ii) initial Q follows initial COMP; and iii) COMP is initial and Q is final. The fourth option of COMP being final and Q being initial is rare and it is not attested in the languages of the South Asian region. The reason is that an initial Q blocks the borrowing/development of a final COMP as this would lead to a FOFC violation (Biberauer et al. 2009). Biberauer et al. 2009 observe that Sinhala is the only NIA language which has a clause final/medial (not initial) Q and final only COMP (option (i) above), along with Dravidian, as shown in (42). Table 4.9 (figure 2, p. 14, ibid.) shows the patterns of the positions of polar particles (Pol) and complementizer (C) in the languages of South Asia:

Type	**Position of Pol**	**Position of C**	**Languages**
A	Initial	Initial Only	Hindi-Urdu, Panjabi, Kashmiri, Sindhi, Maithili, Kurmali
B	Final/Medial	Initial and Final	Marathi, Gujarati, Assamese, Bangla, Dakhini Hindi, Oriya, Nepali (plus some North Dravidian languages, i.e. Brahui)
C	Final/Medial	Final Only	Sinhala (plus most Dravidian languages)
D	**Initial**	**Final**	**Unattested in the area**

Table 4.9: Position of polarity heads and complementizers in South Asian languages

In the two languages, the default position of Q is clause-final; only in constituent polar questions does Q occur medially, following the constituent which is questioned ((31a/b) in 2.3.1.6), while COMP is always clause-final. In the two languages, Q always precedes COMP in default complex sentences with embedded polar questions (42). If

---

[71] Bieberauer et al. (2009: 5) note that FOFC is ‘an absolute principle which acts as a universal constraint on synchronic grammars. Based on Holmberg’s (2003) proposal, they postulate FOFC as follows: If α is a head-initial phrase and β is a categorically non-distinct phrase immediately dominating α, then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α, then β can be head-initial or head-final. The implications of FOFC are not discussed here owing to constraints of space (see Biberauer et al. 2009 for discussion).

Biberauer et al.'s (2009) findings, especially with regard to the positions of Q and COMP in South Asian languages, are correct, the similarity between the correlation of Q with COMP in Dravidian and Sinhala indicates that the latter may have undergone contact-induced restructuring from option (ii) or (iii) to (i)[72]. Contact-induced change, as Biberauer et al. (2009) rightly point out, is of particular interest in the present context as it has sometimes been suggested that this type of change, in contrast to the non-contact-induced variety, may result in typologically unusual linguistic systems.

So far, as has been evident, Tamil and Sinhala share with other South-Asian languages all the morphosyntactic features of the region, including SOV word order (2.2.1), morphological causatives (3.3.6) and compound verbs (4.3.3).[73] The two languages also share other accepted areal features discussed in 4.3.11–4.3.13. Most of these are thought to be Dravidian in origin, having diffused from Dravidian into the NIA languages (Schiffman 2010, Emeneau 1956, Ebert 2010 and Thomason 2000). As noted above, Dravidian is also the likely source of the *say*-complementizer (Biberauer et al. 2009) and compound verbs (Herring 1989).

Ebert (2010: 997) notes that Dravidian 'is usually considered a likely source' of the dative constructions (4.3.11), while Schiffman (2010) identifies it as one of the phenomena which was borrowed intensively, and has had cross-family influence. Examples (39/40) in 4.3.10 show that the dative constructions in the two languages are isomorphic. Furthermore, in the two languages, the same kinds of verbs license dative subjects. It is, therefore, likely that Sinhala has adopted the dative verb construction from Tamil. Like the other areal features, it is difficult to determine when and where Sinhala acquired dative constructions.

Emeneau (cited in Ebert 2010) considers reduplication (4.3.12) 'to be borrowed from Dravidian...' The reduplication is attested in the NIA languages too. As has been shown, a wide range of items from verbs to reciprocal pronouns that involve reduplication are similar in the two languages (see table 4.6). Also similar are the semantic functions these reduplicated items perform. It is, therefore, plausible to conclude that the Sinhala reduplicated items may have been modelled on the corresponding Tamil ones.

---

[72] Biberauer et al. (2009) take the same stand in their work.

[73] The retroflex consonants, distinguished as an areal feature are not discussed in this dissertation. According to Ebert (2010: 995) the increasing frequency of retroflex consonants in the successive books of the *Rigveda* 'Vedic Sanskrit hymns' is generally accepted as proof of Dravidian substratum influence.'

The source of converbs or conjunctive participles (4.3.13)—whether NIA or Dravidian— is not known. What is important, however, is that converb constructions like (45)–(48) which have the same structure cannot evolve in the two languages of diverse origins owing to coincidence or endogenous changes. Note that in (45)–(48), the non-finite verb of the subordinate clause precedes the finite verb of the main clause. Though converbs are a distinct areal feature, the view that Sinhala may have replicated them either directly from Tamil or indirectly from another NIA language which has replicated them from Dravidian cannot totally be ruled out.

4.5 Conclusion

Further to the morphosyntactic features analyzed in Chapters 2 and 3, the 20 features analyzed in this chapter clearly show that these features in Sinhala are remarkably similar to the corresponding ones in Tamil, possibly due to the contact-induced restructuring of the former on the model of the latter. It is true that the values of all the seven features related to complex sentences compared in 4.2 are the same not only in Tamil and Sinhala, but also in most of the languages of the region (see Appendix). However, the structures of all these features in the two languages are isomorphic, and the other properties related to the seven features discussed in 4.4.1, for example, the diverse uses of the complementizer in the two languages, are distinctively Dravidian. It is, therefore, plausible to conclude that these features in Sinhala have undergone contact-induced morphosyntactic changes.

The 13 miscellaneous features compared in 4.3 are also strikingly similar, and close examination of them shows that many features in Sinhala have been modelled on those of Tamil. Features like topicalization, some grammaticalized auxiliaries, some auxiliaries conveying modality, the two kinds of coordination, alternate questions and the formation of the two kinds of quantifiers are noticeably similar in the two languages. Even the features which are considered to be areal exhibit remarkable resemblance in the two languages when it comes to fine details of structure and formation. Overall, therefore, the features examined in this chapter further support the claim that contact between the two languages has induced Sinhala to replicate fully or partially various features from Tamil. Changes in Sinhala induced by its contact with Tamil, as the analysis shows, are evidenced in almost all the areas of Sinhala morphosyntax. It is difficult to determine which language family Sinhala belongs to purely based on the analysis of its morphosyntax as it has diverged from the other NIA languages, which makes Gair (1998a) call Sinhala ‘an Indo-Aryan isolate’.

This chapter brings to completion the macro-level analysis of the morphosyntactic features in the two languages. The features analyzed in Chapters 2–4 show extensive convergence between Sinhala and Tamil. These changes cannot plausibly be ascribed to chance or to internally motivated causes either in Sinhala or Tamil or both. Instead, they point to the contact between these two languages of diverse origins as the factor that accounts for this convergence. Having established this, in Chapters 5 and 6, two phenomena in the two languages are analyzed at micro-level. Chapter 5 zooms in on one specific feature that Sinhala and Tamil have in common, i.e., the occurrence of null arguments. It was mentioned above that Sinhala and Tamil both allow null arguments. However, there is a much fuller story to be told about this feature and Chapter 5 examines this topic in detail.

Chapter 5

Null arguments

5.1 Preliminaries

Within generative linguistics, ‘pro drop’ or ‘null arguments’ is an important topic in that it has been regarded as a prime example of a parameter of UG. Widely known as the null subject parameter, this together with other parameters such as the head-complement parameter and the polysynthesis parameter (Baker 1996) accounts for cross-linguistic variation in the model of Principles and Parameters Theory proposed under the GB framework (see 1.4). The null subject parameter is taken to account for language variation with regard to the fact that, as Roberts and Holmberg (2010: 5) point out, ‘some languages allow a definite pronominal subject of a finite clause to remain unexpressed as a nominal bearing the subject function, while others do not.’ The former would be Null Subject Languages (NSL), e.g., Spanish, Italian, whereas the latter would be non-NSLs, e.g., English, Swedish. This binary distinction does not mean that the two ways of expressing subject pronouns are mutually exclusive; note that the NSLs typically sometimes use overt pronominal subjects, while the non-NSLs, e.g., English require an obligatory null subject (PRO) in non-finite clauses. Arguably all languages allow both overt and covert (null) subjects, but what is important is to establish in what contexts languages allow these two kinds of subjects.

Tamil and Sinhala rather freely allow null subjects, as in (1) below (brackets indicate that the pronouns are optionally null):

(1) T. (avan) nalla hindi kathai-kir-aan
He well Hindi speak-PRS-3SGM
S. (eyaa) hondətə hindi kathakərənəwa
He well Hindi speak.PRS

‘He speaks Hindi well.’

The important question with regard to null subjects is what licenses null subjects in some languages but not in others. The traditional conception of the null subject parameter (Rizzi 1986) has been that null subjects are possible because of rich agreement. A distinctive feature of classical NSLs is inflectional richness, specifically SV agreement. It is important to note, however, that SV agreement is the most salient morphosyntactic difference between Tamil and Sinhala, as evidenced in the values of three features (3.2.4–3.2.6; see table 3.1); Tamil has rich SV agreement, while Sinhala does not have agreement. Given the difference in patterns of agreement between the two languages, it would be expected that they do not display the same behaviour and that the presence or absence of SV agreement should be reflected in different contexts where null pronouns occur. Note, however, that (1) above indicates that both Tamil and Sinhala behave the same, i.e., allow optional null subjects in at least some contexts. What needs to be investigated, therefore, is whether the two languages display the same behaviour with respect to licensing null subjects and null non-subjects in other potential contexts as well; whether languages allow null arguments other than subjects is another aspect of the pro-drop parameter. This chapter analyzes the occurrence of null subjects/null non-subjects in these two languages in detail to determine what kind of NSLs they are and to examine the circumstances in which null arguments occur and the factors that account for the occurrence of null arguments in the two languages. It is expected that the results obtained in this chapter will help assess the possible implications for the idea that there has been contact-induced restructuring of Sinhala.

In this chapter it will be argued that Tamil and Sinhala can be classified as discourse pro drop languages which allow null subjects/non-subjects under specific discourse conditions. In Section 5.2, the tests for null subjects used in Cole (2010) are applied to Tamil and Sinhala to study the intra-/inter-language differences between the occurrence of null subjects. Section 5.3 focuses on the null subject parameter within the minimalist framework and describes briefly the different kinds of NSLs distinguished in the literature with the view to determining what kind of null subject languages Tamil and Sinhala are. This section also provides a summary of the major theoretical frameworks within which the occurrence of null arguments in NSLs are analyzed. In Section 5.4, diverse works in the literature are used to analyze the contextual conditions under which the two languages allow null subjects/non-subjects. Based on the findings and also the results obtained in 5.2, it concludes that Tamil and Sinhala are discourse pro drop languages, providing further confirmation that they behave the same with respect to allowing null arguments despite the difference between them in SV agreement. Section 5.5 summarizes the findings and discusses the problems with the existing taxonomy of null argument languages.

5.2 Tests for thematic null subjects

Cole (2009) and (2010) may be profitably used for an initial survey to determine the circumstances in which null subjects occur in these two languages. A traditional problem for the classical null subject parameter is that there are languages that allow null subjects, and also null objects, but do not have any agreement, for example, Chinese and Japanese (Huang 1984, Rizzi 1986). In these languages, the interpretation of the null pronoun is entirely a matter of recovering the reference of the pronoun from the context.

Cole (2009) and (2010: 280) argues that the fundamental feature distinguishing NSLs from non-NSLs 'is not rich subject verb agreement, [but] the possibility of recovering thematic null subjects from context when recovery by agreement up to the point of morphological maximality fails to isolate them.' The failure to recover thematic null subjects from agreement may be due to the absence of verbal morphology, as in Chinese, or the agreement available being syncretic, as in Spanish (in some cases). Cole (2010), following the findings in Cole (2009), attempts to establish within a suitable framework the contextual strength and weakness which, according to him, afford the distinction between languages that have thematic null subjects and those which do not. He identifies the accessibility theory developed in the works of Ariel (2001) as the theory that best fits the notion of contextual strength. [74]

A survey was conducted by Cole (2010) to account for the intra-/inter-language differences with respect to the occurrence of thematic null subjects. The intra-language differences refer to the particular circumstances in which languages allow thematic null subjects, whereas the inter-language differences refer to the cross-linguistic variation in the licensing of thematic null subjects. To determine the intra-language differences between the occurrences of thematic null subjects in languages, Cole (2010) tests a set of sentences with dislocated/hanging topics or sentence pairs: in the former the main sentence contains a null subject and in the latter the second sentence contains it. The dislocated topic or the first sentence contains an antecedent for the null subject, but the structural position/function of the antecedent varies in these sentences (see below). To describe the inter-language differences in the occurrence of thematic null subjects in NSLs and the difference in contextual strength, he selected eight languages, five of which have rich SV agreement, namely Greek, Italian, Serbian, Spanish and Icelandic, while three of them, Japanese, Chinese and Swedish lack SV agreement. In the survey, the native speakers of the languages investigated were asked to indicate how the third person singular pronoun with an antecedent in a set of sentences is expressed—whether 'Overt', 'Optional' or 'Null. Note that the antecedents are lexical NPs and, as mentioned above, their syntactic function and syntactic relation to the corresponding pronouns vary.

---

[74] See references in Cole (2010).

Types A–Iii. in which the antecedents occur with the third person pronouns used in Cole's (2010: 285–286) survey are used here to explain the intra-language differences in the occurrence of thematic null subjects in Tamil and Sinhala. The results are then compared with those from the other languages available in Cole (2010) to determine what kind of null subject languages Tamil and Sinhala are. Types A–Iii. which correspond to the ten instances ((2)–(10ii) below) are divided into three categories: (i) instances where the antecedent of the relevant pronominal element is a topic (examples (2), (3), (4), (5) and (10i)); (ii) instances where the antecedent of the relevant pronominal is a non-topic argument of a verb (example (6)); and (iii) instances where the antecedent of the relevant pronominal is a non-argument of a verb (examples (7), (8), (9) and (10ii)). Cole (2010: 284) notes that '[e]ach of these [three categories] could be seen as being of different salience and accordingly of different accessibility.'

Tamil, like the five languages in Cole's survey, has rich agreement and Sinhala, like the three languages in Cole's survey, lacks agreement. I use my native intuition in Tamil and my knowledge of Sinhala to formulate the Tamil and Sinhala equivalents given below for each English form. Note that a) the critical pronominal elements, showing intended interpretations (which are not meant to imply focus), are in italics; b) the intended interpretations are subscripted and c) null subjects are indicated by Ø, while optional null subjects are in brackets.

(2) two single clause sentences with the subject of the second sentence co-referent with a topic antecedent in the first.
'Every morning, the minister$_1$ visits the museum. In the evening *he*$_1$ visits the university.'

T. ov-oru naal kaalamai-um manthiri$_1$ noothanasalai-kku varukai
One-one(INDF) day morning-INCL minister museum-DAT visit
tha-(ruk)ir-aar pinnerath-il (*avar*$_1$) palhalikkalaha-(thu)kku
give-PRS-3SGM(HON) eveining-LOC he(HON) university-DAT
varukai tha-(ruk)ir-aar
visit give-PRS-3SGM(HON)
S. hamədamə udee-tə manthrithuma$_1$ kauthukaagaarəyə balannə
Everyday morning-DAT minister(HON) museum look.INF
yanəwa hawəsə-tə (*eyaa*$_1$) vishvəvidyaaləyə balannə yanəwa
go.PRS evening-DAT he university look.INF go.PRS

(3) two single clause sentences with the subject of the second co-referent with the topic of the first, which has a possessor.

‘John$_1$’s mother$_2$ is frightful. *She*$_2$ hates him$_1$.’ [75,76]

T. john$_1$-inda amma$_2$ payankaram-aana-vaa (*avaa*$_2$) avan$_1$-ai
John-GEN mother frightful-be.PRS-3SGF(HON) (she(HON)) he-ACC
veru-kkir-aa
hate-PRS-3SGF(HON)

S. john$_1$-ge amma$_2$ bhayanəkai (*eyaa*$_2$) eyaa$_1$-tə vairəyə kərənəwa
John-GEN mother frightful she he-DAT hate.PRS

(4) a main clause containing a subordinate clause and beginning with a topic, with the subject of the subordinate clause co-referent with the topic.[77]

‘John$_1$’s mother$_2$, we know the reason why *she*$_2$ criticised him$_1$.’

T. john$_1$-inda amma$_2$, engal-ukku therium *avaa*$_2$ enna
John-GEN mother we-DAT know.PRS.3SGN she(HON) what
kaarana-(thu)kk-aha avan$_1$-ai vimarsi-th-aa endu
reason-DAT-ADV he-ACC criticize-PST-3SGF(HON) COMP

S. john$_1$-ge amma$_2$ api dannəwa monə heethuwə udə *eyaa*$_2$
John-GEN mother we know what reason on she
eyaa$_1$-wə vivechənəyəkər-e kiyəla
he-ACC criticize.PST-NMLZ COMP

(5) a one clause sentence with its subject co-referent with a left dislocated topic, which has a possessor.

‘John$_1$’s mother$_2$, *she*$_2$ hates him$_1$.’

T. john$_1$-inda amma$_2$ *avaa*$_2$ avan$_1$-ai veru-kkir-aa
John-GEN mother she(HON) he-ACC hate-PRS-3SGF(HON)

S. john$_1$-ge amma$_2$ *eyaa*$_2$ eyaa$_1$-tə vairəyəkərənəwa
John-GEN mother she he-ACC hate.PRS

---

[75] In Type B ((3) above), Cole (2010) incorrectly subscripted ‘him’ (John) as ‘2’ which refers to ‘She’ (‘John’s mother’). If it were ‘2’, the possessor ‘John’ which has subscript 2 would not be the antecedent of ‘him’ (John). In Type D ((5) below) ‘him’ (John) is correctly subscripted.

[76] It is important to note that in Tamil and Sinhala, in instances like (3), (4) and (5) the pronoun ‘he’ is replaced by ‘John’ as ‘John’ is not the topic, but the possessor of the topic.

[77] (4) and (5) which involve dislocated/hanging topics characteristic of Romance languages are not common in Tamil and Sinhala, but are not ungrammatical. This kind of topicalization is different from the topicalization/L- movement (see 4.3.2) in Tamil and Sinhala in which the topic is moved from its base position to the topic position.

(6) two single clause sentences with the subject of the second co-referent with the agent of the passive in the first.[78]
‘Every morning, the museum is visited by the minister. In the evening, *he*$_1$ visits the university.’

T. ov-oru naal kaalamai-um noothanasalai manthiri$_1$-aal varukai
One-one(INDF) day morning-INCL museum minister-INS visit
thar-a-padu-kir-athu pinnerath-il *avar*$_1$ palhalikkalaha-(thu)kku
give-INF-PASS-PRS-3SGN eveining-LOC he(HON) university-DAT
varukai tharu-kir-aar
visit give-PRS-3SGM(HON)
S. hamədamə udee-tə manthrithuma$_1$ visin kauthukaagaarəyə
Everyday morning-DAT minister(HON) by museum
balənu labənəwa hawəsə-tə *ohu*$_1$ vishwəvidyaaləyə
look.PTCP get.PRS evening-DAT he university
balannə yanəwa
look.INF go.PRS

(7) two single clause sentences with (i) the subject of the second sentence co-referent with the possessor of the topic in the first and (ii) an object pronoun in the second sentence co-referent with the topic of the first and capable of differentiating between that topic and its possessor.
‘John$_1$’s mother$_2$ is terrible, *H*e$_1$ hates her$_2$.’

T. john$_1$-inda amma$_2$ payankaram-aana-vaa *avan*$_1$ avaa$_2$-(v)ai
John-GEN mother frightful-be-PRS-3SGF(HON) he she(HON)-ACC
veru-kkir-aan
hate-PRS-3SGM
S. john$_1$-ge amma$_2$ bhayaanəkai, *eyaa*$_1$ eyaa$_2$-tə vairəyəkərənəwa
John-GEN mother frightful he she-DAT hate.PRS

(8) as (4) but with (i) the subject of the subordinate clause co-referent with the possessor of the topic and (ii) an object pronoun in the subordinate clause co-referent with the topic and capable of differentiating between that topic and its possessor.
‘John$_1$’s mother$_2$, we know the reason why *he*$_1$ hates her$_2$.’

T. john$_1$-inda amma$_2$, engal-ukku therium *avan*$_1$ enna
John-GEN mother we-DAT know.PRS.3SGN he what
kaarana-(thu)kk-aha avaa$_2$-(v)ai veru-kkir-aan endu
reason-DAT-ADV she(HON)-ACC hate-PRS-3SGM COMP
S. john$_1$-ge amma$_2$ api dannəwa monə heethuwə udə *eyaa*$_1$
John-GEN mother we know.PRS what reason on he
eyaa$_2$-tə vairəyəkərann-e kiyəla
she-DAT hate.PRS-NMLZ COMP

---

[78] The verb constructions in the passive clauses sound more formal because the passive voice is rarely used in speech in the two languages.

(9) a one clause sentence with (i) the subject co-referent with the possessor of a left dislocated topic and (ii) an object pronoun co-referent with the topic and capable of differentiating between that topic and its possessor.
‘John$_1$’s mother$_2$, *he*$_1$ hates her$_2$.’

T. john$_1$-inda amma$_2$ *avan*$_1$ avaa$_2$-(v)ai veru-kkir-aan
John-GEN mother he she(HON)-ACC hate-PRS-3SGM
S. john$_1$-ge amma$_2$ *eyaa*$_1$ eyaa$_2$-tə vairəyəkərənəwa
John-GEN mother he she-DAT hate.PRS

(10) a question, with a subject/topic that has a possessor, and two replies. The first reply has a subject co-referent with the subject/topic of the question. The second reply has a subject co-referent with the possessor of that topic and an object pronoun co-referent with that topic and capable of differentiating between that topic and its possessor.
Question: Has John$_1$’s mother$_2$ left?
Answer i) Yes, *she*$_2$ has left.
Answer ii) No, *he*$_1$ has invited her$_2$ to dinner.

T. john$_1$-inda amma$_2$ poo-it-t-aa-(v)aa
John-GEN mother go.PTCP-leave-PST-3SGF(HON)-Q
i. oom Ø$_2$ poo-it-t-aa
Yes she go.PTCP-leave-PST-3SF(HON)
ii. illai avan$_1$ raa saapaadu-kku avaa$_2$-(v)ai
No he night meal-DAT she(HON)-ACC
koopitu-(v)it-t-aan
call.PTCP-leave-PST-3SGM

S. john$_1$-ge amma$_2$ pitath una də
John-GEN mother leave happen.PST Q
i. ou Ø$_2$ pitath una
Yes she leave happen.PST
ii. nææ *eyaa*$_1$ eyaa$_2$-tə rææ kææmə-tə kathakərəla thiyenəwa
No he she-DAT night meal-DAT speak.PTCP have.PRS

The results of examples (2)–(10i, ii) in Tamil and Sinhala are summarized in table 5.1 below (Null= Null; Opt = optionally covert pronoun; and Ov’t = overt).

The results obtained from Tamil and Sinhala are included in—table 8, Sections A and B in Cole (2010: 298) given as—tables 5.1 and 5.2 respectively (see last rows for the results of Tamil and Sinhala):

Ante-cedent status	Topic					Non topic argu-ment	Non argument			
Example	(2)	(5)	(10) i	(3)	(4)	(6)	(8)	(10) ii	(7)	(9)
Greek	Null	Null	Null	Null	Null	Opt	Opt	Opt	Opt	Opt
Spanish	Null	Null	Null	Null	Null	Opt	Opt	Opt	Opt	Opt
Serbian	Null	Null	Null	Null	Null	Opt	Opt	Opt	Ov't/ Opt	Ov't
Italian	Null	Null	Null	Null	Null	Ov't	Ov't	Ov't	Ov't	Ov't
Icelandic	Ov't	Ov't	Ov't	Ov't	Ov't	Ov't	Ov't	Ov't	Ov't	Ov't
Tamil	Opt	Ov't	Null	Opt	Ov't	Ov't	Ov't	Ov't	Ov't	Ov't

Table 5.1: Rich agreement languages

Ante-cedent status	Topic					Non topic argu-ment	Non argument			
Example	(2)	(5)	(10) i	(3)	(4)	(6)	(8)	(10) ii	(7)	(9)
Japanese	Null	Null	Null	Null	Null	Opt	Ov't	Ov't	Ov't	Ov't
Chinese	Opt	Opt	Opt	Ov't	No result	Ov't	No result	Ov't	Ov't	Ov't
Swedish	Ov't	Ov't	Ov't	Ov't	No result	Ov't	No result	Ov't	No result	Ov't
Sinhala	Opt	Ov't	Null	Opt	Ov't	Ov't	Ov't	Ov't	Ov't	Ov't

Table 5.2: Languages lacking verbal morphology for person or number

Following Cole (2010), the results are given in tables 5.1 and 5.2 in decreasing accessibility from left to right: a topic antecedent is more accessible than a non-topic one, which is still more accessible than a non-argument antecedent, Cole's term for a possessor argument. As a consequence, the requirement for overt pronoun also increases, which holds true for Tamil and Sinhala. As for intra-language differences in the occurrence of thematic null subjects, the results of the survey show that Tamil and Sinhala display identical behaviour with regard to the occurrence of null subjects despite their difference in SV agreement. They require a null subject pronoun in the answer if the antecedent is the subject of a preceding yes-no question (10i), and allow optional null subject if the antecedent is the topic of the matrix clause or a preceding sentence ((2) and (3)). They require overt pronouns if the antecedent is a dislocated topic (5) and (4) or a non-topic argument of the verb (6) or a non-argument of the verb (8), (10ii), (7) and (9).

With regard to inter-language differences in the occurrence of thematic null subjects, the results of Tamil show that it behaves differently from any of the other rich agreement null subject languages; broadly speaking it comes between the other null subject languages and the non-null subject language Icelandic. As such, it may be classified as a partial null subject language, but as will be seen in 5.3 below, it does not conform to the definition of partial null subject languages employed in recent literature either. Compared to the other NSLs in table 5.1, Tamil requires a null pronoun only in one instance and has optional or overt pronouns in others. Therefore, it looks like agreement does not play the same role in Tamil as it does in the other rich agreement NSLs. Sinhala, on the other hand, can be compared to Chinese (table 5.1). Sinhala lacks agreement, and hence, it is the contextual strength, that is, the capacity to recover antecedents in context, which facilitates the occurrence of null arguments in it. Sinhala seems to have moderate contextual strength like Chinese, compared with Japanese which has more contextual strength.

But most strikingly, Tamil and Sinhala have exactly the same values. The results show that of the six NSLs in Cole (2010), Chinese is closer to the two languages under investigation than any other. Though Tamil and Sinhala are crucially different in terms of SV agreement which is supposed to license null subjects, they show identical results in allowing null subjects, which has consequences for the typologies of pro-drop that have been proposed in the literature.

5.3 Null subject parameters within a minimalist approach

In the early 1990s Chomsky proposed the Minimalist Program with the aim of simplifying the GB framework, in effect, to minimize the number of categories and rules involved in syntactic operations. Avoided also in the minimalist program are rules or principles that apply only at certain levels of representation. Another important part of earlier generative grammar which has been re-evaluated in recent years is the status of parameters in the theory. A parameter is a point of variation allowed by a principle of UG, e.g., there is a universal principle that an X' (-bar) phrase consists of a head X and a complement YP. The principle allows variation regarding the linear order of head and complement. This is the so called head parameter. Clearly, whether a language chooses the order head-comp or comp-head will have a huge effect on the surface form of sentences. This is known as a macro-parameter, one which has a wide range of effects in different structures. In recent years, the reality of macro-parameters has been called into question, also by generative linguists. According to Pica (2001, cited in Camacho 2013:6–7), 'Twenty years of intensive descriptive and theoretical research has shown ... that such metaparameters [e.g., the Null-subject parameter, or the Polsynthesis Parameter] do not exist, or, if they do exist, should be seen as artefacts of the 'conspiracy' of several micro-parameters.' Micro-parameters are more superficial parameters with only a restricted range of surface effects. Camacho cites Baker (2008) as having observed that 'if only macro-parameters existed, languages should cluster around a positive or negative value of the parameters with nothing in between and with no mixed cases,' which is not the case for any proposed macro-parameters. For example, there are probably more languages with mixed headedness than there are pure head-final or head-initial languages.

On the other hand, as Baker (2008) also points out, if there were only micro-parameters, we would not predict any clustering, for example, we would expect languages to be evenly distributed along the spectrum from head-final to head-initial, but this is not what we see. As shown by Baker (2008), languages tend to cluster at either end of the spectrum, even if there are also languages that are in between. Therefore, both macro- and micro-parameters are crucial for accounting for cross-linguistic variation. In the case of well-established parameters like the null subject parameter, it is well-accepted that there are several parameters involved, accounting for the variation found among the languages of the world.[79]

Roberts and Holmberg (2010) divide NSLs into four kinds: i. consistent NSLs, e.g., modern Greek, Italian; ii. expletive NSLs, e.g., German, some varieties of Dutch; iii. radical or discourse pro-drop languages, e.g., Chinese, Japanese; and iv. partial NSLs, e.g., Finnish, Hebrew. The diagnostic features of these four kinds are discussed below.

Consistent NSLs (i. above) have rich agreement; person, number and sometimes gender are marked on the verb in these languages. As a consequence, in finite clauses they allow null subjects by default, that is, 'the definite subject pronoun [remains] unexpressed in any person-number combination and in any tense' (Roberts and

---

[79] This has, in fact, been recognized for a long time in serious work on null arguments. For example, Rizzi (1982) assumed that there were two parameters. Rizzi (1986) suggested that there was a separate parameter to account for null arguments in agreementless languages. See Roberts and Holmberg (2010) and Holmberg (2010) for discussion.

Holmberg 2010: 7–8). However, even these languages require overt pronouns when they have an emphatic interpretation and when they indicate a change of topic. Another distinct feature of these languages is that they require an overt third person singular inclusive generic (indefinite) pronoun like English *one* in active finite clauses.

Expletive NSLs (ii. above, also called 'semi-pro-drop languages') allow expletive NSs but do not allow definite (referential) NSs.

Radical or discourse pro-drop languages are distinguished by the following properties: they

(a) allow the antecedent of a null argument to be in a separate sentence;
(b) allow pro drop of subjects and other arguments; and
(c) have no agreement.

Though Roberts and Holmberg (2010) note that the terms radical and discourse pro drop (iii. above) are often used synonymously, this is potentially misleading. Note that the epithet 'radical' suggests that in these languages, the features of the null pronoun are radically absent because they have no agreement. This entails that the features have to be recovered from the discourse. But the epithet 'discourse pro drop' does not, in itself, suggest anything regarding agreement. Therefore, applying the label 'radical pro drop' to languages which have agreement, e.g., Tamil should be avoided because the pro drop is, in fact, not radical: the agreement features on the verb shows which features the missing pronoun has. But the term 'discourse pro drop' can be used because it has properties (a) and (b): the interpretation can be provided in a wide range of discourse contexts. Sinhala is also a discourse pro-drop language, having properties (a) and (b). But it can also be called a radical pro drop language, since, in addition, it has property (c). Chinese, Japanese, Sinhala and Tamil would all be discourse pro drop languages, but only Chinese, Japanese, and Sinhala would be radical pro drop languages. This raises a serious question, though, that is, whether the distinction between radical and non-radical discourse pro drop languages is actually a meaningful one.

Partial NSLs are languages which are intermediate between consistent NSLs and non NSLs. They allow null subjects under certain conditions. For example, Finnish, as Holmberg (2005) shows, allows first and second person null subjects in finite clauses, while it allows a third person null subject only when it is bound by a c-commanding antecedent. Moreover, indefinite, inclusive generic pronouns can, and must be null in partial NSLs (ibid. 540) unlike in consistent NSLs, in which they are phonologically expressed.

Holmberg (2010: 94–95) adopts his (2005) proposal that consistent null-subject languages have a D(efinite) feature as part of the phi (φ)-feature make-up of finite T, which is absent in partial null subject languages. The φ-features are syntactic features such as person, number, gender and also case (PNG in 3.6.1.2; also see below). Further, he makes two assumptions: (a) Pronouns are either DPs, with the structure $[_{DP}\ D\ [_{\varphi P}\ \varphi\ [_{NP}\ N]]$, or φPs; and (b) Null pronouns are φPs. He makes two proposals with regard to these two kinds of languages:

i) in the consistent NSLs, the probe-goal relation between T with its D and a φP results in the union of the φ-features of T and the subject, yielding a definite pronoun;
ii) in the partial NSLs, the probe goal relation between T without D and a null φP does not afford a definiteness value, which results in a D-less, indefinite, subject pronoun. If the φ features are 3SG, the interpretation is that of an inclusive generic pronoun, corresponding to *one* in English.

These two proposals are in compliance with Chomsky's (2001) ideas about feature checking/valuing, specifically the idea that the uninterpretable φ-features of T in order to be valued need to enter into an Agree relation with the subject DP. In consistent NSLs, as in all NSLs, the third person null subjects are dependent on the antecedent, although the conditions on the pronoun-antecedent relation are less strict in consistent NSLs than in partial NSLs. Holmberg (2010) concludes that consistent NSLs allow null definite subject pronouns, but requires overt generic subject pronouns, whereas partial NSLs allow null definite pronouns only when they are locally c-commanded by an antecedent, but allow null generic subject pronouns.

Holmberg's proposals outlined above adequately explicate the occurrence of null subjects in consistent and partial NSLs, but have no obvious consequences for languages without agreement. The following is a theory of null arguments in such languages which has attracted much attention in recent years. In his analysis of null pronouns in Japanese, Tomioka (2003: 335) analyzes them as 'the phonologically null version of bare NPs, which requires a limited set of semantic operations to be interpreted appropriately.' He identifies a wide range of uses of null pronouns with distinct semantic interpretations in Japanese. Classifying what he calls pro-drop languages into ‘agreement pro-drop languages’ and ‘discourse pro-drop languages’, he notes that in the former, the strong inflectional morphology on verbs is believed to license pro, whereas the latter use a condition based on ‘discourse familiarity’. Tomioka (2003: 336) claims that the discourse pro drop languages share a morphosyntactic property which he refers to as ‘Discourse Pro-drop Generalization’ given in (11) below:

(11) All languages which allow discourse pro-drop allow (robust) bare NP arguments.

One of the most widespread forms of null elements across languages, Tomioka (2003) adds, is the ‘phonologically null NP anaphora, also known as N′-Deletion/NP ellipsis’. Observing that ‘every language is equipped with some strategy for not overtly expressing a redundant/familiar portion of a DP,’ he claims that the outcome of this operation can be different from language to language: in languages like English which require obligatory DP projection, the leftover D-heads are overt—in other words, the determiners strand, while in discourse pro-drop languages the entire nominal phrases are phonologically null. Tomioka’s (2003: 336) contention is that ‘[n]ull pronouns in the latter kind of languages are the result of N'-deletion/NP ellipsis without determiner stranding’.

Modesto (2008) tries to show that in languages such as Brazilian Portuguese, Finnish and Chinese, NSs may be licensed and identified without the participation of SV agreement. Proposing that null subjects are elided topics, Modesto (2008: 401) claims that ‘[t]he existence of null subjects in these languages is due to the fact that they are topic-prominent languages [...].’ What sets them apart is that they have a topic deletion rule, deleting fronted topics, in discourse contexts where the deletion can be recovered.

The above discussion summarizes some important theoretical frameworks within which the occurrence of null subjects in languages is analyzed. In 5.4 below, the occurrence of null subjects in Tamil and Sinhala is examined in detail.

5.4 Analysis of null arguments

Examining the property of licensing null subjects in Tamil and Sinhala in light of what has been discussed thus far and the results of the survey above, it can safely be argued that Tamil and Sinhala are neither expletive NSLs nor non-NSLs as they do allow referential null subjects. There are, therefore, three possibilities: Tamil and Sinhala are i) consistent NSLs; or ii) partial NSLs; or iii) discourse/radical pro drop languages. The results of the survey, based on Cole (2010), show that they do not behave like consistent null languages, e.g., modern Greek and Italian. The two languages are, therefore, either partial NSLs or discourse pro drop languages (with Sinhala in the radical pro drop subcategory). This section intends to examine the behaviour of null subjects in the two languages.

Holmberg's (2010) initial tests to distinguish between consistent and partial NSLs are used here with the aim of determining in which contexts the two languages allow or require null subjects. The first test is to find out what kind of subject the two languages have in the embedded clause when the subject is controlled by a higher argument, often the topic in the matrix clause, as illustrated in (12):

(12) T. ravi$_1$ [Ø$_{1/}$thaan$_1$/avan$_{1/2}$ oru yaanai-ai kan-d-aan endu]
Ravi self he one(INDF) elephant-ACC see-PST-3SGM COMP
son-n-aan
say-PST-3SGM
S. ravi$_1$ [Ø$_1$/thaman$_{1/}$eyaa$_{1/2}$ aliy-ek-wə thækka kiyəla] kiwwa
Ravi self he elephant-INDF-ACC see.PST COMP say.PST

'Ravi said that he saw an elephant.'

In Tamil and Sinhala the embedded subject is optionally null. By contrast, in consistent NSLs like Italian, the embedded subject is obligatorily null (Holmberg 2010: 91). Note that another feature which Sinhala shares with Tamil is that its third person singular distal demonstrative pronoun *eyaa* 's/he' can function as an anaphor (reflexive) and a pronoun (pronominal) like the corresponding third person pronoun *avan/l* 'he/she' in Tamil/Dravidian; hence, *avan* /*eyaa* in (12) can co-refer with *Ravi* or another individual. The second test concerns the locality conditions that hold for the null subject-antecedent relation. The objective is to test whether an intervening, c-commanding argument will necessarily block the relation between a null subject and a more distant antecedent. (13) is a case in point.

(13) T. nimal$_1$ ondum soll-a-(v)illai aanal ravi$_2$ [Ø$_{*1/2}$/thaan$_2$/avan$_{1/2}$
Nimal nothing say-INF-NEG but Ravi self he
oru yaanai-ai kan-d-aan endu] sollu-r-aan
one(INDF) elephant-ACC see-PST-3SGM COMP say-PRS-3SGM
S. nimal$_1$ monəwath kiww-e nææ namuth ravi$_2$ [Ø$_{*1/2}$/thaman$_2$/eyaa$_{1/2}$
Nimal nothing say-PST-NMLZ NEG but Ravi self he
aliy-ek-wə dække kiyəla] kiyənəwa
elephant-INDF-ACC see.PST COMP say.PRS

'Nimal hasn't said anything, but Ravi says that (he) saw an elephant.'

Given that the embedded subject in the second clause is null in Tamil and Sinhala, it will co-refer with the immediate antecedent, *Ravi*, and not with the topic of the first clause, *Nimal*.[80] Only an overt pronoun can co-refer with the more distant argument. In this respect, they behave like partial null-subject languages discussed by Holmberg (2010). Note, however, that Tamil and Sinhala like consistent pro-drop languages can allow a null subject optionally, even if the antecedent occurs in a separate sentence (which may be articulated by a different person), as is evident in (2) and (3) above (also see examples below). Thus, they differ from partial pro drop languages, e.g., Finnish which allow third person null subjects only when they have a c-commanding antecedent.

As discussed above, consistent NSLs require overt indefinite (generic) pronouns, corresponding to the English generic pronoun 'one', whereas partial NSLs and discourse pro drop languages allow null indefinite pronouns. The two languages under investigation allow null generic inclusive pronouns (subjects in this case), as in (14):

(14) T. intha kathira-il Ø vasathi-aaha irukk-a-laam
This chair-LOC one comfortable-ADV sit-INF-may
S. mee putuw-ee Ø pahasuwen vaadiwela innə puluwaŋ
This chair-LOC one comfortable.ADV sit.PTCP be-INF can

'One can sit comfortably in this chair.'

---

[80] Note, however, that example (24) below indicates that this locality condition can be violated.

The null generic subject is the default form in instances like these.[81]

On the basis of these tests, it can safely be argued that Tamil and Sinhala are not consistent NSLs. They are similar to partial NSLs in that they allow null definite subjects in some instances and null generic subjects in others. However, they are different from partial NSLs because, unlike them, Tamil and Sinhala can allow null definite pronouns even when these null subjects are not c-commanded by their antecedents, but are available in the sentence-external discourse—provided the antecedents occur within close proximity—as in (15):

(15) T. kamala$_i$ oru thoppi vaangi-n-aal piraku Ø$_i$ ath-ai/antha thoppi-ai
Kamala one(INDF) hat buy-PST-3SGF later (she) it-ACC/that hat-ACC
leela-(vu)kku kudu-th-aal
Leela-DAT give-PST-3SGF
S. kamala$_i$ thoppiy-ak gaththa passe Ø$_i$ eekə/ee thoppiyə
Kamala hat-INDF buy.PST later (she) it/that hat
leela-tə dunna
Leela-DAT give.PST

‘Kamala bought a hat. Later, she gave it to Leela.’

Tamil and Sinhala are also different from the partial pro-drop languages discussed by Holmberg (2005, 2010) in that they have a third person singular inclusive generic pronoun which can be overt (see fn. 81 below).

The above discussion provides reasonable evidence for the conclusion that Tamil and Sinhala are discourse pro drop languages. They show identical behaviour with regard to the occurrence of null subjects despite the difference between them in agreement. This strongly implies that agreement has no role to play in the occurrence of pro drop in Tamil, certainly not the crucial role that it has traditionally been claimed to have.

The following reasons can be stated in support of the claim that agreement has little or no role to play in Tamil. Although only the subject agrees with the verb in Tamil, it allows other arguments such as direct and indirect objects, as shown in the

---

[81] However, the overt generic pronoun derived from indefinite ‘one’ can also be found, as in (i):

(i) T. intha kathira-il (oru-var) vasathi-aaha irukk-a-laam
This chair-LOC one(INDF)-3SG(HON) comfortable-ADV sit-INF-may
S. mee putuw-ee ((ek)-ken-ek-(u)tə) pahasuwen vaadiwela innə puluwaŋ
This chair-LOC one-3SG(HUM)-INDF-DAT comfortable.ADV sit.PTCP be.INF can

‘One can sit comfortably in this chair.’

exchange between A and B in (16), to be null (see below for more examples); Sinhala, too, allows these arguments to be null:

(16) T. A: enna-tta puththakam iru-kkir-athu
I-LOC book be-PRS-3SGN
B: oru naalai-kku Ø$_{\mathrm{IO}}$ Ø$_{\mathrm{DO}}$ thaa-ree-ngal-aa
One(INDF) day-DAT (me) (it) give-FUT-2SG(HON)/PL-Q
S. A: maŋ-gaavə pothə thiyenəwa
I-LOC book have.PRS
B: (ekə) davəs-ək-(ə)tə Ø$_{\mathrm{IO}}$ Ø$_{\mathrm{DO}}$ denəwa də
one day-INDF-DAT (me) (it) give.FUT Q

A: ‘I have the book.’ B: ‘Will you give it to me for a day?’

In (16), the direct object and the indirect object are null, with antecedents in the discourse context. Moreover, Tamil allows null subjects in clauses which do not show any agreement, e.g., in clauses containing modals and negatives.

It could be noted that Barbosa (to appear) rejects Holmberg’s (2005, 2010) distinction between partial pro drop and discourse pro drop languages. Barbosa (to appear: 11) claims that null subjects in consistent NSLs are dependent on agreement with T, the idea being that the phi(φ)-features of T are interpretable in these languages. In other words, the null subject is specified by the φ-features of T in these languages. In partial pro-drop languages, on the other hand, subject pro drop would not be directly linked to the properties of agreement inflection (see below for more details). In this respect, partial pro-drop languages would be similar to discourse pro drop languages, e.g., Chinese in which null subjects are licensed without any agreement morphology.

The discussion above affords the following properties which characterize the occurrence of null subjects in Tamil and Sinhala:

(i) they allow null subjects in embedded clauses if they have a c-commanding topic antecedent in the matrix clause;
(ii) they also allow null subjects if the antecedents are topics in a separate sentence within close proximity in the discourse;
(iii) they allow null generic (indefinite) subjects;
(iv) agreement morphology in Tamil has no role to play in licensing null subjects;
(v) they also allow null objects.

On the basis of (i)–(v), it can safely be concluded that Tamil and Sinhala are discourse pro drop languages. Sinhala is also a radical pro drop language, as it has no SV agreement. No evidence has been seen so far, though, that this distinction would play any role in relation to pro drop.

In the rest of this section, the occurrence of null arguments in the two languages is analyzed in terms of the syntactic operations proposed in the literature for the explication of the occurrence of null subjects in discourse pro drop languages.

Adapting Abney's (1987) conceptualization of pronouns, Jayaseelan (1999: 21) characterizes pronouns, particularly the third person pronouns, as 'determiners inflected for the PNG-[φ-] features of the "missing" noun phrase complement.'[82] Following from the same characterization of pronouns, Jayaseelan (1999) claims that the pronouns in English are formed from the definite articles, e.g., *he* is in effect *the* + 3SGM. Dravidian, however cannot form their pronouns from definite articles, because the definite article is null in Dravidian. More precisely, the overt pronouns cannot be formed from the definite article. This is why the third person pronouns in Dravidian are derived from demonstratives formed by the combination of the demonstrative stem and the pro form base encoding the φ-features (see 4.3.1). All these facts hold true for Sinhala, as well (see 2.3.1.2 and 4.3.1). The formation of the third person pronouns in the two languages is shown in (17):

(17) T. a + -an / -al → a(v)an / a(v)al
that [+ singular,+ masculine] [+singular, + feminine] he she
S. ee + yaa → eyaa
that [+ singular] s/he

Since the definite article is null in these languages, the φ-feature base cannot be phonologically realized because it does not have a phonetic matrix to attach itself to, resulting in empty pronouns, subjects as well as non-subjects (Jayaseelan1999), as shown in (18):

(18) T. Ø + an → Ø
[+singular, +masculine]
S. Ø + yaa → Ø
[+singular]

As Jayaseelan (1999) points out, this analysis presupposes that Dravidian languages have an underlying definite article—although it is phonologically null—with the set of

---

[82] Abney (1987 cited in Jayaseelan 1999) conceptualizes pronouns as 'intransitive determiners' in the sense that they do not require any complement, in contrast to a usual DP in which the D(eterminer) cannot stand alone, but requires a complement, an NP, in which case it is transitive.

φ- features. Sinhala too have an underlying definite article, phonologically null, with the set of φ- features**.**

Tomioka (2003) argues that the diverse uses of full-fledged NPs are derived from one basic meaning, property anaphora (type <e,t>) and their differences are the result of two independently needed semantic operations, namely Existential Closure (yielding indefinite interpretation) and Type Shifting to an individual (yielding definite interpretation) (see Barbosa to appear). He proposes that null arguments in Japanese, a discourse pro-drop language, are derived by what he refers to as null NP anaphora a.k.a N' or NP-ellipsis (see also (11) above). In a language like Japanese which lacks determiners, this deletion leaves a null D behind, giving rise to null arguments. In this respect, Jayaseelan's (1999) theory is related to Tomioka's (2003) theory. In languages with overt definite and indefinite articles, like English, the determiner remains after NP ellipsis; hence, this process will never yield a phonetically null argument. Tomioka's and Jayaseelan's theories both link two properties that known discourse pro drop languages have, along with Tamil and Sinhala: (i) that the two languages lack overt definite articles (2.3.1.2), and (ii) that the languages allow pro drop with subjects as well as non-subjects independently of agreement. This idea is also countenanced in Barbosa (to appear).

NP deletion, i.e., deletion under identity, is claimed to be one of the mechanisms, by which null arguments are derived in discourse pro drop languages, taken to be languages where agreement plays no role in relation to pro drop and there are no definite markers. The most significant diagnostic for NP deletion in discourse pro drop languages such as Japanese, Korean and Chinese, according to Takahashi (2013), is whether it affords a sloppy reading in addition to a strict reading in instances like (19):

(19) T. A. prasanna than-da mahan-ai kadai-kku anupi-n-aan
Prasanna self-GEN son-ACC shop-DAT send-PST-3SGM
B. kumar-um Ø anupi-n-aan
Kumar-INCL send-PST-3SGM
S. A. prasanna thaman-ge putha-wə kadee-tə yæwwa
Prasanna self-GEN son-ACC shop-DAT send.PST
B. kumar-uth Ø yæwwa
Kumar-INCL send.PST

'Prasanna sent his son to the shop. Kumar also sent his son to the shop.'
Strict reading: Kumar sent Prasanna's son.
Sloppy reading: Kumar sent his own son.

As Takahashi (2013) rightly points out, that the underlying full-fledged object NP, the anaphor, in the embedded clause in (19T/SB) is elided under identity with the object NP in the embedded clause in (19T/SA). The fact that (19T/SB) yields both sloppy and strict readings suggests that the null objects arise from NP-deletion. A problem with this NP-deletion analysis is that it can be used to explain only the occurrence of null objects, not null subjects, in Tamil and Sinhala because instances like (20T/SB) which involve null subject NPs yield only a strict reading and not a sloppy reading:

(20) T. A. prasanna [than-ra mahan nalla inglish kathai-kir-aan endu]
Prasanna self-GEN son well English speak-PRS-3SGM COMP
ninai-kir-aan
think- PRS-3SGM
B. kumar [Ø nalla french kathai-kir-aan endu] ninai-kir-aan
Kumar well French speak-PRS-3SGM COMP think-PRS-3SGM
S. A. prasanna thaman-ge putha hondətə ingriisi kathaakərənəwa kiyəla]
Prasanna self-GEN son well English speak.PRS COMP
hithənəwa
think.PRS
B. kumar [Ø hondətə french kathaakərənəwa kiyəla] hithanewə
Kumar well French speak.PRS COMP think.PRS

'Prasanna thinks that his son speaks English well. Kumar$_2$ thinks that his$_{1/2}$ son speaks French well.'
Strict reading: Kumar thinks that Prasanna's son speaks French well.
Sloppy reading: *Kumar thinks that his own son speaks French well.

According to Barbosa (to appear), of the discourse pro drop languages, only Japanese and Korean yield both strict and sloppy readings in instances like (20B), while others including Chinese can yield only a strict reading. The fact that (20T/SB) cannot yield a sloppy reading implies that the null subjects in the embedded clause in (20T/SB) do not result from NP ellipsis.[83]

NP deletion also cannot explain cases of generic null arguments, as in (21):

(21) T. Ø kaasu irunth-aal, Ø ellam vaang-a-laam
Money be.PTCP-if all buy-INF-may
S. Ø salli thibbo-th, Ø okkomə gandə puluwaŋ
Money have-PTCP-if all buy.INF can

'If one has money, s/he can buy everything.'

---

[83] The reason why instances like (20T/SB) do not yield sloppy reading is not discussed here owing to constraints of space.

These null arguments have no antecedent, that is, the null subject here is non-anaphoric; hence, it is unclear what would be elided. Note that an overt generic pronoun can also occur in (21). Holmberg and Phimsawat (2015) propose that null inclusive generic pronouns, counterparts of English generic *one*, are minimal nouns, which have neither D nor φ-features. As such, they have unrestricted reference: the speaker, the addressee, and everyone else.

Barbosa (to appear) argues that not all instances of empty subjects can be reduced to NP ellipsis because there is no way the impersonal (indefinite) null subjects can be an elided form, given that it has no antecedent. Therefore, she concludes that to account for this kind of impersonal null subjects and the occurrence of null subjects which do not involve NP ellipsis in languages like Chinese, some mechanism other than NP ellipsis is needed. According to Barbosa (to appear: 24) there are two ways of deriving null arguments: 'by ellipsis of a full-fledged nominal or by base generating an empty pro-form.' The former applies to objects in Japanese, Korean, Chinese, Turkish, etc. as well as subjects in Japanese and Korean, while the latter is independently needed in the case of the non-anaphoric (indefinite) null subject ((14) above) and in addition, it is also required for the anaphoric null subject in languages with agreement between the subject and T (Chinese included).

As for definite null subjects in languages which lack definite articles, Barbosa (to appear: 32), following Tomioka (2003), hypothesizes that it is topicality that enables the null NP, a function of type <e,t>, to be shifted to an individual (a denotation of type <e>) and hence be interpreted as definite. Based on the assumption that topics denote individuals that the sentence as a whole is about, she observes that there should be a relation between topicality and type-shifting to an individual.

It is instructive to note that Modesto (2008) tries to establish a correlation between the occurrence of null arguments, subjects as well as non-subjects, in discourse/radical pro drop languages and the fact that these are Topic Prominent Languages (TPLs) like Japanese and Chinese. He draws upon the following characteristics which Li and Thompson (1976) identify in TPLs as opposed to Subject Prominent Languages (SPLs) like English (Modesto 2008: 402):

(i) There is a "surface coding" for topics in those languages; they are coded by a special topic marker or appear in sentence-initial position.

(ii) Passive constructions, which are common in SPLs, are absent or marginal and rarely used in speech, in TPLs.

(iii) Dummy subjects are not found in TPLs.
(iv) TPLs use "double subject" constructions.
(v) The topic typically controls co-referential constituent deletion.
(vi) TPLs tend to be V-final.
(vii) In TPLs, there are no constraints on what constituent may be the topic.
(viii) Topic-comment sentences are the basic sentence type in TPLs.

Tamil and Sinhala have most of the characteristics of TPLs. The two languages have property (i) in that they have a special topic marker which occurs adjacent to the topic of the sentence (4.3.2). With regard to property (ii) passive constructions of the two languages are not frequently used in speech (3.3.2). The two languages have property (iii) because they do not have dummy subjects. As for property (iv), although they do not have double subject constructions like those in Chinese or Japanese (Li and Thompson 1971), they have a construction, as in (22), which resembles double subject constructions. This construction could conceivably be viewed as a double subject construction because it has a dative subject and a nominative subject (Shibatani 1999):

(22) T. padma-(u)kku velai kidai-th-athu
Padma-DAT job get-PST-3SGN
S. padma-tə rassaawə hambuna
Padma-DAT job find.PST

'Padma got the job.'

Arguing that this construction is similar in structure and meaning to the double subject construction, Shibatani (1999: 48) points out that it has the structure '[NP-DAT NP-NOM PRED]'. This structure looks plausible for (22) because in Tamil, which has SV agreement, it is the nominative NP 'job' in (22), being a third person singular neuter NP, that agrees with the verb. By virtue of the occurrence of the second subject, (22) is different from the regular dative subject construction found widely in South Asian languages (see (43) in 4.3.11).

The two languages also have property (v), but this kind of topic is not commonly used in the two languages (see fn. 77). The difference between the two languages and a TPL like Chinese is that in the former the topic does not typically control co-referential constituent deletion, as opposed to Chinese in which the topic takes precedence over the subject in controlling the co-referential constituent (see Li and Thompson 1971). Instead, in Tamil and Sinhala, topic antecedents control null arguments in the embedded clause. Properties (vi) and (vii) hold true for the two languages because they are consistently ‘V-final’ and they have no constraints on what constituent can be the topic. Tamil and Sinhala do not have property (viii), for the topic-comment type of sentences is not the basic type in them.

If Modesto is right, it may be safe to conclude that Tamil and Sinhala —given the characteristics of TPL they have—are discourse pro drop languages in the manner in which they allow null subjects/non subjects. One of the likely characteristics of TPLs which Modesto (2008) fails to include in the list of characteristics is that generally these languages do not have definite markers. Barbosa (to appear) claims that there is a correlation between topicality and definiteness. The definite articles play a vital role in referring to an entity whose identity has already been established, thus definite articles distinguish old information from new information. Note that by default, the topic encodes old information, and hence has definiteness inherently. A likely reason that TPLs do not typically have definite articles is that TPLs can move any nominal constituent to the topic position (4.3.2), where they will receive a definite interpretation with no need for an article.

Barbosa (to appear) claims that topic prominence is responsible for two related features. First, the third person null subject may occur in a matrix clause, as in the exchange between A and B in (23):[84]

(23) T. A. mala enge B. Ø veet-ai poo-itt-aal
Mala where home-ACC go.PTCP-leave-PST-3SGF
S. A. mala koo B. Ø gedərə giya
Mala where home go.PST

‘Where is Mala?’ ‘She has gone home.’

The second feature is that a third person NP may be bound by a salient discourse topic (overt or null) across a subject, as in the exchange between A and B in (24); ((23) and (24) are adapted from exs. (67) and (68), p. 44 in Barbosa to appear):

---

[84] This analysis is concerned only with third person null subjects because according to Sigurdsson’s (2004b) hypothesis cited in Holmberg (2010), with regard to first and second person null subjects, every clause has features representing the speaker and the addressee in the C-domain. In this way, the speaker and the addressee are always available as local antecedents, facilitating recoverability of these null pronouns.

(24) T. A. mala B. rani/sanam [Ø london-(u)kku pooi-tt-aal
Mala Rani/People London-DAT go.PTCP-leave-PST-3SF(HON)
endu] sollu-r-aal/ kathai-k-(k)inam
COMP say-PRS-3SGF talk-PRS-3PL
S. A. mala B. rani/minissu [Ø londonvələ-tə giya kiyəla]
Mala Rani/people London-DAT go.PST COMP
kiyənəwa
say.PRS

'What about Mala? Rani says/People say that she has gone to London.'

Despite the intervening subject of the matrix sentence 'Rani/People' (24), the third person null subject, as Barbosa (to appear: 44) observes, 'can arguably be derived by movement of the zero subject to the matrix topic position.'[85] This movement of the null subject to the topic position, as mentioned above, facilitates the recoverability of reference of the null NP, in effect, the semantic function of type shifting of the function type <e,t> to an individual of type <e>, 'Mala' being interpreted as definite, resulting in the null NP. In the two languages, therefore, the referential null arguments entail null NP anaphora, whereas the null indefinite pronouns ((14/21) above), being non-anaphoric, are base generated.

The above discussion indicates that there is a correlation between the occurrence of null subjects in discourse pro drop languages and the fact that they are topic prominent. It should be noted, though, that Barbosa (to appear) as well as Modesto (2008) argue (on partly different grounds) that these properties are also shared with partial pro drop languages. If this is right, the distinction between these two types may be invalid. Barbosa (to appear) also shows that at least some of the languages classified as partial NSLs allow object pro-drop as well as subject pro-drop. Moreover, it still seems to be the case that the discourse pro drop languages allow overt (see fn. 81) or null inclusive generic subjects, while the partial NSLs discussed in the literature only have a null 'one'.

---

[85] Note that (24) may contradict a generalization established above with respect to the occurrence of null subjects in the two languages, on the basis of (13). In (13), the kind of non-local binding seen in (24) was not possible. The structure is different, though, in a way which is likely to be crucial: in (24) the embedded clause containing the null subject is in a position where it can be expected to be more accessible to the antecedent than it is in (13). Investigating this difference further will be left for future research.

5.5 Conclusion

The difference in agreement between the two languages notwithstanding, Tamil and Sinhala display the same behaviour with regard to the occurrence of null arguments. From a comparative perspective they are close to Chinese, based on Cole's (2010) criteria and the results of the tests discussed in this chapter: they allow null subjects with a c-commanding antecedent in the same sentence, as well as with a topic antecedent in a separate, preceding sentence, given that certain locality conditions are met. Further, like Chinese, they allow not just null subjects but also null non-subject arguments. The fact that the two languages share with Chinese the same behaviour with respect to the occurrence of null arguments even though they are not genetically related and have not been in contact, further underlines the conclusion that the typology of null subjects/null arguments needs to be re-examined. In particular, the lack of or no agreement in languages can no longer be postulated as a defining property of discourse pro-drop languages. A type called discourse pro-drop languages may still be established with the following distinct characteristics:

(a) They allow not only null subjects but null non-subject arguments;
(b) Null subjects/non subjects can have an antecedent in a separate, preceding sentence.

An additional characteristic may be (c):

(c) They have an optionally null inclusive generic pronoun.

Properties (a) and (c) set them off from consistent NSLs. Properties (b) and (c), and arguably (a), set them off from partial NSLs. Cole's (2010) results, combined with the results obtained from other tests in this chapter, show that there are significant differences among these languages, too: Japanese allows null subjects in several contexts where Chinese, Tamil and Sinhala do not allow them (note again that there is no correlation with agreement). It is crucial to note that despite the difference between SV agreement in Tamil and Sinhala, the occurrence of null subjects in them involves the same processes, namely the null NP anaphora and base generation of empty pro forms: by and large, the former accounts for the referential null arguments, while the latter accounts for the null indefinite pronouns in the two languages (see Barbosa (to appear)).

What may also be distinguished within the category of the discourse pro-drop languages is a subcategory of agreementless languages; this would be the true radical pro-drop languages, comprising Chinese, Japanese, and Sinhala (see table 5.2). However, comparing their properties with those of Tamil indicates that this distinction is not typologically useful because the languages seem to have nothing in common which would set them off from Tamil. What the preceding discussion has shown instead is that the discourse pro-drop languages have certain other properties like (d) and (e) below in common, which may well be crucial:

(d) They have no definite markers;
(e) They are topic-prominent.

Property (d) may be crucial if Jayaseelan's (1999), Tomioka's (2003), and Barbosa's (to appear) theories are on the right track, while property (e) may be crucial if Modesto's (2008) and Barbosa's (to appear) theories are on the right track. But note that properties (d) and (e) are, arguably, found also among the partial pro-drop languages identified by Holmberg (2005, 2010), including Brazilian Portuguese, Finnish, Hebrew, and Marathi. The information discussed in this section indicates that Roberts and Holmberg's (2010) typology (expletive pro-drop, consistent pro-drop, partial pro-drop, and discourse pro-drop; see 5.3 above) is not the right basis for an explanatory theory of pro-drop.

What perceptions of the relation between Tamil and Sinhala do the facts discussed in this chapter provide? The fact that they show identical behavior with respect to the null subject-antecedent relation, as seen in the experiment based on Cole (2010) and the results of the tests in other works, is obviously consistent with the more general situation, which is that the two languages are syntactically similar, and the close contact between the languages is likely to be at least part of an explanation for the same behaviour they display with regard to the occurrence of null subjects/non subjects. Works on the occurrence of null arguments on Malayalam (Jayaseelan 1999 and Takahashi 2013) show that with regard to the licensing null arguments, Malayalam displays almost similar properties to those of Tamil and Sinhala.[86] It could probably be that this is a more general South Asian Sprachbund (areal) feature. Establishing whether

---

[86] The only detailed investigation that I am aware of is Holmberg, Nayudu, and Sheehan (2009) on Marathi, an Indo-Aryan language. Marathi has the properties of a partial pro drop language, except that unlike some other languages of this type, it allows a null subject to have an antecedent in a separate sentence, like a discourse pro drop language.

this is so requires a detailed comparative study covering a range of other Dravidian and NIA languages.

Chapter 6

Focus constructions

6.1 Preliminaries

No morphosyntactic phenomenon which Tamil and Sinhala share can provide more evidence for contact-induced restructuring of Sinhala than the two kinds of focus constructions which represent the two ways in which exhaustive focus is produced in the two languages. Focus singles out a sentential constituent and presents it either as new information in the discourse or as an exhaustively identified element out of an understood list of alternatives. The former is called information focus, the latter, exhaustive focus. Exhaustive focus can be assigned morphosyntactically in Tamil and other Dravidian languages in two ways (Sarma 1999 and 2003, Krishnamurti 2003) which Sinhala too possesses (Gair 1980, Gair and Paolillo 1988 and Chandralal 2010): (i) by clefting whereby the constituent focused is moved to clause-final position, as in (1b); or (ii) by cliticizing *-ee* in Tamil and *-may* or -(*u*)*y* in Sinhala to the constituent focused or adjoining the focus particle *thaan* in Tamil and *thamay* in Sinhala 'in deed' to the right of the constituent focused in situ— the focus clitic/particle has scope over either only the constituent focused, as in (1c) or the entire sentence, as in (1d) below (the unfocused sentence is (1a)):[87]

(1) a. T. kumar nettu mala-(u)kku antha puthakath-ai kudu-th-aan
Kumar yesterday Mala-DAT that book-ACC give-PST-3SGM
S. kumar iiyee mala-tə ee pothə dunna
Kumar yesterday Mala-DAT that book give.PST

'Kumar gave that book to Mala yesterday.'

(1) b. T. kumar nettu mala-(u)kku kudu-th-athu antha
Kumar yesterday Mala-DAT give-PST-NMLZ that
puthakath-ai-**(ee)/(thaan)**
book-ACC FOC
S. kumar iiyee mala-tə dunn-e ee
Kumar yesterday Mala-DAT give.PST-NMLZ that
pothə-**(may)/(thamay)**
book FOC

'It is that book that Kumar gave to Mala yesterday.'

---

[87] Along with the two kinds of focus, Sarma (1999, 2003) distinguishes a third way of assigning focus to a constituent in Tamil, which is phonological in that the constituent focused is articulated with a rising intonation. Sinhala too can express focus through intonation (Chandralal 2010).

(1) c. T. kumar-(**ee)**/(**thaan**) nett-(**ee)**/(**thaan**) mala-(u)kku-(**ee)**/(**thaan**)
Kumar FOC yesterday FOC Mala-DAT FOC
antha puthakath-ai-(**ee)/**(**thaan**) kudu-th-aan
that book-ACC FOC give-PST-3SGM
S. kumar-(**may)**/(**thamay)** iiyee**-(may)**/(**thamay)** mala-tə**-(may)**/(**thamay)**
Kumar FOC yesterday FOC Mala-DAT FOC
ee pothə-(**may)**/(**thamay)** dunn-e
that book FOC give.PST-NMLZ

‘It was/is Kumar that gave Mala that book yesterday.’
‘It was/is yesterday that Kumar gave Mala that book.’
‘It was/is to Mala that Kumar gave that book yesterday.’
‘It was/is that book that Kumar gave Mala yesterday.’

(1) d. T. kumar nettu mala-(u)kku antha puthakath-ai kudu-th-aan/
Kumar yesterday Mala-DAT that book-ACC give-PST-3SGM
kudu-th-avan **thaan**
give-PST-PRON.3SGM FOC
S. kumar iiyee mala-tə ee pothə dunna **thamay**
Kumar yesterday Mala-DAT that book give.PST FOC

‘In deed, Kumar gave that book to Mala yesterday.’

In the two kinds (1b,c), exhaustive focus can be assigned to any constituent DP/NP. [88] Note that the finite form of the verb in the unfocused sentence (1a) is changed into the nominalized verb form in (1b). The focus clitic/particle may be optionally added to the constituent clefted (as shown in (1b)). A third case is when the focus particle occurs clause-finally, in which case the entire sentence receives exhaustive focus, as in (1d), unlike only the constituent which the focus particle is adjacent to receives exhaustive focus (1c)**.**[89] In this kind of focus construction (1d), the finite form of the verb is used in the two languages and the focus clitic *-ee/-may* is not used. Also used in Tamil is the finite verb form with the pronominal head (glossed as PRON; see (1d))—corresponding to the PNG features of the subject—affixed to it, possibly to give extra emphasis to the

---

[88] In Tamil and other Dravidian languages, the verb (the event encoded) too can be assigned exhaustive focus via clefting (1b) using the dummy verb *sei* ‘do’ which becomes the nominalized form of the verb (see Sarma 1999, Jayaseelan and Amritavalli 2005). This is possible in Sinhala too. Assigning focus to the verb is not common in the two languages unlike assigning focus to other constituents in constructions like (1b). Note also that it is not possible to assign focus to the verb using the focus clitic/particle in constructions like (1c) in the two languages.

[89] An alternative analysis is that the clause-final particle takes scope over the polarity of the sentence, positive or negative. It would be clause-final, c-commanding the entire clause, because polarity is the highest head of the clause (Holmberg in press).

proposition conveyed in the sentence.[90] This finite verb with the pronominal head is identical to the one in headless relative clauses (see 6.3 below).

Almost all the morphosyntactic features of the cleft construction (1b) and the focus particle constructions (1c,d) are the same (for semantic and syntactic implications, see 6.2). The only difference between the two constructions in Tamil and Sinhala is that in the focus particle construction (1c), the finite verb is used in Tamil, whereas the nominalized form of the verb is used in Sinhala.

It will be argued in this chapter, in part based on Gair (1980, 1998a, 1998c, 2007 and 2010) and Paolillo (1994), that the syntax of present day Sinhala has a number of properties which have resulted from the replication of the two kinds of focus constructions from Tamil. This chapter examines the changes which the two focus constructions replicated from Tamil have brought about in Sinhala morphosyntax.

The main focus of this chapter is the cleft construction. Owing to constraints of space, the implications of the focus particle construction are analyzed only where necessary.

In the two languages, the two kinds (1b,c/d) have been widely accepted in the literature (see Sarma 1999, 2003 for Tamil and Gair 1980, 2007, 2010 and Paolillo 1994) as constructions which convey the same meaning as the cleft or pseudo (wh-) cleft construction in many languages. In English, the semantic function of exhaustively identifying a constituent within a sentence requires an equative structure of the type ‘DP COP(ula) (be) DP’, as in ‘The one I admire is John’. The structure of the Tamil/Sinhala cleft construction too is of an equative type ‘DP DP be’, as shown in (2):[91]

(2) T. [$_{CP}$[$_{DP}$ kumar nettu mala-(u)kku kudu-th-athu] [$_{DP}$ antha puthakath-ai] Ø]
Kumar yesterday Mala-DAT give-PST-NMLZ that book-ACC (COP)
S. [$_{CP}$[$_{DP}$ kumar iiyee mala-tə dunn-e] [$_{DP}$ ee pothə] Ø]
Kumar yesterday Mala-DAT give.PST-NMLZ that book (COP)

‘It was/is that book that Kumar gave to Mala yesterday.’/‘What Kumar gave to Mala yesterday was/is that book.’

Under this analysis, the COP, ‘be’ which is null (as in nominal predicate clauses (3.5.3) above) occurs at the end of the clause because these languages are head final. Note that the structure of the English *It*-cleft and pseudo cleft construction is [DP/NP be DP/NP] in which ‘be’ occurs in the middle because English is a head initial language. The null

---

[90] See fn. 44.

[91] As will be shown in (6.2.2), clefting in monoclausal constructions like (2) is TP internal (see Jayaseelan and Amritavalli 2005 and Selvanathan 2012).

COP 'be' in Tamil and Sinhala occurs, by hypothesis, only in cleft constructions like (1b) and does not occur overtly or covertly in the focus particle constructions like (1c) although it has a focus reading (see 6.2.1).

The Tamil/Sinhala construction is different from the English *It*-cleft construction in that in the latter, the focused constituent is constructed with a relative CP, as shown in (3):

(3) [[$_{\text{DP}}$ It] is [$_{\text{DP}}$ that book [$_{\text{CP}}$ that Kumar gave to Mala yesterday]]]

(3) is different from the structure of Tamil/Sinhala cleft construction (2). The latter (2) is more similar to structure of the English pseudo-cleft construction (4):

(4) [[$_{\text{DP}}$ What Kumar gave to Mala yesterday] is [$_{\text{DP}}$ that book]]

Hagstrom (2004) and Kishimoto (2005) distinguish the Sinhala cleft construction as a pseudo-cleft construction. Comparing (2T) with (4) shows that the Tamil cleft construction is also of the pseudo-cleft kind. The difference between the English pseudo cleft and the Tamil/Sinhala cleft construction is that the former has a wh-operator, whereas the latter does not have it.

The almost identical features of the two kinds of Tamil/Sinhala focus constructions (1b,c,d)—given their diverse origins—are too many to have occurred due to chance. Two possible reasons why Tamil and Sinhala share the same features involved in the two kinds of focus are: (a) both languages may have replicated these features from another language with which they were in contact, or (b) either of the two languages may have replicated them from the other. The former can be ruled out because the languages with which Tamil and Sinhala came into contact (see 1.5) do not have one or both kinds of focus constructions, whereas the latter is conceivably plausible given the contact between the two languages (see paragraph below). Moreover, since the two languages are of two different families, it is not possible that they got them from a common ancestor.

Examining the two kinds of focus in Tamil and Sinhala in terms of the 'catalogue of diagnostics' that Heine (2007; see brief description in 1.3**)** proposes for identifying instances of contact-induced grammatical replication confirms that Sinhala has replicated the two structures on the model of those in Tamil. As for 'genetic patterning', all the languages of the Dravidian family which Tamil belongs to have one or both kinds (Krishnamurti 2003, Jayaseelan and Amritavalli 2005), whereas none of the other languages of the NIA family to which Sinhala is believed to belong has the two kinds. In Gair's (1980: 29) words 'I have been unable to find such constructions in any other Indo-Aryan language, and this lack, together with the unmistakable resemblance to Dravidian, justifies the conclusion that the [cleft] construction was a syntactic borrowing from Dravidian …'

The diagnostic of genetic patterning can also be framed in terms of 'genetic inheritance' in that the fact that all Dravidian languages possess one or both kinds of focus indicates that the proto Dravidian should have had at least one if not both kinds of focus, in other words, one or both of these kinds can be reconstructed back to earlier stages of the Dravidian languages. Classical Sinhala, on the other hand, had a kind of focus construction which is different from the two kinds in the modern Sinhala (Paolillo 1994).

In terms of 'intertranslatability' the two focus constructions are mutually intertranslatable word by word, if not morpheme by morpheme (see 1b,c), that is, the two constructions are isomorphic.

Paolillo (1994: fn. 12, p.168) notes that the focus particle *thamay* is *thama* (the anaphor in Sinhala) 'self' + *-y* which is identified historically with focus marking in Sinhala, while the focus clitic *may* is *ma* an emphasising enclitic + *-y*. Just like the anaphor in Sinhala *thaman* which has been modelled on that in Tamil *than*, so too has the focus particle *thamay* in Sinhala been modelled on *thaan* in Tamil which are almost homophonous, like the anaphors *than* and *thaman*. The generation of the focus particle involves matter replication. This is also an instance of replica grammaticalization in that the model for the grammaticalization of the focus particle *thamay* in Sinhala, that is, adapting its anaphor to form its focus particle is provided by the formation of the corresponding focus particle *thaan* adapted from the anaphor in Tamil. A fact in support of the claim that Sinhala has replicated the focus particle construction from Tamil is that the focus particle occurs at the same positions as the corresponding Tamil focus particle (1c/d). In summary, the adoption of the two kinds of focus constructions by Sinhala from Tamil involves both MAT and PAT replication.

The two kinds of focus constructions indicate 'paired grammaticalization' (1.3) in that assigning exhaustive focus to a constituent is the general grammatical function and the focus particle and cleft constructions are the two processes. However, only the former involves grammaticalization (see above), whereas the latter does not involve any grammaticalization (cf. the two topic markers discussed in 4.3.2 and 4.4.2).

Sinhala uses the two kinds of focus constructions far more frequently than Tamil (Paolillo 1994: fn. 17, p. 168) from which it replicated them. The frequent use of these two kinds in Sinhala may have induced significant changes in its morphosyntax. Gair's (1980: 39) conclusion is of importance:

> The Sinhala borrowing of the focusing construction and its subsequent history goes beyond mere adaptation… to illustrate a process for which I have suggested the term "naturalization" … by which a borrowed form enters into the grammar of the borrowing language in an intimate way, participating in its rule structures and even, as in this case, serving as a model for further internal change.

The process of naturalization that Gair distinguishes in Sinhala accounts for the use of the nominalized form of the verb in the focus particle construction and the other morphosyntactic phenomena, e.g., wh-questions, negatives etc. unlike in the corresponding Tamil phenomena.

The analysis of the two kinds of exhaustive focus, especially the cleft construction in relation to other morphosyntactic phenomena shows that the two kinds of focus constructions are either related to or have made changes in the following morphosyntactic phenomena in Sinhala:

(i) it brought about changes in the agreement system that Sinhala once had, resulting in two kinds of finite verbs: the unmarked (default) finite verbs which have lost SV agreement and the marked finite forms which have SV agreement (see 6.4; also Paolillo 1994).

(ii) it led Sinhala to restructure its wh-questions and negative statements in that the nominalized form of the verb used in the Tamil/Sinhala focus constructions has become the default form of the verb in the majority of wh-questions (for exceptions, see 6.5), and negative sentences, though no constituent—neither the wh-phrase nor the negative particle—receives exhaustive focus.

(iii) the focus particle construction has been extended to polar, negative and constituent quotative constructions and also wh-questions in which the respective particle involved occurs either clause-internally adjacent to the constituent/wh-phrase or clause-finally, marking narrow and wide scope respectively.

The aim of this chapter is to substantiate these claims. In what follows, Section 6.2 discusses the semantic and syntactic properties of the cleft construction; Section 6.3 examines the relationship between clefting and relativization, especially the role that relativization plays in clefting; Section 6.4, following Paolillo (1994), traces the development of the present agreement system in Sinhala ever since it replicated the Tamil cleft construction; Section 6.5 deals with the morphosyntactic changes induced by the replication of the two kinds of focus constructions in questions in Sinhala; and Section 6.6 summarizes the findings related to the morphosyntactic changes occurred in Sinhala resulted from the replication of the two kinds of Tamil focus constructions.

6.2 Cleft constructions

An awareness of the two kinds of focus, especially the cleft construction in the two languages is important for the understanding of the ways in which the cleft construction which Sinhala replicated from Tamil interacted with other morphosyntactic phenomena in Sinhala to bring about changes in them. In this section, the cleft construction in the two languages is analyzed in detail. As noted above, here and below, the study is exclusively on the cleft construction, except for instances which require the explication of the focus particle construction. In 6.2.1 below, the semantic and syntactic properties of cleft construction are discussed, detailing its interpretation, its structure and the constraints on clefting. In 6.2.2 below, the syntactic operations involved in the Tamil/Sinhala cleft construction are analyzed.

6.2.1 Properties of cleft constructions

É. Kiss (1998) identifies two kinds of focus, namely identificational focus and information focus. The former draws the attention of the interlocutor to a segment of an utterance by exhaustively identifying it in an understood list of alternatives, while the latter conveys new information, as illustrated in (5) and (6), respectively:

(5) T. mala padi-th-athu **vingnaanam**
Mala study-PST-NMLZ science
S. mala igenəgathth-e **vidyaavə**
Mala study.PST-NMLZ science

‘It was science that Mala studied.’

(6) T. mala VINGNAANAM padi-th-aal
Mala science study-PST-3SGF
S. mala VIDYAAVƏ igenəgaththa
Mala science study.PST

‘Mala studied science.’

In (5) the object is exhaustively focused, i.e., identified, in E. Kiss’s terms, occurring in postverbal position. In (6), the word order is the default SOV order. The object can convey information focus simply by intonation, indicated with uppercase letters. An identificational focus, according to É. Kiss (1998: 245), ‘represents a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate phrase actually holds.’ In (5) ‘science’ is exhaustively focused in that ‘of a set of [subjects] present in the domain of discourse’ (ibid. 247), it was ‘science’ and no other subject that ‘Mala’ studied. The exhaustive focus becomes semantically explicit if it is a response to a (constituent polar) question, as in the exchange between A and B in (7):

(7) T. A. mala padi-th-athu suhaathaaram-aa
Mala study-PST-NMLZ health science-Q
S. A. mala igenəgathth-e soukyə vidyaavə də
Mala study.PST-NMLZ health science Q

‘Is it health science that Mala studied?’

T. B. illai ~~mala~~ ~~padi-th-athu~~ vingnaanam
No Mala study-PST-NMLZ science
S. B. nææ ~~mala~~ ~~iganagathth-e~~ vidyaavə
No Mala study.PST-NMLZ science

‘No, it is science that Mala studied.’

Instances like those of B by default spells out only the focused constituent; the constituents struck through in the responses of B can be elided.

The focus particle is used—as in the second turn in the exchange between A, B in (8) below—to further emphasize or endorse the constituent focused earlier with extra stress on the constituent (indicated with upper case letters) in response to a question repeated:

(8) T. A. mala padi-th-athu vingnaanam-aa
Mala study-PST-NMLZ science-Q
S. B. mala igenəgathth-e vidyaavə də
Mala study.PST-NMLZ science Q

‘Is it science that Mala studied?’

T. B. oom ~~mala~~ ~~padi-th-athu~~ vingnaanam
Yes Mala study-PST-NMLZ science
S. B. ow ~~mala~~ ~~iganagathth-e~~ vidyaavə
Yes Mala study.PST-NMLZ science

‘Yes, it is science that Mala studied.’

T. A. vingnaanam-aa
Science-Q
S. A. vidyaavə də
Science Q

‘Is it science that Mala studied?’

T. B. oom VINGNAANAM thaan
Yes science FOC
S. B. ow VIDYAAVƏ thamay
Yes science FOC

‘It is certainly science that Mala studied.’

The constituent that is not elided, i.e., the clefted constituent, in B’s response in (7) and the clefted constituent which co-occurs with the focus particle in B’s response in the second turn in (8) bear exhaustive focus.

An introduction to the structure of the cleft construction is in order. In cleft constructions like (5), there are two crucial notions, namely presupposition and focus posited by Chomsky (1971) (see also Rizzi (1997). Presupposition is a proposition contained in/conveyed by a sentence which is assumed to be true or accepted by the speaker and his/her interlocutor, while focus is the constituent that is semantically highlighted/emphasized by the speaker. The construction in (5) above means that the focus of the sentence is ‘science’, while it conveys the presupposition ‘Mala studied something/some subject’. According to Chomsky (1971: 199), the focused constituent ‘receives the main stress and serves as the point of maximal inflection of the pitch contour’, as shown in the English *it*-cleft of (5) given in (9):

(9) ‘It was SCIENCE that Mala studied.’

As mentioned above, the equative structure of the cleft construction in Tamil and Sinhala (see 6.1) involves two DPs being equated: (i) the nominalized segment without the focused constituent, a.k.a the cleft clause and (ii) the clefted constituent. The former encodes the presupposition, while the latter, the focus. These two DPs are the constituents of the construction which is similar in structure to the nominal predicative construction (3.5.3) in which the copula is null.

The initial DP (the cleft clause) in the Tamil/Dravidian cleft construction ends with the (tensed) nominalized form (see 4.3.8; Sarma 1999, 2003; Jayaseelan 2001, 2008; Jayaseelan and Amritavalli 2005). Gair (1980) refers to the verb form which the Sinhala cleft clause ends with as 'a special, emphatic, tensed form of the verb'. It is glossed as EMPH/E- (Hagstrom 2004, Kishimoto 2005), FOC (Chandralal 2010) in the literature and it is popularly known as *e*-form of the verb, e.g., *dunne* 'gave' because this kind of verb ends with the morpheme *-e*. This *e*-form of the verb is in contrast to the *a*-form of the verb, the finite form, e.g., *dunna* 'gave' because this kind ends with the morpheme *-a*. The term 'emphatic form of the verb' is problematic for a number of reasons. First, in none of the different kinds of grammar is there any reference to the emphatic form of the verb, unlike 'participle', 'finite', and 'infinite'/'non-finite' forms of verbs which perform particular grammatical functions. These verbs do not show any overt marking of emphasis, such as stress or reiteration, because they do not denote emphasis per se. Neither do these verbs convey positive polar emphasis, like the *do* in *I do apologize for any inconvenience caused*. Finally, it is the focused constituent that bears exhaustive focus, and not the verb.

Gair (1980: 29), however, notes that 'although the origin of the Sinhala form is not entirely clear, it did function earlier as a nominal form with a third person masculine/neuter ending…' In his view, it seems likely that this form was derived from the same participial sources as the Sinhala attributive (tensed) verb, just as the corresponding Tamil verb form. Therefore, he concludes that 'the resemblance between the Sinhala and the Dravidian constructions would have been even closer at an earlier stage.' It is plausible to assume that this earlier form is the tensed verbal nominal form (see (1) in table 4.3), e.g. *dunnə ekə* in which *dunnə* is the participle form of the verb 'give'—the form of the verb which occurs in the Sinhala relative clauses—and *ekə* 'one' is the third person (inanimate) neuter ending (see 6.3). This form may have subsequently been reduced to *dunne* (see 6.3 for the description of the process).

Based on the analysis, it is assumed that the (tensed) nominalized verb form which occurs in differnt morphosyntactic phenomena in the two languages is of two kinds: (i) that which co-occurs with copula; and (ii) that which does not co-occur with copula. The constructions in which these two forms occur are given in table 6.1 below (these two forms require further research):

Nominalized form	**Tamil**	**Sinhala**
(i) that which co-occurs with (null) copula	cleft construction (6.1)	cleft construction (6.1)
(ii) that which does not co-occur with (null) copula	———	focus particle construction (6.1–6.2)
	wh-question which questions the subject (2.3.1.6)	all default wh-questions (2.3.1.6)
	constituent polar question which questions the subject (2.3.1.6)	all constituent polar questions (2.3.1.6)
	———	default negative construction (3.4.1)
	all constituent negative constructions (3.6.3)	all constituent negative constructions (3.6.3)
	———	all constituent quotative constructions (4.4.2)

Table 6.1: Distribution of tensed nominalized verbs

Note that the nominalized forms in Sinhala have a wider distribution than those in Tamil because in the former they have been extended to other contexts. This kind of extension is a structural effect of contact-induced restructuring.

The other important feature of the Tamil/Sinhala cleft construction is the occurrence of the null copula. The cleft construction has the equative structure characteristic of the feature encoding nominal predication (3.5.3) in which the copula is null in the two languages. Moreover, like the written form of the different kinds of predication, the cleft construction in the written form of the two languages has an overt copula, as in (10):[92]

---

[92] *weyi* (also *veyi*) and *yæ* (also *ya*) in (10S) are glossed, as in exs. (25), (12) respectively, pp. 255, Gair (1998e).

(10) T. kumar nettu mala-(u)kku kudu-th-athu antha puthakath-ai
Kumar yesterday Mala-DAT give-PST-NMLZ that book-ACC
aakum
be.PRS.3SGN
S. kumar iiyee mala-tə dunn-e emə pothə weyi / yæ
Kumar yesterday Mala-DAT give.PST-NMLZ that book be.3Pers /3Pers

‘It is that book that Kumar gave to Mala yesterday.’

In (10T) *aakum* ‘be/become’ is the third person singular neuter form of the copula (one of the two copulas, see 3.6.4). As for instances like (10S), Gair (1998e: 255) notes that in literary Sinhala, the focused element must co-occur either with ‘the lexical copula in the third person’ *weyi* (see (47s) in 3.6.4; see also 2.3.1.1) or ‘the third person agreement suffix’ *yæ*, a.k.a ‘the predicate marking form’ which marks finite predicates (see 6.4). The details discussed here provide evidence for the fact that the colloquial form of the cleft construction in the two languages uses the equative structure in which the copula is null.

The features of the cleft construction discussed above become clear when it is compared with a movement parallel to the cleft construction in Tamil, which Sarma (1999, 2003) distinguishes as R(ightward)-extraction, as in (11) (cf. 1b):

(11) T. kumar nettu mala-(u)kku kudu-th-aan antha puthakath-ai
Kumar yesterday Mala-DAT give-PST-3SGM that book-ACC
S. kumar iiyee mala-tə dunna ee pothə
Kumar yesterday Mala-DAT give.PST that book

‘Kumar gave that book to Mala yesterday.’

The R-extraction is possible in Sinhala too because of its relatively free-word order, as shown in 2.3.1. Clefting is similar to R-extraction in that both move the constituent to the right past the verb. However, there are differences between the two operations. First, in clefting, the verb of the clause from which the constituent is clefted becomes nominalized, while in R-extraction, the default finite form of the verb occurs. Secondly, in clefting, the constituent moved receives exhaustive focus, while in R-extraction the constituent moved does not receive exhaustive focus (for empirical facts in support of this claim, see Selvanathan 2012).

Discussed thus far in this subsection are the basic properties of the cleft constructions which Tamil and Sinhala share. Note also that the other Dravidian languages also possess these properties (see Jayaseelan and Amritavalli 2005; Krishnamurti 2003).

Sarma (1999, 2003) discusses some syntactic properties of clefting which are illustrated below with examples to show the similarities between the cleft constructions in the two languages. The first property of clefting is that it is possible to cleft entire clauses which are finite in (12a) and infinitival in (12b) respectively:

(12) a. T. prasanth t_i so-nn-athu $[_{CP}$ kamala soothanai pass panni-n-aal endu$]_i$
Prasanth say-PST-NMLZ Kamala exam pass do-PST-3SGF COMP
a. S. prasanth t_i kiww-e $[_{CP}$ kamala vibhage pass kəla kiyəla$]_i$
Prasanth say.PST-NMLZ Kamala exam pass do.PST COMP

‘What Prasanth said was that Kamala passed the exam.’

(12) b. T. vasanth t_i ninai-th-athu $[_{CP}$ PRO antha paper-ai vaasikk-a$]_i$
Vasanth think-PST-NMLZ that paper-ACC read-INF
b. S. vasanth t_i hithuw-e $[_{CP}$ PRO ee paththərəyə kiyəvandə$]_i$
Vasanth think.PST-NMLZ that paper read.INF

‘What Vasanth thought was that he would read the newspaper.’

Note that in (12a/b) the embedded clause in (12a) and the infinitival clause in (12b) are clefted in that the nominalized form of the verb ‘think’ is used, and the clefted clauses occur clause finally to the right of the nominalized form of the verb. Clefting can be local in that the DP of an embedded clause can be clefted within that clause, while the matrix finite verb remains finite, as in (13):

(13) T. avan $[_{CP}$ rani t_i eesi-n-athu mary-ai$_i$ endu] ninai-th-aan
He Rani scold-PST-NMLZ Mary-ACC COMP think-PST-3SGM
S. eyaa $[_{CP}$ rani t_i bann-e mary-tə$_i$ kiyəla] hithuwa
He Rani scold.PST-NMLZ Mary-DAT COMP think.PST

‘He thought that it was Mary that Rani scolded.’

Clefting can also be long distance, that is, from within an embedded clause to the matrix clause, as in (14):

(14) T. avan $[_{CP}$ rani t_i eesi-n-aal endu] ninai-th-athu mary-ai$_i$
He Rani scold-PST-3SGF COMP think-PST-NMLZ Mary-ACC
S. eyaa $[_{CP}$ rani t_i banna kiyəla] hithuw-e mary-tə$_i$
He Rani scold.PST COMP think.PST-NMLZ Mary-DAT

‘What he thought was that it was Mary that Rani scolded.’

It is possible to cleft a constituent only once. The same constituent cannot be clefted sequentially, crossing two verbs as in (15a). However, it is possible to cleft within different clauses, as in (15b) in which one constituent ‘Kumari’ is clefted within an embedded finite CP and then the entire embedded CP is clefted out to the matrix clause.

(15) a. T. * luxman [[$_{CP}$ ravi t_i adi-th-athu t_i endu] ninai-th-athu]
Luxman Ravi hit-PST-NMLZ COMP think-PST-NMLZ
kumari-kku$_i$
Kumari-DAT
S. * luxman [[$_{CP}$ ravi t_i gahuw-e t_i kiyəla] hithuw-e]
Luxman Ravi hit.PST-NMLZ COMP think.PST-NMLZ
kumari-tə$_i$
Kumari-DAT

Intended meaning: ‘Luxman thought that it was Kumari that Ravi hit.’

(15) b. T. luxman t_i ninai-th-athu [$_{CP}$ ravi t_j adi-th-athu
Luxman think-PST-NMLZ Ravi hit-PST-NMLZ
kumari$_j$-kku endu]$_i$
Kumari-DAT COMP
S. luxman t_i hithuw-e [$_{CP}$ ravi t_j gahuw-e
Luxman think.PST-NMLZ Ravi hit.PST-NMLZ
kumari$_j$-tə kiyəla]$_i$
Kumari-DAT COMP

‘What Luxman thought was that it was kumari that Ravi hit.’

The reason for the ungrammaticality of (15a) must be that once the constituent is clefted within the embedded clause, the remnant structure becomes a complex NP, headed by a nominalized verb; the second extraction out of this NP (the embedded CP) to the matrix clause will be illicit owing to subjacency violation or violation of the Complex NP constraint (CNPC); hence, (15aT,S) are infelicitous. For the same reason, it is not possible to extract constituents out of relative clauses, as in (16b); the unclefted form is (16a):

(16) a T. aval [$_{NP}$ [$_{TP}$ padma-(u)kku udavi seith-a] teacher-ai]
She Padma-DAT help do.PTCP-RELAT teacher-ACC
santhi-th-aal
meet-PST-3SGF
S. eyaa [$_{NP}$ [$_{NP}$ padma-tə udawu kərəp-u] teacher-wə]
She Padama-DAT help do.PTCP-RELAT teacher-ACC
hambuna
meet.PST

‘She met the teacher who helped Padma.’

(16) b. T. *aval [$_{NP}$ [$_{TP}$ padma-(u)kku udavi seith-a] t_i]]
She Padma-DAT help do.PTCP-RELAT
santhi-th-athu teacher-ai$_i$
meet- PST-NMLZ teacher-ACC
S. *eyaa [$_{NP}$ [$_{TP}$ padma-tə udawu kərəp-u] t_i]]
She Padama-DAT help do.PTCP-RELAT
hambun-e teacher-wə$_i$
meet.PST-NMLZ teacher-ACC

Intended meaning: 'It was the teacher who helped Padma that she met.'
Or 'Who she met was the teacher that helped Padma.'

The ungrammaticality of (16b) results from CNPC violation; the relative clause is a complex NP in itself, and clefting 'Padma' out of this complex NP will be illicit. The ungrammaticality of (15a) and (16b) demonstrates that clefting in the two languages strictly adheres to island constraints.

Finally the clefted constituent is related to the clause in which it originates. The θ-role and the case of the clefted constituent is determined by the verb of the clause to which it belongs. Moreover, the binding relations of the constituent with the verb remain unaltered, as in (17):

(17) T. mohan$_i$ t_j veru-th-athu than$_i$-da mahan-ai$_j$
Mohan hate-PST-NMLZ self-GEN son-ACC
S. mohan$_i$ t_j vairəyə kəre-e] thaman$_i$-ge putha-tə$_j$
Mohan hate do.PST-NMLZ self-GEN son-DAT

'It was his own son that Mohan hated.'

In summary, clefting in the two languages is a syntactic process with semantic implications. Clefting can move clauses, both finite and non-finite; while it can be local or long distance, it obeys island constraints, which shows that clefting in the two languages involves movement. The verb which the moved constituent is an argument of assigns the θ-role and determines its case and binding relations. What is important from the point of view of this research is that these languages, which are genetically unrelated, share this kind of clefting. Their properties, as the discussion above shows, are the same.

6.2.2 Syntactic analyses of cleft constructions

This section examines the syntactic operations involved in the cleft construction in the two languages under study. Two studies, namely Jayaseelan and Amritavalli (2005; J&A hereafter) and Sarma (1999, 2003) propose operations related to clefting that can be used to explicate clefting in Tamil and Sinhala; the former analyzes the movement involved in clefting in Dravidian languages, particularly, Malayalam as leftward, while the latter analyzes clefting in Tamil as rightward. In the rest of this section, first the operations proposed in these works are briefly discussed and then, these operations are adapted so as to explain Tamil/Sinhala clefting.

Based on the proposals made by Jayaseelan (2001; see also 2008 and 2010), J&A analyze the short (local) and long distance clefting in Malayalam and Kannada, two Dravidian languages. Their proposals can adequately be used to analyze the syntactic operations involved in clefting in Tamil and Sinhala (see below). In order to explain the clefting in Wh-questions in Malayalam, Jayaseelan (2001) makes two claims: a) there is an IP internal focus phrase immediately above vP/VP into which wh-phrases and other focused phrases are moved; and b) there are topic positions above this focus position but below the IP. Assuming the universal order of Specifier-Head-Complement in languages postulated by Kayne (1994), Jayaseelan claims that though the two kinds of languages, head initial and head final, have a FocP above VP, the position at which it shows up is different in these two kinds: in an SVO language, it shows up at the end of the clause, i.e., the right periphery of VP, while in an SOV language, it shows up to the immediate left of V.

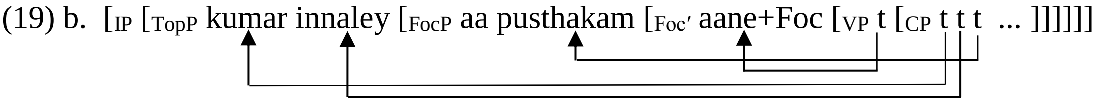
Following these two proposals, J&A explain the short distance cleft construction in Malayalam, that is, clefting a constituent locally within a single clause, as shown in (18), the equivalent of Tamil and Sinhala (1b):

(18) kumar (**aane**) innaley (**aane**) malakke (**aane**) aa pusthakam (**aane**)
Kumar COP yesterday COP Mala-DAT COP that book (is)
kodu-tt-atea
give-PST-NMLZ
'It was/is Kumar that gave that book to Mala yesterday.'
'It was/is yesterday that Kumar gave that book to Mala.'
'It was/is to Mala that Kumar gave that book yesterday.'
'It was/is that book that Kumar gave to Mala yesterday.'

According to J&A, in the Malayalam cleft construction, *aane* is copula which is overt—unlike the covert one in the Tamil/Sinhala cleft construction—occurring adjacent to the focused constituent, seemingly floating into the cleft clause that expresses the presupposition. This kind of cleft construction is relatively similar to the focus particle construction in Tamil and Sinhala (1c) in that the copula occurs immediately after the focused constituent in Malayalam, while the focus particle occurs immediately after the focused constituent in Tamil and Sinhala.

J&A make two proposals to explicate the syntactic operations involved in the cleft construction in the Dravidian languages in terms of the cleft constructions in English and Malayalam. Firstly, clefting in the Dravidian languages involves two syntactic operations: (i) the copula ‘be’ which occurs clause-finally takes a clausal complement; and (ii) a constituent (to be focused) from within the clausal complement is moved into the IP internal Spec FocP above the VP (claim (a) of Jayaseelan 2001). For instance, the structure of (19a) would be (19b):

(19) a. $[_{CP}[_{DP}$ kumar innaley mala-kke kodu-tt-atea] $[_{DP}$ aa pusthakam] **aane**]
Kumar yesterday Mala-DAT give-PST-NMLZ that book COP
‘It is that book that Kumar gave to Mala yesterday.’


(19) b. $[_{IP}$ $[_{TopP}$ kumar innaley $[_{FocP}$ aa pusthakam $[_{Foc'}$ aane+Foc $[_{VP}$ t $[_{CP}$ t t t ...]]]]]]

According to J&A, in English, the copula raises to I (presumably adjoining to Foc^0 as an intermediate step) and the expletive ‘it’ is merged with Spec IP. In Malayalam, the copula *aane* raises to Foc^0 but not to I^0. The second proposal stipulates that the apparent effect of the focused constituent together with the copula floating in the cleft clause (see (18) above) is created by the movement of elements from within the cleft clause to Spec of multiple topic positions above the focus position, as shown in (19b):[93]

Having made these two proposals with the aim of explicating the syntactic operations involved in the short distance clefting, J&A claim that long distance clefting in the Dravidian languages obligatorily employs relativization in extracting the constituent to be focused from the cleft clause (see 6.3). They note that there are two cleft clauses, one is an IP and another is a CP and in the long distance cleft construction, only the latter can be employed (fn. 18, p.160, ibid.). The short distance cleft construction of the type (19a or its Tamil/Sinhala equivalent 1b) does not employ relativization because, as shown above, it is an IP, and it does not involve the C(omplementizer)-system. The short distance cleft construction moves constituents—

---

[93] The ellipsis in the CP (19b) stands for the nominalized form of the verb ‘give’ (not shown owing to constraints of space) which is also moved to Spec of TopP above FocP.

the one to be clefted and those that are topicalized—to its left (IP-internally), directly from the cleft clause into the matrix clause.[94]

These proposals need to be slightly adapted to explain the syntactic operations involved in short and long distance clefting in Tamil and Sinhala for two reasons: 1) the copula in the cleft construction in the two languages is always null unlike the overt copula in Malayalam; and 2) the cleft clause in Tamil and Sinhala obligatorily occurs to the left of the focused constituent unlike that in the Malayalam construction. Following J&A's two proposals and Jayaseelan's (2001, 2008, 2010) proposals together with Kayne's underlying universal order of Specifier-Head-complement, the following syntactic operations are assumed to take place in the short distance cleft construction in Tamil and Sinhala (note that IP is changed into TP in consistent with the minimalist framework):

(i) the copula which is null in Tamil and Sinhala takes a nominalized (tensed) clause as its complement which ends with the nominalizer suffix (*-athu* in Tamil and *-e* in Sinhala) in accordance with Kayne's underlying universal order (J&A's operation (i)).[95] The tensed nominalized clause becomes fully transparent to extraction;
(ii) the constituent to be focused is moved to the TP internal Spec of FocP (J&A's operation (ii)) in which position the constituent moved is interpreted as being exhaustively focused;
(iii) the null copula is moved into Foc head;
(iv) to derive the surface SOV order, either the complex NP as a whole or its constituents individually move to Spec of TopPs; J&A propose the latter of the two; here the former is assumed.

(20b) below shows operations (i)–(iv) which take place in the cleft construction in Tamil and Sinhala given in (1b) repeated as (20a). Under this analysis, the entire NP containing the remnant CP moves to a position preceding the focused constituent. It is

---

[94] Selvanathan (2012) presents a similar view, arguing that the Tamil short distance cleft involves only TP-internal movement.

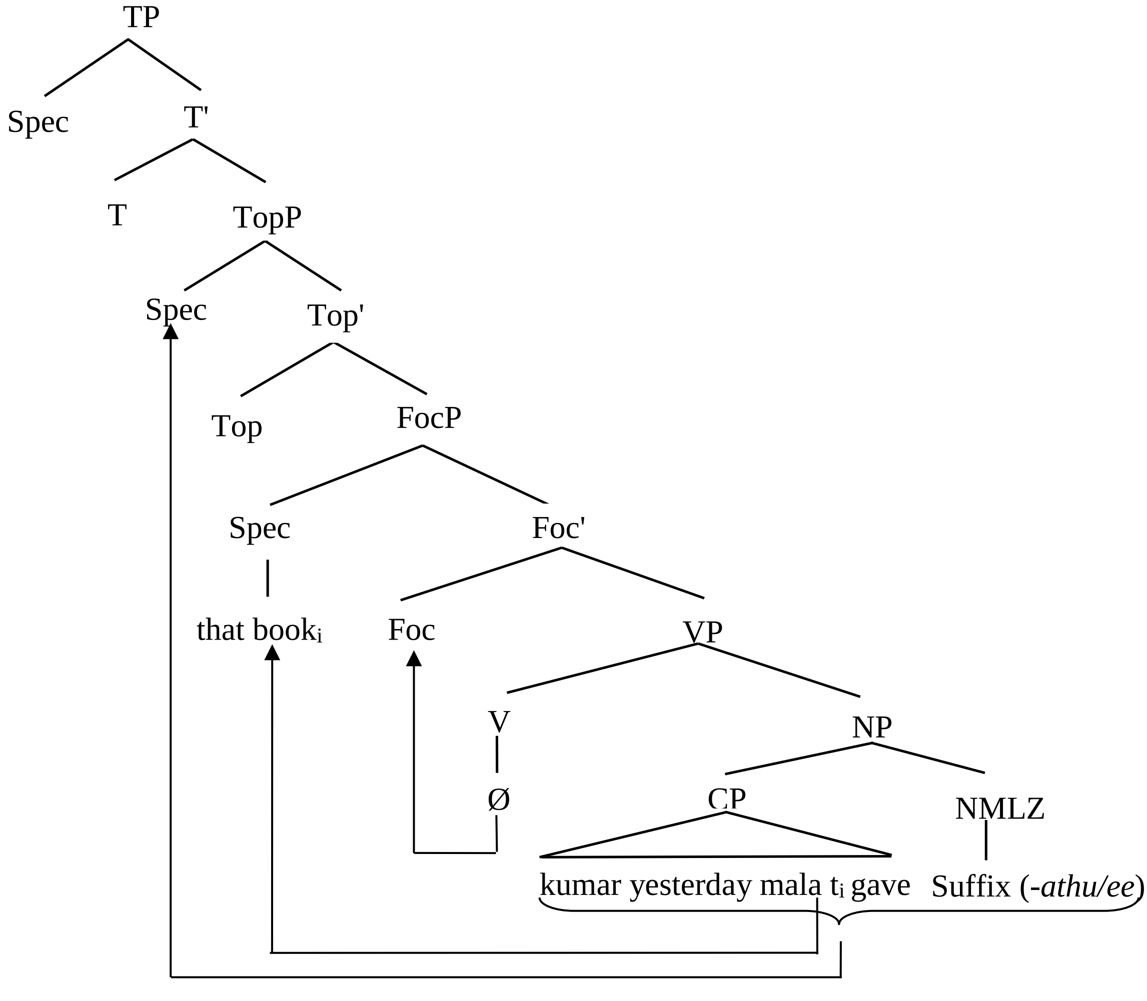
[95] According to Jayaseelan (2001a), the nominalized clause with the suffix *-athu* in Dravidian historically derived from the demonstrative stem *a-* which belongs to the category of determiner and the third person singular neuter suffix *-thu* becomes a tensed clause fully transparent to extraction. Note that a similar process may be assumed to take place in Sinhala because the nominalized clause ends with the demonstrative (distal/non situational) stem *e*(*e*)- which too belongs to the category of determiner.

proposed here that this is the subject position, which is also a topic position, since the default information-function of the subject is topic.

(20) a. T. kumar nettu mala-(u)kku kudu-th-athu antha puthakath-ai Ø
Kumar yesterday Mala-DAT give-PST-NMLZ that book-ACC (COP)
S. kumar iiyee mala-tə dunn-e ee pothə Ø
Kumar yesterday Mala-DAT give.PST-NMLZ that book (COP)

'It is that book that Kumar gave to Mala yesterday.'

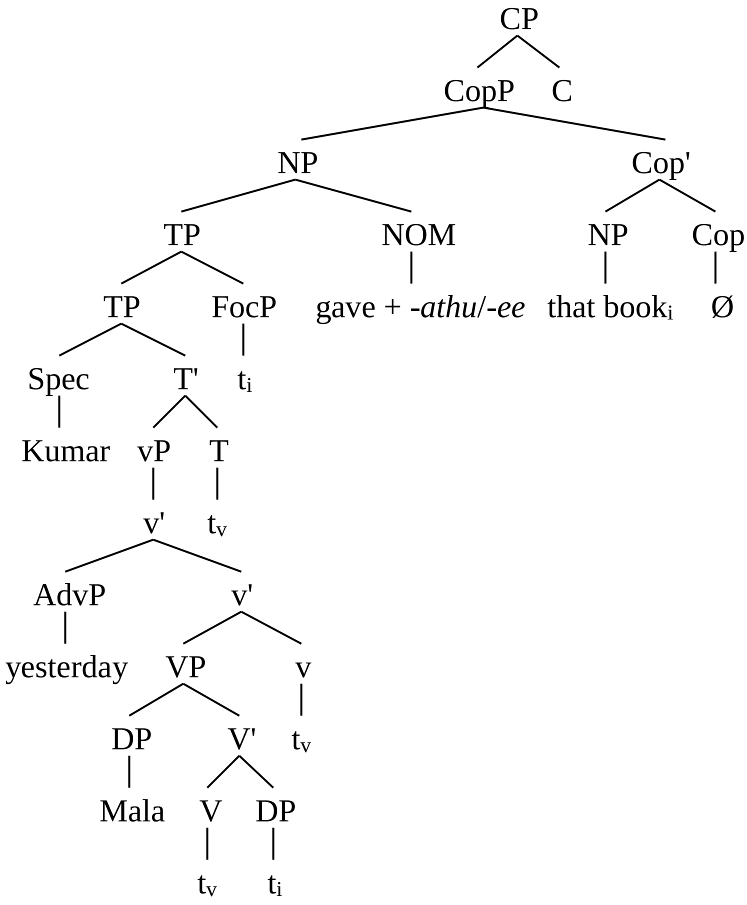
(20) b.



The nominalization can now be seen as linked to the topicalization: a VP cannot be a subject (or a topic), an NP/DP obviously can.

The movements Sarma (1999, 2003) proposes to explain the syntax of clefting in Tamil are diametrically opposite to those of J&A. As shown above, the operation(s) involved in clefting proposed by J&A are leftward, whereas the operations proposed by Sarma are rightward. Sarma (1999: 88–89) points out that the fact that in Tamil, binding, case and theta properties of the focused constituent are determined by its base position—in effect, by the verb that the constituent is originally an argument of—indicates that clefting is a movement to a non-argument position, i.e., non-case assigning position; hence, it is an A-bar movement. Claiming that the relevant movement in clefting is NP movement in a copular structure, she assumes that the NP raises to FocP to the right of TP. Part of the process of clefting is the nominalization of the TP by the addition of what she calls a NOM head, which results in the creation of a nominal predicative construction of the type NP NP (COP). Further, she posits an empty copula head which takes the nominalized clause and the clefted NP as its complements. She also assumes that the head of CP selects as its complement the Cop(ula)P whose head is empty. The cleft construction of (20a), following what Sarma proposes, is shown in (20c):

(20) c.



The leftward and rightward movements proposed in the two works discussed above adequately explicate the operations involved in the Tamil/Sinhala cleft construction. In effect, the two analyses are notational variants, making the same predictions with different theoretical assumptions; note that the constituents and the structure in the two analyses are the same. J&A's analysis which is based on a more restrictive theory than Sarma's analysis is used in this dissertation because it neatly explains the syntactic operations involved in the Tamil/Sinhala cleft construction including relativization which takes place in the long distance cleft construction as discussed in 6.3 below.

6.3 The role of relativization in long distance clefting

Long distance clefting is similar to relativization in that the two phenomena extract a constituent from its base position and move it to the C-domain. As mentioned in 6.2, relativization plays a significant role in long distance clefting in Dravidian and Sinhala. The two languages share the two phenomena—the prenominal relative clause (2.2.10) which uses the gap strategy (4.2.1 and 4.2.2), and long distance clefting ((14) in 6.2.1)—because these phenomena in Sinhala have been modelled on the corresponding Tamil phenomena. It is reasonable to assume that the two languages share the same strategy of relativization in long distance clefting in Dravidian which J&A propose. Following J&A, this section analyzes the role of relativization in long distance clefting in the two languages.

The prenominal relative clause construction and the cleft construction are almost similar in the two languages. An important difference between them, though, is that the relative clause in the two languages is non-finite, whereas the cleft construction is finite. As mentioned above, the prenominal relative clause in Tamil and Sinhala is non-finite, which is attested typologically (Kayne 1994, Jayaseelan 2014). It is also important to note that the relativizer suffix *-a* in Tamil and *-ə* and *-u* in Sinhala is at the right edge of the relative clause, as in (21):

(21) T. [kumar nettu mala-(u)kku _____ kuduth-a] puthakam
Kumar yesterday Mala-DAT give.PTCP-RELAT book
S. [kumar iiyee mala-tə _____ dunn-ə] pothə
Kumar yesterday Mala-DAT give.PTCP-RELAT book

'The book that Kumar gave to Mala yesterday'

In (21), the relative clause which functions as the modifier of the head noun occurs pre-nominally, linearly to the left of the head noun. As mentioned in 4.2.1, the gaps in the two constructions indicate the position from which the constituent has been moved to the C-domain. Evidence for movement comes from island effects. Compare (22a) with (22b) (adapted from exs. (16a,b) p. 146, J&A); both are long distance relative clauses:

(22) a. T. [[kumar nettu mala-(u)kku ________ kudu-thth-aan endu]
Kumar yesterday Mala-DAT ________ give-PST-3SGM COMP
nee ninaikir-a] puthakam
you think.PTCP-RELAT book
S. [[kumar iiyee mala-tə ________ dunna kiyəla]
Kumar yesterday Mala-DAT ________ give.PST COMP
oyaa hithən-ə] pothə
you think.PTCP-RELAT book

'The book that you think that Kumar gave to Mala yesterday'

(22) b. T. *[[kumar nettu mala-(u)kku ________ kudu-thth-aan
Kumar yesterday Mala-DAT ________ give-PST-3SGM
end]-a karuththu nee oththukond-a] puthakam
COMP-RELAT claim you accept.PTCP-RELAT book
S. *[[kumar iiyee mala-tə ________ dunna kiyəp]-u
Kumar yesterday Mala-DAT________ give.PST COMP-RELAT
mathəyə oyaa piligathth-ə] pothə
claim you accept.PTCP-RELAT book

'The book that you accepted the claim that Kumar gave to Mala yesterday'

As shown in the cleft construction (see (15a) and (16b) above), (22b) is infelicitous because of a subjacency/CNPC violation.

The gap in the relative clause, e.g., (23) below as Jayaseelan (2014) claims, is the trace of the relative head which has been raised successive cyclically. Following Kayne's (1994) raising analysis, a determiner selects a CP as a complement as shown in (23):

(23) [$_{\text{DP}}$ the [$_{\text{CP}}$ ________ that [IP Kumar gave Mala book]

Note that the CP which is headed by *that* contains the relative clause TP. Jayaseelan (2014) notes that the relativizer -*a* (22aT) in Dravidian which is the demonstrative stem (4.3.1) cannot correspond to the determiner *the* in English because the definite article in Dravidian is null. He analyzes the relativizer -*a* as a reduced form of the distal demonstrative which can correspond to *that* in Kayne's (1994) analysis, occupying a position comparable to that of the complementizer of the embedded CP.

Dravidian languages also have 'headless relative clauses' in which the head NP is replaced by 'a pronominal suffix' added to the relative participle. This pronominal suffix is ‘appropriate to the replaced NP in number and gender, from the third person demonstratives …’ (Krishnamurti 2003: 446), as in, for example, (24bT) below; (24aT) is the relative clause with the head noun. J&A note that when the head noun is null as in free (or headless) relatives, the common relativizer -*a* in the Dravidian languages shows agreement with the missing head. Sinhala uses the generic pronoun ‘one’ instead of the head, unlike Tamil. The agreement of the relativizer with the missing head is shown in (24b,d,f) and the relative clauses with the head noun is shown in (24a,c,e) (adapted from exs. (17a,b)–(20a,b), p. 147, J&A):[96]

(24) a. T. [________ vanth-a] manushan/manushi
Come.PTCP-RELAT man/woman
S. [________ aaw-ə] miniha/gaani
Come.PTCP-RELAT man/woman

‘The man/woman who came’

(24) b. T. ________ vanth-a-van/val (human: man/woman)
Come.PTCP-RELAT-3SGM/3SGF
S. ________ aaw-ə ekkena (human)
Come.PTCP-RELAT one

‘The one/person who came’

(24) c. T. [________ vanth-a] naai
Come.PTCP-RELAT dog
S. [________ aaw-ə] balla
Come.PTCP-RELAT dog

‘The dog which came’

(24) d. T. ________ vanth-a-thu (NON-HUM)
Come.PTCP-RELAT-3SGN
S. ________ aaw-ə ekaa (NON-HUM)
Come.PTCP-RELAT one

‘The one which came’

(24) e. T. [________ vanth-a] lorry
Come.PTCP-RELAT lorry
S. [________ aaw-ə] lorrya
Come.PTCP-RELAT lorry

‘The lorry which came’

---

[96] Owing to constraints of space, the plural forms of these examples are not given (see 25b).

(24) f. T. ________ vanth-a-thu (INAN)
Come.PTCP-RELAT-3SGN
S. ________ aaw-ə ekə (INAN)
Come.PTCP-RELAT one

'The one which came'

In the Tamil relative clauses given in (24b,d,fT), the head nouns are covert and it is the agreement on the verb that indicates the person (third: invariant), number (singular/plural) and gender (masculine/feminine/neuter) of the head noun of the relative clause. In the Sinhala relative clauses given in (24b,d,fS), the head nouns are present in that the actual head noun is replaced by indefinite (generic) pronouns. The possible reason for the use of the generic pronoun as head in Sinhala is that it does not have SV agreement. The difference between the Tamil relative clauses (24b,d,fT) and their Sinhala counterparts (24b,d,fS) is that in the former, the head of the relative clause is null, while in the latter, the generic pronoun shows up as the head instead of the actual head in (24a,c,eS). It is significant to note that (24fS) is the default tensed verbal nominal form in Sinhala *aawə ekə* (see (1) in table 4.3) which is assumed to be 'the earlier nominal form with masculine/neuter ending,' as Gair (1980) argues (see 6.2.1).

The endings of verbs in (24b,d,fT) which are marked for agreement are syncretic with the third person pronouns, *avan* 'he', *aval* 'she' and *athu* 'it' (see table 3.4). It is crucial to note that the relativizer *-a* in Tamil, the distal demonstrative stem *a-*, combines with the third person bound pronominal base *-van* (masculine), *-val* (feminine) and *-thu* (neuter) (see 4.3.1) to mark the agreement features of the missing head in (24b,d,fT). In the Sinhala relative clauses (24b,d,fS), the third person generic pronouns —human *ekkena* 's/he', non-human *ekaa* 'it' and inanimate *ekə* 'it'—which follow the verbs stand for the missing head. The two relativizers *-ə* and *-u* in Sinhala are different from the nominalizer suffix *-e* which is the distal demonstrative stem *e*(*e*)-.

It is also important to note that the relative clauses (24bT,S) convey generic inclusive pronoun 'one' (human) meaning in that the singular marking *-avan/-aval* (Tamil) or *ekkena* (Sinhala) can co-occur with any singular human pronoun (first/second/third person), as shown in (25a). If the marking is plural like *-avarkal* in Tamil or *ayə* in Sinhala, it can co-occur with any plural human pronoun (first/second/third person), as shown in (25b):

(25) a. T. ________ vanth-a-van/val naan/nee/avan/aval (HUM)
Come.PTCP-RELAT-3SGM/3SGF I/you(S)/he/she
S. ________ aaw-ə/aap-u ekkena mamə/oyaa/eyaa (HUM)
Come.PTCP-RELAT one I/you(S)/s/he

‘The one/person who came is I/you(S)/he/she.’

(25) b. T. ________vanth-a-varkal naangal/neengal/avarkal (HUM)
Come.PTCP-RELAT-3PL we/you(PL)/they
S. ________ aaw-ə/aap-u ayə api/oyaala/eyaala (HUM)
Come.PTCP-RELAT people we/you(PL)/they

‘The people who came are we/you(PL)/they.’

Instances like (25a,b) which are frequently used in the two languages are similar to the cleft construction in that the subject extracted can be of any person (human) either in singular or plural form. Note that the gap strategy is used to extract the constituent in these constructions.

Given that the gap strategy is used in clefting in the two languages, when a clause is nominalized as a result of clefting, the relative proform *-a* in Tamil combines with the third person singular neuter bound pronominal base *-thu* ‘it’ to form the nominalized ending *-athu*—in effect the tensed verbal nominal form in (1) in table 4.3—occurs at the end of the nominal clause, as shown in (26a–cT) below. Note that as J&A (148) point out, in Dravidian *-athu* in the cleft clause is invariant in that it ‘no longer counts as agreement’. Therefore, it can co-occur with a clefted constituent of any person, number or gender. A similar process is apparent in Sinhala in which the relativizers *-ə* or *-u* merge with the third person proforms *ekkena*/*ekaa/ekə*, and the bound pronominal base *-kena/-kaa/-kə* gets elided and the remaining *e-* becomes the nominalizer suffix *-e* which thereby makes the finite verb the (tensed) nominalized form that occurs at the end of the cleft clause. This, in principle, accounts for the occurrence of the tensed nominalized verb *aawe* in (26a–cS):

(26) a. T. vanth-a + thu manushan → va-nth-athu manushan
Come.PTCP-RELAT 3SGN man come-PST-NMLZ man
S. aaw~~-ə~~ + e~~kkena~~ miniha → aawe miniha
Come.PTCP-RELAT 3SG(HUM) man come.PST-NMLZ man

‘It is the man that came.’

(26) b. T. vanth-a + thu naai → va-nth-athu naai
Come.PTCP-RELAT 3SGN dog come-PST-NMLZ dog
S. aaw~~-ə~~ + e~~kaa~~ balla → aawe balla
Come.PTCP-RELAT 3SG(NON-HUM) dog come.PST-NMLZ dog

‘It is the dog that came.’

(26) c. T. vanth-a + thu lorry → va-nth-athu lorry
Come.PTCP-RELAT 3SGN lorry come-PST-NMLZ lorry
S. aaw~~-ə~~ + e~~kə~~ lorrya → aaw-e lorrya
Come.PTCP-RELAT 3SG(INAN) lorry come.PST-NMLZ lorry

‘It is the lorry that came.’

Just as the nominalized ending *-athu* (the invariant third person singular neuter ending) in Tamil, the nominal ending *-e* (26a–c) in Sinhala too has no agreement features; hence, it too can co-occur with a clefted constituent of any person, number or gender, as shown in (27S), just as the Tamil nominalized form in (27T):

(27) T. oodi-n-athu naan(gal); nee(ngal); avan/aval/avarkal/athu/avai(kal)
Run-PST-NMLZ I/we you(PL) he/she/they(HUM)/it/they(NON-HUM)
S. diww-e mamə/api; oyaa(la); eyaa(la)/uu/ung
Run.PST-NMLZ I/we you(PL) s/he/they(HUM)/it/they(NON-HUM)

‘It is I/we/you(PL)/s/he/they/it/they that ran.’

The nominalized form of the verb in Tamil (27bT) ends with the third person singular neuter ending *athu*, whereas the nominalized form of the verb in Sinhala (27bS) ends with the third person pronominal (distal demonstrative) stem *-e*, resulting from the elision process shown in (26a–cS). At the end of the sentence, by hypothesis, is a null copula (see (20)).

The above discussion shows that relativization plays an important role in long distance clefting in that the gap strategy involved in the former is used to extract the constituent to be focused in the latter. In addition, in the two languages, relativization generates the invariant nominalized verb form which occurs at the end of the cleft clause. What is important to note is that this kind of nominalization discussed in this section, as J&A argue, occurs only in long distance clefting in Dravidian (6.3).

This seems to hold true for Sinhala too. J&A add that although in Dravidian the short distance clefting does not use relaivization, ‘it retains the relative clause morphology as a historical residue.’ How the nominalized verb form is generated in the Sinhala monoclausal cleft construction cannot be verified. What may perhaps be likely, though, is that the nominalized form is the reduced form of the earlier nominal form with the third person neuter ending (assumed to be the verbal nominal (1) in table 4.3) mentioned in 6.2.1. As shown in 6.2.2, the two languages, however, do not use relativization in short distance clefting in that clefting does not move the constituent focused to the C-domain. Instead, it moves it to the Spec FocP below TP or above vP/VP (6.2.2). There are, therefore, two kinds of clefting in the two languages: the short distance clefting which does not involve relativization and the long distance clefting which involves relativization. The facts discussed in this section show that Sinhala uses almost the same strategies for the extraction of the constituent in long distance clefting, just as Tamil does, which provides evidence for the claim that the strategy used in the former has been modelled on that of the latter. In 6.4 the role of clefting in the restructuring of the Sinhala agreement system is discussed.

6.4 Diachronic analysis of the agreement system in Sinhala

Finiteness plays an integral role in syntax in that it determines such features as tense, agreement (conveying person, number and gender where necessary), mood and assignment of case to the argument DP(s) of verbs. As mentioned above, one of the important differences between the morphosyntax of the two languages is that Sinhala lacks SV agreement, while Tamil has it (3.6.1.2), as shown in (28):

(28) T. amma kada-il saamaan vaang-in-aa
mother shop-LOC things buy-PST-3SGF(HON)
S. amma kad-en badu gaththa
mother shop-ABL things buy.PST

'Mother bought things in the shop.'

Sinhala seems to have had an agreement system which it has lost; in effect, it has restructured its agreement system over time. As mentioned towards the end of 3.6.1.2, Sinhala has several marked finite verbs with agreement which are believed to be the vestiges of the agreement system Sinhala once had. These marked forms need a subject of a particular person and/or number kind, as in (29–33) below ((29–32) adapted from exs. (2)–(5); and (33) is ex. (6), p. 152, Paolillo 1994):[97]

---

[97] For consistency, these sentences and those below from classical Sinhala are glossed as those in Paolillo (1994).

(29) S. kumar kərəpu-de
Kumar do-PAST-CONC
'Let Kumar do (whatever he likes I am not bothered).'
(expressing indifference; third person singular or plural)

(30) S. api raa-tə bath ka-mu
We night-to rice eat-HORT
'Let us have rice tonight.'
(hortative verb suggesting/encouraging; first person plural)

(31) S. leela duwa-i
Leela run-INVOL
'Leela may (probably) run, (seeing this chaos).'
(involitive optative verb expressing likelihood/assumption; third person singular or plural)

(32) S. mamə salli de-nnang
I money give-VOL
'I'll give money.'
(volitive optative verb expressing promise; first person)

(33) S. uu mata kaa-pi
It I-DAT eat-PFV
'(Darn!) it (up and) bit me!'
(expressing contempt/distress/surprise; third person inanimate)

These forms are marked not only in grammatical form but also in the kinds of meaning they express, with the speaker's attitude towards the event encoded in the verb (given in parentheses below each sentence). Paolillo (1994: 154) claims that the absence of verbal agreement and the existence of finite forms with agreement, indexing communicative attitudes are the products of the naturalization of a Dravidian focus cleft and the expansion of its role in the grammar of Sinhala. In this section, based on the arguments that Paolillo (1994) develops, I show that the Tamil cleft construction which Sinhala calqued led to the restructuring of the Sinhala agreement system, resulting in unmarked finite verbs which lack agreement and the limited number of marked finite verbs with agreement, as those in (29–33), which express communicative attitudes.

Paolillo (1994: 154–155) observes that since classical Sinhala has verbs with agreement, but does not have attitudinal verb forms as those in (29–33), these forms of the present day Sinhala evolved from classical verbs as verbs with SV agreement, expressing communicative attitudes. This development from agreement markers to attitude-marking forms, according to Paolillo (1994), exemplifies Traugott's (1987, 1988) 'tendency III' paraphrased (ibid. 155) given in (34):

(34) In grammaticalization, meanings tend to become more situated in the speaker's mental attitude toward the situation.

On the basis of (34), Paolillo (1994: 155) hypothesizes that a person-number-gender (PNG) agreeing form conventionally implicated a particular communicative attitude in certain tense/aspect categories and argues that this attitude is then reanalyzed as the actual content of the form with the original PNG content perhaps being retained as a conventional implicature. Thus, the content and its conventional implicature exchange their status in the reanalysis.

The present Sinhala agreement system with its unmarked (28S) and marked (29–33) finite forms developed the way it did owing to the co-development of finiteness and focus marking in Sinhala. Following (Paolillo 1994), the facts related to finiteness and the two kinds of focus given in (i–v) below show that in Sinhala the focus is associated with communicative attitudes syntactically and semantically.

(i) The nominalized form of the verb in either of the focus constructions (1b/c) in Sinhala is similar to the unmarked finite form of the verb in sentences like (28S) in that the former has no SV agreement and it expresses a neutral communicative attitude; it is different from the latter in that the clause which contains the nominalized form must have a focused constituent, whereas the clause which contains the finite form cannot have a focused constituent.

(ii) The focus, question, negative and quotative markers in Sinhala which occur clause finally or clause internally adjacent to the constituent are considered focus marking forms (see Paolillo 1994) which like the marked finite forms express communicative attitudes. Note that all these focus marking constructions are almost identical to the corresponding Tamil constructions.

(iii) These focus markers are different from the endings of the marked forms which express communicative attitudes: the focus markers are syntactically and morphologically free, while these endings are bound to the verbs.

(iv) The unmarked finite forms participate in the focusing alternation, whereas the marked finite forms which express communicative attitudes have inherent focus of the attitude they express and, therefore, do not participate in focusing alternation, which accounts for the ungrammaticality of (35):

(35) a. S. *kumarmay kərəpu-de
Kumar-FOC do-PAST-CONC
(35) b. S. *kumar-de kəre
Kumar-CONC do-PAST-EMPH

Intended meaning: ‘Let it be Kumar that does (I am not bothered).’

Since the clauses which have marked finite verbs (29–33) cannot have focused constituents, these verbs do not have the corresponding EMPH(atic) (nominalized) form (35bS).

(v) The predicator, that is, the verb of the clause that contains the unmarked and marked forms, is the unmarked focus of the sentence. [98]

These facts show that the two kinds of focus interact with the finite forms, both unmarked and marked.

Further explication of the changes that the Sinhala agreement system has undergone owing to the interaction of focus marking with finiteness requires analysis of classical Sinhala. The attitude implicatures of the agreeing forms which occur in classical Sinhala are the antecedents of the modern Sinhala marked finite forms; (36) can be considered as the direct antecedent of the hortative given in (30), and (37) as the antecedent of the volitive optative given in (32). ((36)–(42) below are from *Gurulugomi* 1967, a classical Sinhala text, cited in Paolillo 1994):

(36) S. eva maha terun sarana yamha
thus great elder refuge go-PRES-1PL
‘Thus, let us take refuge under the great elder.’ (lit. ‘We take refuge…)

(37) S. Wahanse ohu kara gos kiyam da?
Lord him near go-PPLE tell-PRES-1SG Q
‘My lord, shall I go and tell him?’ (lit. ‘Do I go and tell him?’)

The question marker *da* and the reportative (quotative) *la*, the antecedents of modern Sinhala *də* and *lu*, are not ‘focus marking forms’ but are forms which mark predicates in classical Sinhala. One of the important features of the agreement system in classical Sinhala is that the third person singular finite predicates are not normally marked by an overt agreement morpheme like predicates with subjects of other person-number categories are. However, in the absence of the overt agreement morpheme or other predicate marking form, *də or la*, a predicate marking form *yæ*—whose function is to

---

[98] Following Gair and Paolillo (1988), Paolillo (1994: fn. 13, p. 168) distinguishes predicator as ‘the major constituent of the clause to which the other constituents are syntactically or semantically related.’

mark finite predicates as assertive—is obligatory, as in (38), while in (39), *da* marks the finite predicate as a question:

(38) S. raja … tama ge raja ovunata pilipevi yæ
king self's kingdom them-DAT offer-PAST PRED
"The king offered them his own kingdom.'

(39) S. ta …anan indul-mas kanu yede da
you-ACC …others' leavings eat-PART-NOM suitable-PRES Q
'Is it proper for you to eat the leavings of others?'

Note that *yæ* is the alternant *weyi* which occurs in the modern Sinhala written form of the cleft construction (see (12) in 6.2.1); also (47S) in 3.6.4). Moreover, instances where the predicative is a nominal or adjective, '*yæ* alternates with the third person singular form of the copula *weyi*' (Paolillo 1994: 160).

Classical Sinhala has a cleft construction similar to its modern Sinhala descendant, as in (40a), the structure of which Paolillo (1994) represents as (40b):

(40) a. S. tata karune mahat labha yæ
you-DAT do-PAST.PART-3SG great fortune PRED
'What has been done for you is a great fortune.'

(40) b. S. $[_{\text{S}}\ [_{\text{S}}(_{\text{nominalized}})\ldots \text{Ø}_{i} \ldots \text{V}_{\text{-TNS-PART-3SG}}]\ [_{\text{PRED}}\ \text{XP}_{i}\ \textit{yæ}]]$

What Paolillo (1994: 161) glosses as PART here and the other examples below is 'an invariant third person singular participle (a nominalized verb form)', like the one used in the Tamil/Sinhala cleft construction (1b/20a). In (40a) the focus is marked both by the post-verbal position of the focused constituent and the predicate marker *yæ*, functioning as a focus marker because the focused constituent is represented as a predicate. Note that the subject 'you' does not agree with the 'participle' form because the subject of (40a) is not 'you' but 'what has been done for you' which requires a third person singular neuter verb form. Note also that (40b), which is similar to the structure of the literary Tamil/Sinhala cleft construction ((12) in 6.2.1), according to Paolillo (1994: 161–162), 'seems to be a calque from Dravidian.'

One difference between the classical and modern Sinhala focus constructions is that only non-subject items are focused in the classical construction like (40a/b). If the subject of a sentence is focused, then *yæ* marks the subject as focus, as in (41a), the structure of which Paolillo (1994) represents as (41b): [99]

(41) a. S. mama yæ mandava piye-mi
I FOC trample-PAST-PART-1SG
‘It is I who trampled it.’

(41) b. S. [$_{S}$ NP$_{Subject}$ -yæ ... V-PART-AGR]

The subject agrees with the participle (nominalized) form of the verb in the subject focused construction (41a/b), unlike in the non-subject focused construction (40a/b).
In this type of sentence, no form of lexical copula like *weyi* can substitute for *yæ*, nor can the other predicate markers such as *da* or *la* occur. What is significant to note is that this construction does not involve clefting. Paolillo (1994: 162–163) assumes that since (41a/b) above does not follow the Dravidian pattern, the source of *yæ* here is different from the source of the predicate marking *yæ* in (40a/b), ‘but the two have converged in form in Classical Sinhala’ (ibid. 163). The validity of Paolillo’s assumption is questionable because (41a/b) resemble the Tamil/Sinhala focus particle construction (1c) in which the focus particle occurs adjacent to the focused constituent in the two languages and the finite form of the verb which is marked for PNG is used only in Tamil (1cT), not in Sinhala (1cS). The classical Sinhala focus construction (41a/b) has agreement like the Tamil focus particle construction (1cT) (contra its modern Sinhala counterpart (1cS)). On the other hand, if Paolillo’s (1994) assumption that the source of the subject marking *yæ* (41a/b) is different from predicate marking *yæ* (40a/b) which follows the Dravidian pattern is correct, the former may have been replaced subsequently by *thamay* modelled on the corresponding Tamil particle *thaan* and then extended to non-subject focus constructions (see 1cS) in modern colloquial Sinhala, while the latter, the predicate marking *yæ*, may have been retained in formal/literary Sinhala (10S).

The two focus constructions that Paolillo 1994 distinguishes—namely, the non-subject focused construction (40a/b) and the subject focused construction (41a/b)—co-exist in classical Sinhala. This coexistence, Paolillo (1994: 163) assumes, ‘results in an analytical ambiguity in the case of the third person singular focused subject clauses, since (bare) participial form and a morpheme *yæ* would be licensed by both

---

[99] AGR stands for person number marking, namely 1SG/PL, 2SG/PL or 3SG/PL (see examples 26a–c in Paolillo 1994: 162).

structures'— the non-subject focused sentence (40a/b) and the third person subject focused sentence with the structure like (41a/b)—as in (42):

(42) S. enne saccaka yæ
come-PRES.PART-3SG Saccaka PRED
'The one who comes is Saccaka.'

In the overall paradigm of the classical Sinhala focused sentence types, a focused sentence has a third person singular participle (the nominalized form) in the unmarked case, for this is the form that appears in all the non-subject focused clauses, as well as the third person singular subject clauses. Paolillo, therefore, concludes that classical Sinhala appears to have undergone a paradigm levelling in the direction of the third person singular category, motivated in part by the structural ambiguity of the third person focused-subject clauses and the general unmarkedness of the third person participle in focused sentences. This paradigm levelling would lead to the ultimate replacement of the structure in (41b) by (40b).

Paolillo (1994) hypothesizes the following changes in the agreement system owing to the paradigm levelling of the classical Sinhala focus system. The agreement-marked forms of the present and past participles would be retained only in non-focused finite clauses, that is, those in which the only focus was the predicate. The direct consequence of retaining these agreement marked forms in non-focused clauses is the exceptional behaviour of the modern Sinhala marked finite forms with respect to non-focused sentences because certain of these forms which derive from agreeing verb forms are restricted to non-focused clauses after the paradigm levelling of the focus system. The paradigm levelling motivates the semantic shift characteristic of those from (36) to (30) and (37) to (32) through the development of such focus marking forms *də* and *lu* from predicate markers like *da* and *la*. The status of the focused element is reanalyzed from that of predicate and focus to that of merely focus. The property of the classical Sinhala predicate markers *yæ, da* and *la* that expresses speaker's attitudes is retained in the new focus markers. After the paradigm levelling, the agreement suffixes in classical Sinhala which have a predicate marking function occur only in non-focused clauses. In these clauses, the meanings of the agreement suffixes which function as focus markers could be regularized with those of the other focus markers by a reanalysis of their implicatures as their new content. The old agreeing forms are retained in the contexts in which they had communicative attitudes as invited inferences (Paolillo 1994: 163–164).

A factor which crucially impinges on the development of the Sinhala agreement system, as identified by Paolillo (1994), is the expansion of the discourse role of the focused constructions to other phenomena such as questioning—referred to as naturalization by Gair (see 6.1). Paolillo (1994: 164) is of the view that since the focused sentences lose agreement, 'an increase in their discourse frequency would seriously weaken the agreement remaining in non-focused sentences, possibly rendering it functionally opaque.' Therefore, the two kinds of focus constructions—cleft construction and the focus particle construction (the former to a great extent and the latter to a lesser extent) —which Sinhala replicated from Tamil interacted with the classical Sinhala agreement system to produce the present agreement system with unmarked forms without agreement and the marked forms which require subjects of particular person-number kind.

The hypothesized diachronic development in the Sinhala agreement system (as schematized in (29) in Paolillo 1994: 165) is given in figure 6.1:

Stage1
Classical Sinhala
Finite verb agreement (36)
Non-agreeing, non-subject focussing structure (40b)
Agreeing subject-focussing structure (41b)
predicate-marking *da* and *la* (37)

↓ paradigm levelling

Stage 2
Intermediate stage
(hypothesized)
Non-focussing finite verb agreement (36)
Non-agreeing subject/non-subject
focussing structure (40b)
predicate/focus-marking *da* and *la*

↓ regularization of focus/finiteness meanings

Stage 3
Modern Sinhala
(Colloquial)
Non-agreeing finite forms (28S)
Emphatic sentence/Focus construction (1bS/20aS and1cS)
Attitude-marking non-focussing finite forms ((29)–(33))
Focus-marking *də* ((12bS) in 2.2.12/(31bS) in 2.3.1.6);
and *lu* ((24S) in 4.3.4/(52S) in 4.4.2)

Figure 6.1: Diachronic development in the Sinhala agreement system

The morphological paradigm levelling of the focused sentence constructions motivates the syntactic reanalysis of *yæ*, *da* and *la* from predicate to focus marking and the grammaticalization of agreement markers into markers of speaker's attitude. As a consequence, SV agreement in classical Sinhala becomes restricted to non-focus sentences (from stage 1 to 2). This together with the expansion of the discourse functions of the focused constructions contributes to the weakening of the grammatical function of agreement which becomes restricted only to the marked finite forms, expressing speaker attitudes (from stage 2 to 3). The association of focus markers and agreement markers with the unmarked focused predicates permits a regularization of their meanings: the focus markers have attitude meanings and the agreement markers have attitude meanings as possible implicatures. Thus, it is evident that the present agreement system in Sinhala with the unmarked and marked finite forms resulted from the co-existence of and the consequent interaction between the classical Sinhala focus construction—which Sinhala replicated from Tamil—and the agreement system it had. The final section of this chapter analyzes the effect of the two kinds of focus constructions on polar- and wh-questions and the changes the latter have undergone.

6.5 The impact of focus constructions on questions in Sinhala

One kind of contact-induced changes among many others attested in this research is that once the features which the replica language replicated became established, they interacted with other morphosyntatic phenomena and brought about changes in them (recollect Gair's (1980) naturalization in 6.1). As mentioned above, the two kinds of focus constructions which Sinhala replicated from Tamil interacted with such phenomena as polar-/wh-questions, negatives and quotatives and brought about the changes in them (see below). In this section, the changes brought about in polar and wh-questions by the two kinds of focus constructions are discussed in 6.5.1 and 6.5.2 respectively.

6.5.1 Polar questions

The only difference between the polar questions in Tamil and Sinhala is that the former uses the question clitic *-aa*, while the latter uses the common question particle *də* (2.2.12/2.3.1.6), as in (43) below:

(43) T. appa kooda kaasu selavalli-th-aar-**aa**
Father a lot of money spend-PST-3SGM(HON)-Q
S. thaaththa huŋgak salli viyədamkəra **də**
Father a lot of money spend.PST Q

‘Did father spend a lot of money?’

The structure of the two questions is the same; note that the question marker occurs clause finally, and the finite form of the verb is used. Note that the question particle *də* has been modelled on the Dravidian correlative particle *-oo* (2.3.1.5). Also similar in structure are the constituent polar questions, as in (44a) (questioning the subject) and (44b) (questioning the object):

(44) a. T. appa-(v)**aa** kooda kaasu selavalli-th-athu / selavalli-th-aar
Father-Q a lot of money spend-PST-NMLZ / spend-PST-3SGM(HON)
S. thaaththa **də** huŋgak salli viyədamkər-e
Father Q a lot of money spend.PST-NMLZ

‘Was/is it Father who spent a lot of money?’

(44) b. T. appa kooda kaas-**aa** selavalli-th-aar
Father a lot of money-Q spend-PST-3SGM(HON)
S. thaaththa huŋgak salli **də** viyədamkər-e
Father a lot of money Q spend.PST-NMLZ

‘Was/is it a lot of money that Father spent?’

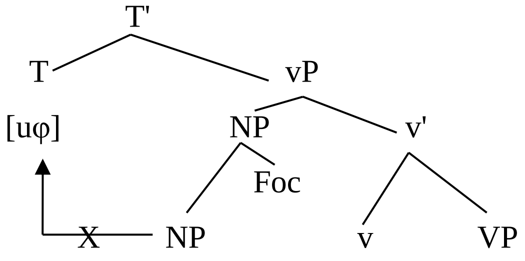
In (44a,b) the question clitic/particle occurs adjacent to the constituent which is questioned. Instances like these in the two languages, which are similar to the focus particle constructions (1c), question the constituent which the clitic/particle is adjacent to, like the focus particle focuses on the constituent which it is adjacent to. By assumption, since they are not cleft constructions, they do not involve a copula overt or covert and the tensed nominalized verbs which occur in (44aT,S and 44bS) do not co-occur with copula (6.2.1). However, these questions have focus interpretations.

The question marker in the two languages has scope over the whole proposition in (43), whereas it has scope only over the constituent which is questioned in (44a,b). The difference between the constituent polar constructions in the two languages is that in the Tamil constituent polar construction, the nominalized form of the verb is used only if the subject is questioned, while the finite form of the verb is also used (44aT); if the other constituents are questioned, the finite form of the verb is used (44bT). In the

Sinhala constituent polar construction, on the other hand, the nominalized form of the verb is the default form (44a,bS).

Why would the subject and object constituent polar constructions be different in this way in Tamil, but not in Sinhala? The following is a viable hypothesis: the polar clitic which merged with the subject NP blocks agreement between the subject and T, that is, it prevents the subject from assigning phi-feature values to T. This is represented in (45) below (assuming that the subject is first merged as a specifier of vP, as has been standard ever since Chomsky 1995).

(45)



Nominalization is then a way to overcome the problem: the nominalizer morpheme has no phi-features that would need valuing.

The structure of the polar constructions (44a,b) indicates that Sinhala has modelled its constituent polar construction on the focus particle construction (1c) which it replicated from Tamil. Consider the following focus particle construction (46):

(46) T. appa (thaan) kooda kaasu (thaan) selavali-th-aar
Father FOC a lot of money FOC spend-PST-3SGM(HON)
S. thaaththa (thamay) huŋgak salli (thamay) viyədamkər-e
Father FOC a lot of money FOC spend.PST-NMLZ

'It was/is Father that spent a lot of money.'
'It was/is a lot of money that Father spent.'

The difference between the focus particle constructions in the two languages is that the finite form of the verb is used in Tamil (46T), whereas the nominalized form of the verb is used in Sinhala (46S). The fact that the focus particle construction and the constituent polar construction in the two languages share the same semantic and syntactic properties indicates that Sinhala has used the model of the focus particle construction for its constituent polar questions; note that it has used the same model for the constituent negative (3.6.3) and constituent quotative (4.4.2) constructions and also wh-questions (6.5.2 below). Crucially the question and quotative clitics, which performed the function of predicate marking, as in (37), in classical Sinhala have changed into focus marking forms in modern Sinhala owing to the changes brought about by the interaction between the agreement system in classical Sinhala and the two focus constructions which Sinhala replicated from Tamil (see (ii) of (i–v), the five facts related to finite and focus marking forms in Sinhala in 6.4).

6.5.2 Wh-questions

Wh-questions in the two languages are of the in situ type. Sinhala wh-phrases consist of wh-words and the question particle *də*, whereas Tamil wh-phrases consist only of wh-words (2.2.13/2.3.1.6). The latter does not use the question clitic/particle in wh-questions, but uses it in polar interrogatives and alternative questions (4.3.6). The fact that Tamil uses a question clitic in polar interrogatives and alternative questions suggests that Tamil may have had a wh-question particle which it has lost. Possibly there is still a null question particle in Tamil wh-questions. Sinhala, on the other hand, has extended the question particle to wh-questions.

An important contact-induced change that has occurred in Sinhala wh-questions is that the nominalized verb form which occurs in the cleft construction has been extended to the majority of wh-questions. One of the differences between the default wh-questions in Tamil and Sinhala is: Tamil uses the nominalized form of the verb only in questions which seek the subject as the answer, while it uses the finite verb form in other wh-questions, just as the polar questions in it; Sinhala uses the nominalized form as its default form (see below for exceptions). As mentioned in 6.2.1 and 6.5.1 above, it is assumed that the nominalized verb form used in default Sinhala wh-questions does not co-occur with null copula, unlike the nominalized verb form which co-occurs with null copula in the cleft construction.

The structure of the wh-questions in the two languages is almost the same. The fact that the question particle occurs adjacent to the wh-base together with the use of the nominalized form, just as in the Sinhala focus particle construction ((1c) shows that the default Sinhala wh-questions have been modelled on its focus particle construction, in other words, the structure of the focus particle construction has been extended to the default Sinhala wh-questions. Consider (47a,b) which question the subject and direct object respectively:

(47) a. T. yaar nettu mala-(u)kku antha puthakath-ai kudu-th-athu
Who yesterday Mala-DAT that book-ACC give-PST-NMLZ
S. kau də iiyee mala-tə ee pothə dunn-e
Who Q yesterday Mala-DAT that book give.PST-NMLZ

'Who gave Mala that book yesterday?'

(47) b. T. kumar nettu mala-(u)kku ennath-ai kudu-th-aan
Kumar yesterday Mala-DAT what-ACC give-PST-3SGM
S. kumar iiyee mala-tə mokak də dunn-e
Kumar yesterday Mala-DAT what Q give.PST-NMLZ

'What did Kumar give Mala yesterday?'

Although (47aT) and (47a,bS) look similar to the cleft constructions in that the nominalized form of the verb is used in all three, they are not cleft questions. Note that they neither have the clause-final focused constituent characteristic of the cleft nor lend a focus reading.[100] It is, therefore, assumed that these questions do not involve the null copula despite the occurrence of the nominalized verb form. I propose that the reason for the fact that the default subject wh-questions (47aT) have the nominalized verb form is the same as in the case of constituent polar questions (see (44aT), also (45) above): the wh-phrase which questions the subject is incapable of assigning phi-feature values to T. Using the nominalized form is a solution to this problem, since the nominalized form has no phi-features that need valuing.

In the two languages, the wh-phrase can occur at the front (clause-initially) (see fn. 100 below) or at the end (clause-finally) in a cleft question, as in (48a). A post-verbal subject wh-phrase which co-occurs with a finite verb is not an option in the two languages, as in (48b).

(48) a. T. nettu mala-(u)kku antha puthakath-ai kudu-th-athu yaar
yesterday Mala-DAT that book-ACC give-PST-NMLZ who
S. iiyee mala-tə ee pothə dunn-e kau də
yesterday Mala-DAT that book give.PST-NMLZ who Q

'Who is it that gave the book to Mala yesterday?' also
'Who gave that book to Mala yesterday?'

---

[100] Note that due to the possibility of scrambling, wh-phrases can be moved optionally to the front in the two languages, as in (i):

(i) T. yaar-(u)kku kumar nettu antha puthakath-ai kudu-th-aan
who-DAT Kumar yesterday that book-ACC give-PST-3SGM
S. kaa-tə də kumar iiyee ee pothə dunn-e
who-DAT Q Kumar yesterday that book give.PST-NMLZ

'Who did Kumar give that book to yesterday?'

(48) b. T. *nettu mala-(u)kku antha puthakath-ai kudu-th-aan yaar
yesterday Mala-DAT that book-ACC give-PST-3SGM who
S. *iiyee mala-tə ee pothə dunna kau də
yesterday Mala-DAT that book give.PST who Q

'Who gave that book to Mala yesterday?'

In (48a) the nominalized verb form is used as in clefts generally (see (1b/20a)): the wh-phrase moves (leftwards) to a focus position, and the remnant nominalized predicate moves to the topic position above (and preceding) the focus position. Given the word order of (48a), it lends a focus reading compared to (47a), its unclefted form. Surprisingly, however, (48aT) may also give an unfocused reading similar to (47aT), so may (48aS) similar to (47aS).[101] The reason why (48bT,S) is ungrammatical is that a verbal constituent cannot be the subject; also note that in (48bS) the verb form needs to be the nominalized form and not the finite form. In Sinhala default wh questions (47a,bS), the wh-phrase occurs in situ, just as in Tamil wh-questions. However, owing to the use of the nominalized verb form as the default form in Sinhala wh-questions, any wh-phrase, subject, direct object or any other function, can occur clause-initially or clause-finally, while in default Tamil wh-questions only the wh-phrase which seeks the subject as the answer can occur clause-initially and clause-finally, those seek non-subjects can occur only clause-initially, but not clause-finally for the same reason stated for the ungrammaticality of (48bT,S).

Clefting can be used in Tamil and Sinhala wh-questions for questioning non-subjects, as in (49):

(49) T. kumar nettu mala-(u)kku kudu-th-athu ennath-ai
Kumar yesterday Mala-DAT give-PST-NMLZ what-ACC
S. kumar iiyee mala-tə dunn-e mokak də
Kumar yesterday Mala-DAT give.PST-NMLZ what Q

'What is it that Kumar gave Mala yesterday?' also
'What did Kumar give Mala yesterday?'

Since the wh-phrase occurs clause-finally in (49T,S), they give a focus reading and the analysis is, again, the same as that of (20/48a). However, as mentioned above, the

---

[101] Sarma (1999: 84) gives an unfocussed reading (like that of (47aT)) to her examples of cleft wh-questions similar to (48aT) in Tamil for which she has not given any reason.

unfocused interpretation as that of (47T,S) is also possible.[102]

A comparison of the Sinhala and Tamil wh-questions, as mentioned above, shows two striking differences between them: one is that the Sinhala wh-phrases have an adjoined question particle which is absent in Tamil and the other is that the default verb form in Sinhala wh-questions is consistently the nominalized form. In the rest of this subsection how these two properties are related to each other is discussed.

Two facts need to be considered in the analysis of Sinhala wh-questions. First, the nominalized form of the verb is obligatory if *də* is adjacent to the wh-phrase; this is the default form of Sinhala wh-questions ((50aS) below and others above) (see Gair and Sumangala 1992, Hagstrom 2004 and Kishimoto 2005). A wh-base in Sinhala, e.g., *kochchərə* 'how much' (quantity) becomes a wh-phrase only when it is combined with the question particle *də* (2.3.1.6). Secondly, in fairly restricted contexts in Sinhala—e.g., (i) in quantifier wh-questions, involving 'how much', 'how many' etc.; and (ii) in embedded clauses which occur as the complement of verbs like *dannəwa* 'know', *hoya balenəwa* 'examine/[search after]', *pariksha kərənəwa,* 'look into/inspect', *teerenəwa* 'understand' etc.—*də* can occur in the clause final position (at the right edge of the clause) in a matrix clause with a finite verb, as in (50bS) or to the left of the complementizer in an embedded clause again with a finite verb (Kishimoto 2005: 8), as in (51bS): [103] Note that in these contexts *də* can occur adjacent to wh-phrases, as in (50aS, 51aS).

(50) a. S. Ø haal kochchərə də gathth-e
Rice how much Q buy.PST-NMLZ
(50) b. S. Ø haal kochchərə gaththa də
Rice how much buy.PST Q

'How much rice did you buy?'

---

[102] Malayalam is another language among the South Asian languages which uses clefting in wh-questions. According to Jayaseelan (1996, 2001, 2003 and 2008), in Malayalam, apart from clefting, a non-cleft question is possible under only one strict condition: the question word must be placed immediately to the left of the verb. He concludes that clefting is a device for positioning the question word next to the verb, the copula *aane*. Jayaseelan (1993, 2003) gives focus reading to his examples of cleft wh-questions, while unfocused reading to the non-cleft questions.

[103] Kishimoto (2005) gives the Sinhala verb for 'examine/[search after]' as *bərənəwa* [sic, *beerənəwa*] 'rescue/save' which is incorrect; the correct verb is *balənəwa* so that *hoya* 'search' together with *balənəwa* 'look for' expresses the meaning 'examine/search after'.

(51) a. S. leela [kau də pingaanə kæduw-e kiyəla] dannəwa
Leela who Q plate break.PST-NMLZ COMP know.PRS
(51) b. S. leela [kauru pingaanə kæduwa də kiyəla] dannəwa
Leela who plate break.PST Q COMP know.PRS

‘Leela knows who broke the plate.’

Note that in (50bS) and in the embedded clause of (51bS) in which *də* occurs clause-finally, the finite form of the verb is used instead of the nominalized form of the verb.

A wh-phrase can occur clause-finally with a nominalized form of the verb in a cleft question (52aS) below, whereas a clause-final wh-phrase which occurs with the finite form of the verb is ungrammatical, as in (52bS), for the same reason discussed for the ungrammaticality of such questions in (48b).

(52) a. S. Ø haal gathth-e kiiyətə də
Rice buy.PST-NMLZ how much Q
‘How much is it that you bought the rice for?’ also
‘How much did you buy the rice for?

(52) b. S.*Ø haal gaththa kiiyətə də
Rice buy.PST how much Q
‘How much did you buy the rice for?’

Note also that given its word order, (52aS) can have a focus reading, while it may also have an unfocused reading, as seen in (49T,S).

Kishimoto’s (2005) account of the syntax of Sinhala wh-questions explicates the movement of the question particle and the use of the nominalized form of the verb. Noting that the scope of wh-in-situ must be specified by one of the two, that is, by the question particle or the nominalized verb form, Kishimoto (2005: 2) assumes that ‘it is Q-movement, rather than movement of a wh-phrase, which is used to form an operator-variable structure in a wh-question, and that a Q-element, while delimiting a wh-constituent in its Merge position, serves as an operator that assigns scope to its host *wh*-in-situ.’ For example, in (50aS) the Q element (*də*) delimits a wh-constituent, while in (50bS) it assigns scope to its host wh-in-situ.

In (53aS) below *də* specifies matrix scope, whereas in (53bS) it specifies embedded scope:

(53) a. S. mary [oyaa kochchərə t_i viyədamkəra kiyəla] dannəwa $də_i$
Mary you how much spend.PST COMP know.PRS Q
‘How much does Mary know that you have spent?’ also
‘Does Mary know how much money you have spent?’

(53) b. S. mary [oyaa kochchərə t_i viyədamkəra $də_i$ kiyəla] dannəwa
Mary you how much spend.PST Q COMP know.PRS
‘Mary knows how much money you have spent.’

In (53aS) the particle *də* moves from the base embedded wh-constituent to the matrix clause, whereas in (53bS) it moves from its base position to the embedded clause-final position. [104] It is important to note that in (53a,bS) the verb in the matrix and embedded clauses respectively is finite. In Tamil, the wh-questions do not have a question clitic/particle and distinct intonation patterns are used to express the embedded and matrix scope in instances like (53a,bS), as in (54T):

(54) T. mary-kku [nee evvalavu selavali-th-aai endu] therium
Mary-DAT you how much spend-PST-2SG COMP know.PRS.3SGN
‘How much does Mary know that you have spent?’ also
‘Mary knows how much money you have spent.’

The movement of the question particle in Sinhala questions, according to Kishimoto (2005), occurs either in overt syntax or in LF. He notes that the fact that the particle *də* which is used to delimit a wh-constituent can be attached to a number of different types of constituents without changing their categorical properties suggests that the operation to generate *də* involves adjunction. Moreover, he argues that since *də* is a particle which does not project any further, it can be assumed to be ‘a non-projected head’ which can in principle be either a maximal or minimal projection. However, Kishimoto (2005) argues that since *də* must be adjoined to a maximal projection when it is merged, it is plausible to hypothesize that *də* counts as a phrasal element (XP) in the syntax. Note that when *də* is merged in its merge position, e.g., in (50aS, 51aS), it is merged as an XP element with a maximal projection which includes the wh-form because only an XP can be adjoined to an XP.

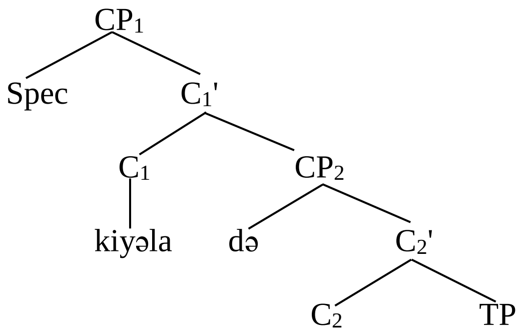
As an operator binding a wh-phrase, *də* occurs to the immediate right of the finite verb in matrix wh-questions (50bS; also 53aS), and in embedded questions *də* occurs to the immediate right of the finite verb and to the left of the complementizer

---

[104] Note that (53aS) is ambiguous between ‘How much does Mary know that you have spent? and ‘Does Mary know how much money you have spent?’ The possible reason for the latter, the polar question reading, is the occurrence of the Q-particle in the clause-final position adjacent to the finite verb. In instances like these, intonation is used to express the intended meaning.

kiyəla, as in (51bS; also 53bS). The coexistence of the Q-particle and the complementizer in (51bS; also 53bS) suggests that CP recursion is allowed in Sinhala. To explain the syntax of wh-questions, Kishimoto's (2005: 15–16) assumption that 'Sinhala has two layers of CP projections' in accordance with the CP stacking theory is adapted here as represented in (55):

(55)



Note that the lower CP (CP_2) is occupied by *də*, while the upper CP (CP_1) is occupied by complementizer. Note that the complementizer *kiyəla* occupies the C_1 head position. It is assumed that *də* occupies [Spec, C_2], rather than the C_2 head position because the (clause-final) *də* which assigns scope to a wh-phrase is a non-projected head which acts like a phrasal element in the syntax. Since *də* which is located in scope position (clause-finally) does not mark a delimited wh-constituent, it is not base-generated in the spec of CP. This means that it must have ended up there by movement from a constituent position where it is merged to mark a delimited wh-constituent to the scope position where it can c-command and bind the wh-word (53a,bS).

Given the assumption that *də* may count as a non-projecting head which can be a minimal or maximal projection, it can—as a minimal projection— act like a head affixed to the questioned-constituent on its left at PF, as in (56) below, where it is affixed to 'who'. What is crucially important to note is that if *də* delimits a wh-constituent in situ in this manner, the scope of the wh-phrase is interpreted relative to this verbal marking, producing 'participle-predicate concord', that is, the co-occurrence of the nominalized form of the verb with *də* located adjacent to a wh-phrase (Kishimoto 2005:4), as in (56a,bS):

(56) a. S. nimal [kau də horəkaŋkəra kiyəla] dann-e
Nimal who Q steal.PST COMP know.PRS-NMLZ
‘Who does Nimal know stole?’

(56) b. S. nimal [kau də horəkaŋkər-e kiyəla] dannəwa
Nimal who Q steal.PST-NMLZ COMP know.PRS
‘Nimal knows who stole.’

In (56aS) the matrix clause has the nominalized form; hence, the nominalized verb encodes matrix scope, whereas in (56bS) the embedded question has the nominalized form of the verb; hence, the nominalized form encodes embedded scope. Note that in both, the question particle is adjacent to the wh-phrase; hence, it does not assign scope. The nominalized verb form, therefore, performs the function of specifying the scope of these wh-questions (see Gair and Sumangala 1992, Hagstrom 2004 and Kishimoto 2005). On the other hand, if *də* occurs clause finally, as in (53a,bS), *də* marks the scope and hence the nominalized form does not occur. These are the major operations involved in wh-questions in Sinhala (see Kishimoto 2005 for details). Again in Tamil wh-questions, the nominalized form of the verb is used only when the subject is questioned and it is not used to mark the scope because in Tamil, the scope is indicated by intonation in questions like (57T):

(57) T. nimal-kku [yaar kalavu edu-th-athu endu] therium
Nimal-DAT who theft take-PST-NMLZ COMP know.PRS.3SGN
‘Who does Nimal know stole? also
‘Nimal knows who stole.’

The above discussion on Sinhala wh-questions amply demonstrates that wh-questions have been modelled on the focus particle construction (1c,d; also 46). The question particle *də* in wh-questions has the same function as the focus particle *thamay*. In its merged position, *də* which co-occurs with the nominalized verb form delimits a wh-constituent which is assumed to have inherent focus (56a,bS), just as *thamay* which co-occurs with a nominalized verb form assigns exhaustive focus to a constituent which it is adjacent to. In its scope position (clause-finally) *də*, as an operator, which co-occurs with the finite verb assigns scope to its host wh-in-situ (53a,bS), just as *thamay* which co-occurs with the finite verb assigns exhaustive focus to the entire sentence.

Certain features in Sinhala wh-questions, namely, the presence of the question particle, its capacity for movement and the use of the question particle and the nominalized form to mark scope, make them different from Tamil wh-questions. These distinct features of Sinhala wh-questions, as the discussion above shows, have resulted from the interaction between the wh-questions and the two kinds of focus constructions which Sinhala replicated from Tamil. Polar questions in the two languages, on the other hand, are almost similar, except for the fact that Tamil uses a question clitic, whereas Sinhala uses a question particle; the positions in which the question clitic/particle occurs and their functions, however, are the same. What is important to note is that the polar and wh-questions in Sinhala have, as shown above, been modelled on the Tamil/Sinhala focus constructions.

6.6 Conclusion

The two kinds of focus constructions which Sinhala replicated from Tamil and the changes they brought about in the morphosyntax of Sinhala discussed in this chapter provide interesting insights into the effects of language contact. The two focus constructions in the two languages share the same properties as shown in 6.1–6.3. The replication of the two focus constructions from Tamil brought about significant changes in the morphosyntax of Sinhala. First, as shown in 6.4, owing to the interaction between the two kinds of focus constructions and the agreement system in classical Sinhala, the latter has undergone restructuring which led to the present agreement system in Sinhala, comprising two kinds of finite verbs: the major part of verbs which are unmarked for agreement features and a limited set of verbs which require subjects of a distinct person and number kind. The crucial step in the process of change which occurred in Sinhala may have been the replication of the use of the nominalized verb form which occurs in the Tamil cleft construction.

Secondly, the nominalized form of the verb which is a distinct feature of the cleft construction permeated the other morphosyntactic phenomena and became the default verb form in wh-questions and negatives (see below) in addition to the constituent polar/negative/quotative/focused constructions (see table 6.1). Thirdly, concurrent with or subsequent to the replication of the use of the nominalized form of the verb are the contact-induced changes occurred in Sinhala polar questions and wh-questions, as shown in 6.5.1 and 6.5.2 (also in negatives and quotatives, which have not been discussed here) in that the structure of these constructions shows that these have been modelled on the focus particle construction: the respective clitic/particle occurs clause-finally, providing wide scope or occurs adjacent to a constituent, providing narrow scope.

Finally, as shown in 6.5.2, the nominalized form of the verb together with the question particle marks scope in Sinhala wh-questions: in default wh-questions in which the question particle delimits a wh-constituent in its Merge position, the scope of the wh-constituent is specified by the nominalized form, while in marked wh-questions, the clause-final question particle assigns the scope to its host wh-in-situ. What is important to note is that Tamil wh-questions do not have these phenomena; hence, these phenomena in Sinhala may have been brought about by independent endogenous changes subsequent to the contact-induced changes. These endogenous changes arguably account for the fact that SV agreement and wh-questions, as identified in this dissertation, are synchronically different between Sinhala and Tamil, by and large, owing to the diachronic changes occurred in spoken Sinhala resulted from the replication of the two kinds of Tamil focus constructions and features related to them. It is this process of post-contact endogenous changes in Sinhala that Gair (1980) refers to as ‘naturalization’ (see 6.1).

Chapter 7

Conclusion

A characteristic feature of languages is that they change over time. These changes are either endogenous or exogenous or both. The former occur when a language undergoes language-internal lexical, phonological morphosyntactic and semantic changes. The latter result from language external factors, especially when languages are subjected to changes induced by their contact with other languages. Of the two kinds of language change, the exogenous changes owing to the influence of other languages are rapid and more complex than the endogenous changes. Endogenous changes affecting two languages of diverse origins will seldom make the languages share a wide range of similar features, simply because the complexity of the grammatical system makes it an improbable outcome. Further, no two languages can share the same lexicon just by accident. Therefore, if two languages share a wide range of similar features, then it is either because they have inherited these features from a common ancestor, or it is due to exogenous changes caused by contact between the two languages. These changes induced by contact between languages belonging to different language families range from borrowing simple lexical items to replication of complex morphosyntactic features. The Sinhala-Tamil contact situation involving two languages of NIA and Dravidian origins which this dissertation focuses on demonstrates vast scale contact-induced changes from lexical items to complex morphosyntactic restructuring in Sinhala.

In this research a wide range of morphosyntactic features of modern spoken Tamil and modern spoken Sinhala in Sri Lanka, have been analyzed, via comparison, with the view to establishing that modern spoken Sinhala has undergone contact-induced restructuring on the model of modern spoken Tamil. The main motivation for this research is that the two languages under study share a wide spectrum of morphosyntactic features, which is unusual for languages of two different families (see research questions in 1.1). Since a full scale diachronic study to determine the history of convergence between the two languages was beyond the scope of this dissertation, a synchronic study was undertaken. However, if information about the history of any features was available, it was discussed. A total of 64 morphosyntactic features were examined to determine whether Sinhala morphosyntactically converges with Tamil and to what extent the former has restructured itself on the model of the latter.

The Sinhala-Tamil contact situation can be best classified as borrowing (as opposed to interference through shift) in which speakers of the replica language in the beginning of their contact with another language tend to borrow lexical items from the source language in an attempt to maintain their language and gradually, depending on the intensity of the contact, begin to restructure the former's morphosyntax on the model of the latter. The interaction between the Sinhalese and the Tamils indubitably resulted—and also by virtue of both having had their origins in the same region—in the diffusion of various cultural, religious aspects and practices of the Tamil community into the Sinhala community. Equally, if not more, important is the considerable influence the Tamil language has exerted on the Sinhala language and the changes induced by the former in all linguistic subfields of the latter, namely lexicon, phonology, morphology and syntax as attested in the literature. In the beginning, speakers of Sinhala who aspired to better social status began to learn Tamil—the then language of power and prestige owing to the reign of Sri Lanka by the South Indian Tamil kings and the fact that Tamil was the commercial lingua franca of the region covering the southern part of India and Sri Lanka (see 1.5). In terms of language dominance, therefore, the Sinhala-Tamil situation is of the displacive kind in which influence tends to be unidirectional, that is, the influence of Tamil on Sinhala.[105]

Since the two languages had their origins in India where there had been extensive language contact before they were located in Sri Lanka, the two languages share with other languages of the region the areal features, e.g., the SOV word order, the dative subject construction, compound verbs etc. Dravidian, especially proto Dravidian is reckoned to be the source of most of these areal features; evidence comes from the fact that NIA languages, a sub branch of the Indo-European family of languages, share more features with the Dravidian languages than with the languages which belong to the Indo-European family. Given this, the question, then, is whether Sinhala, at least in the case of some features (see below) acquired them indirectly through its contact, in India, with languages which had already adopted these features from a Dravidian language before it came to Sri Lanka or directly through its contact with Tamil in Sri Lanka.

The Sinhala-Tamil contact situation also displays the characteristics of metatypy, or morphosyntactic convergence discussed in the literature on contact linguistics. As a result of the prolonged Sinhala-Tamil bilingualism prevalent at least among a considerable number of speakers in these two speech communities ever since these languages came into contact, Sinhala has undergone metatypy, a diachronic process of morphosyntactic restructuring. Consequently, Tamil which served as the intercommunity language provided the metatypic model for the morphosyntactic restructuring of Sinhala. Close examination of the morphosyntactic features of the two languages display also the diverse processes proposed in the literature (see 1.3). Sinhala-Tamil contact situation can be distinguished as what

---

[105] What needs to be emphasized repeatedly is that Tamil is no longer a dominant language in Sri Lanka. Being spoken by the majority of the population in Sri Lanka, Sinhala is the dominant language in Sri Lanka now.

Thomason and Kaufman (1988) refer to as ‘extreme structural borrowing under very intense contact’.

That Tamil has restructured itself on the model of Sinhala can safely be ruled out, for almost all the morphosyntactic features that are similar between the two languages are Dravidian in origin. The remarkably similar features which the two languages share—as the discussion of the findings in each chapter indicates (also see below)—disprove the claim that Sinhala may have developed these features via an endogenous process of grammaticalization. Equally invalid is the view that these two languages adopted these features from a common ancestor or a third language with which both have been in contact: the former seems impossible because these languages are of diverse origins and the latter too is unlikely because no language except the two have been in Sri Lanka until they came into contact with the languages of European traders and colonizers, which have not induced morphosyntactic changes to any significant extent in either of them. The reason for the similarities between the two languages is, therefore, the morphosyntactic changes that Sinhala has undergone over time induced by its contact with Tamil.

In what follows in this concluding chapter, a brief summary of the chapters in the dissertation is given in order to show how each chapter contributes to the dissertation as a whole. Only a brief summary of the chapters is given because the wide scale morphosyntactic convergence between Tamil and Sinhala is closely examined in each chapter and the implications for the contact-induced changes occurred in diverse morphosyntactic features in Sinhala are accounted for.

Chapter 1 provides a comprehensive overview of the information necessary for the rest of the dissertation. 1.1 states the research problem and the aim and objectives of the research. 1.2 describes the different kinds of language classification with special reference to the two kinds of classification, genetic and areal, which help account for the similarities and differences between Tamil and Sinhala in particular and the languages of the South-Asia in general. 1.3 discusses diverse processes involved in contact-induced restructuring and their major implications for the morphosyntactic restructuring that Sinhala has undergone. 1.4 focuses on language universals from linguistic typological and generative linguistic perspectives—the two approaches used in the research for analysis—and introduces areal typological features with special reference to South Asia. 1.5 gives an outline of the history of languages in Sri Lanka found in the existing literature with special reference to the contact between the two languages against the Sri Lankan socio-political historical backdrop. Finally, 1.6 discusses the methods used in the research.

Both macro- and micro-level analyses are undertaken to show the morphosyntactic restructuring of Sinhala. In the macro-level analysis a total of 64 features—51 features distinguished in Dryer and Haspelmath (2013) and 13 others—are compared in the two languages to determine the features of Sinhala that have undergone contact-induced restructuring. Chapter 2 compares 17 features related to word order (2.2) and analyzes the implications for the vast scale convergence between Tamil and Sinhala. Chapter 3 compares 27 features related to simple sentences divided into four sections: Case and Agreement (3.2), Valence and Voice (3.3), Negation and Questions (3.4) and Predication (3.5); and analyzes the diverse contact-induced changes in Sinhala manifested in these features. Compared in Chapter 4 are features related to complex sentences, and other features which are deemed important to show the contact-induced changes occurred in Sinhala, resulting from the restructuring/replication of these features of Sinhala modelled on the corresponding features of Tamil. In the micro-level analysis, two phenomena and the features related to them are closely analyzed to study the nature of changes that have occurred in these two phenomena in Sinhala. Chapter 5 closely examines the occurrence of null arguments in the two languages and the conditions under which they occur. Chapter 6 analyzes the two kinds of focus constructions which assign exhaustive focus to a constituent in a sentence and the role played by the cleft construction, one of the two kinds of focus constructions in the two languages, that is believed to have caused significant changes in Sinhala morphosyntax.

Summarized below are the diverse contact-induced changes attested in the morphosyntactic restructuring of Sinhala on the model of Tamil, and facts which point to the convergence between Sinhala and Tamil despite their diverse origins:

(1) Appropriation—a transferred form borrowed wholescale:
the existential quantifier clitic *-oo* in Tamil is used as the existential quantifier clitic in Sinhala (4.3.7). Note that in Tamil, the same clitic is used as the correlative clause marker (2.3.1.4) and the disjunctive coordinator (4.3.5). The near homophonous disjunctive coordinator in Sinhala is *-ho* (4.3.5).

(2) MAT(ter) replication—the replication of form-meaning units often near homophonous:

a) the reflexive *thaman* 'self' in Sinhala has been modelled on the corresponding reflexive *than* in Tamil (3.3.1).

b) the focus particle *thamay* in Sinhala 'indeed' has been modelled on the corresponding focus particle *thaan* in Tamil (6.1).

(3) PAT(tern) replication or grammatical calque—the replication of the structure of constructions:

- **(i) segments at the level of affixes, clitics or particles:**
  - a) the accusative case marker -*wə* in Sinhala has been modelled on the corresponding accusative marker -*ai* in Tamil. Note Masica's (1991) claim that except Sinhala, the other major NIA languages do not have a distinctive accusative marker and may take it in the form of 'dative-accusative' marker (2.3.1.2).
  - b) the quotative evidential clitic -*lu* in Sinhala has been modelled on the corresponding evidential clitic -*aam* in Tamil (4.3.4).
  - c) the causative morpheme -*wə* occurs in both periphrastic (3.3.5) and morphological (3.3.6) causative constructions in Sinhala has been modelled on the corresponding causative morpheme -*vi* which occurs in Tamil morphological causative construction.
  - d) the comparative particle *vadaa* in Sinhala has been modelled on the corresponding comparative particle *vida* in Tamil (3.5.5).
- **(ii) small constituents/constructions:**
  - a) the reciprocal pronoun *denna-tə denna* 'two-DAT two' in Sinhala has been modelled on the corresponding Tamil reciprocal pronoun *oruvar-ai-oruvar* 'one(HON)-ACC-one(HON)' (3.3.1).
  - b) the ordinal numbers are formed in Sinhala by suffixing -*wæni* to the base form of numerals, just as suffixing -*aavathu* to the cardinals to form ordinals in Tamil; these suffixes have been derived from one of the two copulas in the two languages (2.3.1.1).
  - c) the two kinds of verbal nominals, tensed and non-tensed, are formed in Sinhala, just as they are formed in Tamil (4.3.8).
- **(iii) bigger constructions:**
  - a) the cleft construction in which the cleft constituent occurs clause-finally in Sinhala has been modelled on the corresponding Tamil cleft construction (6.1).
  - b) different kinds of conjunctive and disjunctive coordinating constructions in Sinhala have been modelled on the corresponding Tamil coordinating constructions (4.3.5).

(4) **Word order restructuring**—the existing structure of the replica language is rearranged/replaced on the model of the corresponding structure in the model language.

Sinhala may have restructured the following constructions on the model of the corresponding Tamil constructions:

a) the prenominal relative clause (2.2.10 and 2.3.1.3)
b) the clause-final polar and the constituent polar questions (2.2.12, 2.3.1.6 and 6.5.1)
c) the clause-final negative and the constituent negative constructions (3.4.1 and 3.6.3)
d) the clause-final quotative evidential and the constituent quotative evidential constructions (4.3.4 and 4.4.2)
e) the clause-final complementizer in utterance complement clause (4.2.7 and 4.4.1)

(5) **Replica grammaticalization**—the process of grammaticalizing a lexical item to form a functional/grammatical category in the source language is replicated by the replica language for the fomation of the corresponding grammatical category from the same lexical item, in other words, there is a grammaticalization path way for the replication:

a) the formation of indefinite pronouns (both interrogative- and generic pronoun-based) by adding the inclusive clitic/conjunctive coordinator *-th* to wh-words/phrases and the generic pronoun 'one' (denoting a person or animal/thing) in Sinhala, just as the formation of the corresponding indefinite pronouns by adding the inclusive/conjunctive coordinator clitic *-um* to wh-words/phrases and the generic pronoun in Tamil (see 3.6.3)
b) the formation of universal quantifiers and existential quantifiers by adding the inclusive/conjunctive coordinator clitic *-th* (similar to wh-based indefinites in (a) above) and the disjunctive coordinator *-oo* respectively to wh-words/phrases (4.3.7), just as the formation of the corresponding quantifiers by adding the conjunctive, disjunctive coordinators to wh-words/phrases in Tamil
c) the formation of the two copulas by grammaticalizing two lexical verbs, just as the formation of the two corresponding copulas in Tamil (3.6.4)

d) the formation of functional verbs/auxiliaries: the reflexive verb/auxiliary (3.6.2); the applicative auxiliary (3.6.2); and other auxiliaries, e.g. aspectual auxiliaries (4.3.3) by grammaticalizing lexical verbs, just as the formation of the corresponding functional verbs in Tamil

e) the formation of the disjunctive coordinator by grammticalizing the lexical adjective *hari* 'correct' in Sinhala, just as the formation of the correseponding coordinator in Tamil by grammaticalizing the corresponding lexical adjective *sari* with the same meaning (4.3.5)

(6) **Extension or context generalization**—a grammatical category in the replica language being generalized into the same contexts into which the source language had generalized the corresponding category:

(i) constituents extended to other contexts:

a) the clitic *-th* is used as: i. the conjunctive coordinator (4.3.5); ii. the concessive marker (2.3.1.7); and iii. the marker which is used to form the two kinds of indefinite pronouns (see (5a) above) and universal quantifiers (see (5b) above).

b) the dative case has been extended to a variety of semantic contexts, just as the dative marker has been extended to the corresponding contexts in Tamil (3.6.1.1). Note also that this dative marker is also overgeneralized into other contexts in Sinhala, e.g., the formation of adverbs from adjectives, while the corresponding dative case in Tamil has not been generalized into other kinds (3.6.1.1).

c) the locative marker is also used as the possessive marker, just as in Tamil (3.5.1 and 3.6.1.1).

d) the nominalized form of the verb which occurs in the cleft construction replicated from Tamil has been extended to default wh-questions (2.3.1.6); constituent polar questions (2.3.1.6); default and constituent negatives (3.6.3); constituent quotatives (4.4.2); and focused particle constructions (6.1) (see also table 6.1).

(ii) a structure extended to other constructions:

a) the structure of the focused particle construction (6.1) replicated from Tamil has been extended to the following phenomena: i. default wh-questions (6.5.2); ii. constituent polar questions (2.3.1.6); iii. constituent negatives (3.6.3); and iv. constituent quotatives (4.4.2).

(7) Paired grammaticalization—two grammaticalization processes for the same general grammatical function:

a) the two topic markers grammaticalized from the conditional suffix in Sinhala and the lexical verb ‘go’, just as the corresponding topic markers have been grammaticalized from the same lexical items, the conditional suffix together with the complementizer in Tamil and the verb ‘go’(4.3.2, 4.4.2). The only difference, though, is that in Sinhala, the first topic marker consists only of the conditonal suffix unlike in Tamil.

b) the two kinds of focus constructions in Sinhala, namely the cleft construction and the focus particle construction—the former involves restructuring, while the latter involves grammaticalization in that the focus marker is formed from the reflexive pronoun (anaphor), just as in Tamil (6.1). However, both constructions perform the same grammatical function of assigning exhaustive focus to a constituent.

(8) Co-existence—the inherited (old) and the innovated (new) structures being used in one grammatical category. Co-existence is of two kinds:

(i) co-occurrence of the new and the old categories in the same construction, resulting in double marking:

a) double marking of indefiniteness in some expressions—prenominally by numeral ‘one’ and postnominally by the indefinite marker (20aS in 2.3.1.1)

b) causation is double marked in periphrastic construction using the causative morpheme and the causative auxiliary ((12S) in 3.3.5).

c) aberrant teen numerals (see 2.3.1.1)

(ii) the two structures coexist as alternative constructions:

a) preverbal and postverbal negatives in Sinhala ((23) in 2.3.1.1 and (15aS) in 3.4.2)

(9) Remnant NIA features:

a) the dative, accusative, and instrumental constructions ((37aS–cS) respectively in 3.6.1.1) which are similar to those in the split ergative/absolutive system characteristic of NIA languages

b) volitive and involitive verbs (4.3.9)

(10) Diverse features which Sinhala and Tamil share with other languages of the region:

- a) the use of ‘one’ as the indefinite marker (2.3.1.2)
- b) definiteness being not overtly marked (2.3.1.2)
- c) the distal demonstrative being optionally used as the definite marker (2.3.1.2)
- d) the demonstrative based third person pronouns having the same composition: the demonstrative stem being added to the bound pronominal base encoding the feature content (4.3.1)
- e) the third person distal demonstrative pronoun, performing pronominal and anaphoric functions (5.4)
- f) different kinds of predication:
  - i. those in which the copula is null in the present tense ((23a) in 3.5.2; (24a) in 3.5.3; (25) in 3.5.4); and (26) in 3.5.5);
  - ii. those in which the copula is overt in the present tense ((22) in 3.5.1); and ((24b) in 3.5.3);
  - iii. forms in both i. and ii. above in which the copula is overt in the past tense
- g) the alternative question construction (4.3.6)
- h) all complex sentences which have the same structures: the deranked (dependent) forms which have non-finite verbs (4.2.3–4.2.6) and the balanced (independent) forms which have finite verbs (4.2.7).

(11) Attested areal features

- a) the SOV word order (2.2.1)
- b) the morphological causative construction (3.3.6)
- c) compound verbs consisting of diverse auxiliaries (functional categories) (4.3.3).
- d) the dative subject (4.3.11) and double subject (5.4) constructions
- e) reduplication (4.3.12)
- f) converbs (4.3.13)

(1–11) are the diverse contact-induced changes attested in Sinhala morphosyntax compared with Tamil morphosyntax in this dissertation. Note that the majority of constructions in the two languages are strikingly isomorphic and they are intertranslatable most often, morpheme by morpheme. Moreover, in some, if not most, instances, the genetic relatives of Tamil have corresponding constructions, whereas the genetic relatives of Sinhala do not have these constructions. As for areal features and other features believed to be shared by the South Asian languages, since Tamil is believed to be the source language of the region, Sinhala may have acquired these areal features: i. on the South-Asian mainland from other languages which have been in contact with Tamil/Dravidian languages before it evolved as Sinhala in Sri Lanka or ii. through its direct contact with Tamil in Sri Lanka (see 1.5). However, (1–11) above amply demonstrate that ii. is more probable and viable than i. and that Sinhala has adopted considerable morphosyntactic features from Tamil in Sri Lanka.

It is also useful to analyze the following differences arising from the comparative analysis of the morphosyntax of the two languages:

(i) the noun numeral order in Sinhala (2.2.9, 2.3.1.1) may be a remnant substrate feature of the indigenous languages or an endogenous change taken place in Sinhala.

(ii) the values of the feature, ‘the position of polar question particles’ (2.2.12 and 3.4.7) are different not because of the position of the particle but because of the type of the question marker: Tamil uses a question clitic, while Sinhala uses a question particle. The position of the clitic/particle in Sinhala and Tamil polar questions is the same. Though the question particle is one of the distinctive features of Sinhala, it has, as shown in 2.3.1.5, been modelled on the Tamil correlative clitic -*oo* (Slade 2013).

(iii) the values of three features related to SV agreement, regarding marking of agreement on finite verbs (3.2.4–3.2.6) are different due to the lack of SV agreement in Sinhala which has lost most of its agreement owing to the interaction between the agreement system and the two kinds of focus constructions, resulting in its present agreement system (6.4).

(iv) Of the values of three features related to negation, those of ‘Order of negative morpheme and verb’ (3.4.2) and ‘Symmetric and asymmetric standard negation’ (3.4.4) are different because of the preverbal negatives in Sinhala (3.4.2) which in turn accounts for the difference between the values of the feature ‘Symmetric and (a)symmetric negation’ (3.4.4), while the disparity between the values of the feature ‘Subtypes of asymmetric standard negation’ in 3.4.5 is due to the fact that Tamil has a ‘Subtype of Asymmetric Standard Negation (A/Cat)’ which Sinhala does not have.

(v) the volitive/involitive distinction in Sinhala verbs (4.3.9) can be comparable with the affective/effective distinction in Tamil verbs. However, the former has a wider distribution among verbs in Sinhala than the latter among verbs in Tamil. The volitive/involitive distinction in Sinhala verbs is assumed to be an NIA feature that Sinhala has retained.

(vi) the nominalized form of the verb occurs in default wh-questions (2.3.1.6) and negatives (3.6.3) because Sinhala has extended the use of the nominalized verb form which occurs in the cleft construction. The nominalized form of the verb is not used in Tamil default wh-questions and negatives; hence the difference. However, this is a contact-induced change in that subsequent to the replication of the cleft construction in which the nominalized verb form occurs from Tamil, Sinhala extended the nominalized verb form to other phenomena.

(vii) the reason for the difference between the differential object marking (DOM) (2.3.1.2) in the two languages is that in Sinhala, it is determined by animacy, whereas in Tamil, it is determined by definiteness.

Most of the differences ((i)–(vii)) above resulted from varied contact-induced changes in Sinhala.

At the micro-linguistic level, Chapter 5 examines the properties of licensing null arguments in the two languages, both null subjects and other arguments, to find out whether the two languages display the same behaviour when one factor, namely, SV agreement—which is believed to play an important role in the occurrence of null arguments cross linguistically—is different between these languages. In 5.2, the tests proposed in Cole (2010) which aim to account for the intra-/inter-language differences with respect to the occurrence of thematic null subjects are applied to Tamil and Sinhala. The results obtained show that at intra-language level, the two allow null subjects in similar contextual conditions and have exactly the same degree of context-dependence. At inter-language level, they are to a significant extent similar to Chinese, among the languages included in the comparison, which indicates that the difference in SV agreement between the two languages has no effect on subject pro-drop. 5.3 outlines the current typology of null subjects and the major theories in the literature which are used to explain the occurrence of null subjects. In 5.4 on the basis of the results from the tests in the literature applied to Tamil and Sinhala, they are classified as discourse pro drop languages. Moreover, the various contextual conditions in which Tamil and Sinhala allow null subjects/non subjects are examined. The properties of the two languages that account for the occurrence of null arguments are analyzed in terms of the theories outlined in 5.3.

Arising from the analysis are the following properties of the occurrence of null arguments in Tamil and Sinhala: i) they allow null subjects in embedded clauses if they have a c-commanding topic antecedent; ii) they allow generic null subjects; iii) they allow null subjects in main clauses with an antecedent in a separate sentence even if there is no c-commanding antecedent provided that the antecedent is in close proximity; iv) they allow null arguments other than subjects. Tamil and Sinhala thereby show properties of discourse pro-drop languages, with no discernible effect of the rich SV agreement in Tamil on the occurrence of null subjects in it. In the concluding section, the implications of the current typology of NSLs for the theory of pro-drop are briefly discussed. On the basis of the discussion of Jayaseelan (1999), Tomioka (2003), Modesto (2008) and Barbosa (to appear), it is proposed that in the two languages, the referential null arguments are the result of null NP anaphora, while the null indefinite pronouns, being non-anaphoric, are base generated empty pro forms. The crucial syntactic features which determine the occurrence of null arguments in the two languages, especially the null referential ones, are topic prominence and the absence of overt definite marking on noun phrases.

Chapter 6 analyzes the two kinds of focus constructions—the cleft and the focus particle construction—that Sinhala replicated from Tamil. These two constructions have interacted with the features that Sinhala had when it came into contact with Tamil and have caused significant changes in the morphosyntax of Sinhala. 6.2 describes and analyzes the almost identical semantic and syntactic properties of the cleft construction which Sinhala shares with Tamil/Dravidian. A derivation of the cleft construction is proposed which is similar, but not identical to the one proposed for Malayalam by Jayaseelan and Amritavalli (2005). 6.3 discusses the role of relativization in the long distance cleft construction in the two languages for the extraction of the constituent to be clefted. This extraction results in the formation of the complex NP from the rest of the clause. It also analyzes the role of relativization in the derivation of the nominalized verb form which the complex NP in the cleft construction ends with in the two languages. 6.4, following Paolillo (1994), shows that the agreement system that Sinhala once had has undergone restructuring over time owing to its interaction with the two kinds of focus constructions, the syntax of which has been modelled on those of Tamil. The restructuring of the agreement system in Sinhala has resulted in the present agreement system, comprising unmarked finite forms without agreement and a small set of marked finite forms with agreement.

6.5 examines polar questions (6.5.1) and wh-questions (6.5.2) in the two languages and shows that the two kinds of questions like negative and quotative constructions in Sinhala have been modelled on the focus particle construction in that in these constructions the clitic/particle occurs either adjacent to constituents or clause-finally, assigning narrow and wide scope respectively. 6.5.2 shows that the nominalized form of the verb, a distinct feature of the Tamil/Sinhala cleft construction, has been extended to wh-questions. Note also that it has been extended to other phenomena (see (6id) above). In Sinhala wh-questions, the question particle together with the nominalized form of the verb performs an important syntactic/semantic function of assigning scope. In default wh-questions in which the question particle delimits a wh-constituent, the scope of the wh-constituent is specified by the nominalized form, whereas in marked wh-questions in which the question particle occurs clause-finally to the right of the finite verb, the question particle assigns the scope to its host wh-in-situ. These two kinds are structurally similar to the constituent polar/negative/quotative constructions and their default forms in which these markers occur clause-finally. What has happened is that once Sinhala replicated the two focus constructions from Tamil, the features of these constructions have been extended to other morphosyntactic phenomena—and in some phenomena after endogenous morphosyntactic changes—become established in these phenomena.

The comparative analysis of the wide range of morphosyntactic features undertaken in this research shows significant convergence between Sinhala, which is rather unexpected given that they belong to two different language families, NIA and Dravidian respectively. The reason for their convergence is that one language, namely Sinhala has undergone considerable changes induced by its contact with the other language, namely Tamil. At macro-linguistic level, the fact that 55 of the 64 features of Sinhala and Tamil compared in the dissertation are similar (see (1–11) above) shows that a wide range of morphosyntactic features in Sinhala have undergone considerable changes on the model of the corresponding features in Tamil. Even some of the features that are different between the two languages (see (i)–(vii) above) have resulted from contact-induced changes. At micro-linguistic level, the two morphosyntactic phenomena, the occurrence of null arguments and the constructions which assign exhaustive focus, show that the syntactic and semantic properties of these phenomena are almost identical. Overall, the features examined, some of which display striking structural isomorphism to the point where they can be intertranslatable word by word, if not morpheme by morpheme, provide overwhelming evidence for the convergence between Sinhala and Tamil, that is, the core morphosyntax of colloquial (spoken) Sinhala has undergone extensive restructuring on the model of Tamil.

The variety of Tamil spoken in Sri Lanka, by and large, has the same features as Tamil generally and it shares to a large extent these features with other major Dravidian languages, but does not share them with NIA languages except the areal features which the languages of South Asia share. Sinhala, on the other hand, shares the majority of these features with Tamil, but does not share them with other NIA languages except the areal features. This indicates that Sinhala has converged towards Tamil and diverged from other NIA languages which it is believed to belong. The wide scale convergence between these two genetically unrelated languages and, more importantly, the differences between some features resulting from post-contact changes occurred in Sinhala, as shown in the analyses, are due neither to these two languages sharing language universals nor to coincidence, but to the changes which have occurred in Sinhala induced by its contact with Tamil. In Schiffman's (2010: 753) view, 'it seems that Sri Lanka is in fact a microcosm of the whole South Asian Linguistic area' because of 'other interesting Sprachbund topics that Sri Lanka displays (Vedda Creole, Sri Lanka Malay, Indo-Portuguese).' 'One issue,' according to Schiffman (2010), 'that needs to be dealt with in an article on South Asian linguistics is how Sinhala, a language separated widely from other Indo-Aryan languages, has been strongly influenced by Tamil' (ibid.: 753). This issue remains central to the Sir Lankan Sprachbund among others Schiffman mentions. This dissertation, it is expected, will fulfil that long-felt need.

Appendix

Word Order										
		Dravidian			New Indo-Aryan					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
2.2.1	Order of Subject, Object and Verb (81A)	SOV	SOV	SOV	SOV	SOV	SOV	SOV	SOV	SOV
2.2.2	Order of Subject and Verb (82A)	SV	SV	SV	SV	SV	SV	SV	SV	SV
2.2.3	Order of Object and Verb (83A)	OV	OV	OV	OV	OV	OV	OV	OV	OV
2.2.4	Order of Object, Oblique, and Verb (84A)	XOV			XOV	XOV				
2.2.5	Order of Adposition and Noun Phrase (85A)	Post-positions		Post-positions	Post-positions		Post-positions	Post-positions		Post-positions
2.2.6	Order of Genitive and Noun (86A)	Genitive-Noun	Genitive-Noun	Genitive-Noun	Genitive-Noun		Genitive-Noun	Genitive-Noun		
2.2.7	Order of Adjective and Noun (87A)	Adjective-Noun	Adjective-Noun	Adjective-Noun	Adjective-Noun		Adjective-Noun	Adjective-Noun		Adjective-Noun
2.2.8	Order of Demonstrative and Noun (88A)	Demonstrative-Noun	Demonstrative-Noun	Demonstrative-Noun	Demonstrative-Noun		Demonstrative-Noun	Demonstrative-Noun		Demonstrative-Noun

Word Order contd.										
		Dravidian			**New Indo-Aryan**					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
2.2.9	Order of Numeral and Noun (89A)	Numeral-Noun	Numeral-Noun	Numeral-Noun	Noun-Numeral		Numeral-Noun	Numeral-Noun	No dominant order	Numeral-Noun
2.2.10	Order of Relative Clause and Noun (90A)	Relative clause-Noun	Relative clause-Noun	Relative clause-Noun	Relative clause-Noun		Cor-relative	Relative clause-Noun	Cor-relative	
2.2.11	Order of Degree Word and Adjective (91A)	Degree word-Adjective		Degree word-Adjective	Degree word-Adjective		Degree word-Adjective	Degree word-Adjective		Degree word-Adjective
2.2.12	Position of Polar Question Particles (92A)	Final (a clitic)	Final	No question particle	Final		Initial	Final	Final	No question particle
2.2.13	Position of Interrogative Phrases in Content Questions (93A)	Not initial interrogative phrase	Not initial interrogative phrase	Not initial interrogative phrase	Not initial interrogative phrase		Not initial interrogative phrase	Not initial interrogative phrase	Not initial interrogative phrase	Not initial interrogative phrase
2.2.14	Order of Adverbial Subordinator and Clause (94A)	Subordinating Suffix		Mixed	Subordinating Suffix		Initial subordinator word	Subordinating suffix		Initial subordinator word

Word Order contd.

		Dravidian			New Indo-Aryan					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
2.2.15	Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase (95A)	OV and Postpositions		OV and Postpositions	OV and Postpositions		OV and Postpositions	OV and Postpositions		OV and Postpositions
2.2.16	Relationship between the Order of Object and Verb and the Order of Relative Clause and Noun (96A)	OV and RelN	OV and RelN	OV and RelN	OV and RelN		Other	OV and RelN	Other	
2.2.17	Relationship between the Order of Object and Verb and the Order of Adjective and Noun (97A)	OV and AdjN	OV and AdjN	OV and AdjN	OV and AdjN		OV and AdjN	OV and AdjN		OV and AdjN

Case and Person Marking										
		Dravidian			**New Indo-Aryan**					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
3.2.1	Alignment of Case Marking of Full Noun Phrases (98A)	Nominative-accusative (standard)			Nominativ-accusative (standard)		Tripartite	Tripartite		
3.2.2	Alignment of Case Marking of Pronouns (99A)	Nominative-accusative (standard)			Nominativ-accusative (standard)		Tripartite	Nominative-accusative (standard)		
3.2.3	Alignment of Verbal Person Marking (100A)	Accusative alignment			Accusative alignment		Accusative			
3.2.4	Expression of Pronominal Subjects (101A)	Pronominal subjects are expressed by affixes on verbs	Optional pronouns in subject position		Optional pronouns in subject position					
3.2.5	Verbal Person Marking (102A)	Only the A argument			No person marking of any sort		Only the A argument			

Case and Person Marking Contd.

		Dravidian			New Indo-Aryan					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
3.2.6	Third Person Zero of Verbal Person Marking (103A)	No zero realisation of third person			No person marking of any sort					
3.2.7	Order of Person Markers on the Verb (104A)	A and P do not, or do not both, occur on the verb			A and P do not, or do not both, occur on the verb		A and P do not or do not both occur on the verb			
3.2.8	Ditransitive constructions: the verb 'give' (105A)	Indirect-object construction	Indirect-object construction		Indirect-object construction		Indirect-object construction	Indirect-object construction		

Valency and Voice										
		Dravidian			New Indo-Aryan					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
3.3.1	Reciprocal Constructions (106A)	Distinct from reflexive			Distinct from reflexive		Distinct from reflexive			
3.3.2	Passive Constructions (107A)	Present			Present		Present			
3.3.3	Antipassive Constructions (108A)	No antipassive			No antipassive		No antipassive			
3.3.4	Applicative Constructions (109A)	Benefactive object only and transitive base only			Benefactive object only and transitive base only		No applicative construction			
3.3.5	Periphrastic Causative Constructions (110A)	Purposive but no sequential			Purposive but no sequential		Purposive but no sequential	Purposive but no sequential		
3.3.6	Nonperiphrastic Causative Constructions (111A)	Morpho-logical but no compound	Morpho-logical but no compound	Morpho-logical but no compound	Morpho-logical but no compound		Morpho-logical but no compound	Morpho-logical but no compound		Morpho-logical but no compound

Negation and Questions										
		Dravidian			New Indo-Aryan					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
3.4.1	Negative Morphemes (112A)	Negative particle	Negative affix	Negative affix	Negative particle	Negative particle	Negative particle		Negative affix	Negative particle
3.4.2	Order of Negative Morpheme and Verb (143A)	VNeg	[V-Neg]	[V-Neg]	VNeg	VNeg	NegV	VNeg	[V-Neg]	NegV
	Preverbal Negative Morphemes (143E)	No preverbal negative morpheme	None	None	Negative prefix	None	NegV	None	None	NegV
3.4.3	Position of Negative Morpheme with respect to Subject, Object, and Verb (144A)	SOVNeg	MorphNeg	MorphNeg	SOVNeg	SOVNeg	SONegV	SOVNeg	Morph-Neg	SONegV
3.4.4	Symmetric and Asymmetric Standard Negation (113A)	Type Asy	Both		Type SymAsy		Both			

Negation and Questions Contd.										
		Dravidian			**New Indo-Aryan**					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
3.4.5	Subtypes of Asymmetric Standard Negation (114A)	A/Fin and A/Cat	A/Fin and A/Cat		A/Fin		A/Cat			
3.4.6	Negative Indefinite Pronouns and Predicate Negation (115A)	Predicate negation also present	Predicate negation also present		Predicate negation also present		Predicate negation also present	Predicate negation also present		Predicate negation also present
3.4.7	Polar Questions (116A)	Interrogative verb morphology	Question particle	Interrogative verb morphology	Question particle		Question particle	Question particle	Question particle	Interrogative intonation only

Predication

		Dravidian			**New Indo-Aryan**					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
3.5.1	Predicative Possession (117A)	Locational/ Dative			Locational/ Dative	Genitive	Locational			
3.5.2	Predicative Adjectives (118A)	Nonverbal encoding	Nonverbal encoding	Nonverbal encoding	Nonverbal encoding	Nonverbal encoding	Nonverbal encoding			
3.5.3	Nominal and Locational Predication (119A)	Different	Different	Identical	Different	Identical	Identical			
3.5.4	Zero Copula for Predicate Nominals (120A)	Possible	Impossible	Possible	Possible	Possible	Impossible			
3.5.5	Comparative Constructions (121A)	Locational		Locational	Locational		Locational	Locational		

Complex Sentences										
		Dravidian			**New Indo-Aryan**					
Section No.	**WALS Feature**	**Tamil**	**Malayalam**	**Telugu**	**Sinhala**	**Bengali**	**Hindi**	**Marathi**	**Oriya**	**Gujarati**
4.2.1	Relativization on Subjects (122A)	Gap strategy			Gap strategy		Non-reduction			
4.2.2	Relativization on Obliques (123A)	Gap strategy			Gap strategy		Non-reduction			
4.2.3	‘Want’ Complement Subjects (124A)	Subject is left implicit			Subject is left implicit	Subject is left implicit	Subject is left implicit	Subject is left implicit		
4.2.4	Purpose Clauses (125A)	Deranked	Deranked		Deranked			Deranked		
4.2.5	‘When’ Clause (126A)	Deranked			Deranked		Balanced/ deranked	Balanced/ deranked		
4.2.6	Reason Clauses (127A)	Deranked	Balanced/ deranked		Deranked		Balanced/ deranked	Balanced/ deranked		
4.2.7	Utterance Complement Clauses (128A)	Balanced	Balanced		Balanced		Balanced	Balanced		

References

Aikhenvald, A. Y. and R. M. W. Dixon. 2001a. Introduction. In A. Y. Aikhenvald and R. M. W. Dixon (eds.), 1–26.

Aikhenvald, A. Y. and R. M. W. Dixon (eds.). 2001b. *Areal Diffusion and Genetic Inheritance*. Oxford: Oxford University Press.

Aikhenvald , A. Y. 2006. Serial Verb Constructions in Typological Perspective. In A. Y. Aikhenvald and R. M. W. Dixon (eds.). *Serial Verb Constructions: A Cross-Linguistic Typology*. Oxford: Oxford University Press, 1–68.

Aissen, J. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory*, Vol. 21, 3, 435–448.

Amritavalli, R. 2003. Question and Negative Polarity in the Disjunction Phrase. *Syntax* 6: 1, 1–18.

Amritavalli, R. and K. A. Jayaseelan. 2005. Finiteness and Negation in Dravidian. In G. Cinque and R. Kayne (eds.). *The Oxford Handbook of Comparative Syntax*. New York: Oxford University Press, 178–220.

Anderson, J. M. 2002. Historical Linguistics. In K. Malmkjær (ed.), 218–247.

Annamalai, E. 2004. The Constituent Structure of Tamil. In R. Singh (ed.). *The Yearbook of South Asian Languages and Linguistics 2003*. Berlin: Mouton de Gruyter, 3–46.

Ariel, M. 2001. Accessibility Theory: an Overview. In T. Sanders, J. Schliperoord and W. Spooren (eds.). *Text Representation*. Amsterdam: John Benjamins, 29–87.

Asher, R. E. 1982. *Tamil*. Lingua Descriptive Studies, Vol. 7. Amsterdam: North-Holland.

Asher, R. E. 1985. *Tamil* (Croom Helm Descriptive Grammars). London: Croom Helm.

Baker, M. C. 1996. *The Polysynthesis Parameter*. Oxford, New York: Oxford University Press.

Baker, M. C. 2008. The Macroparameter in a Microparametric World. In T. Biberauer (ed.). *The Limits of Syntactic Variation.* Amsterdam: Benjamins, 351–74.

Bakker, P. 2006. The Sri Lanka *Sprachbund*: The Newcomers Portuguese and Malay. In Y. Matras, A. McMahon and N. Vincent (eds.). *Linguistic Areas Convergence in Historical and Typological Perspective*. New York: Palgrave Macmillan, 135–159.

Barbosa, P. (to appear). *Pro* as a Minimal NP: Towards a Unified Approach to *pro*-drop, 1–61.

Biberauer, T., G. Newton and M. Sheehan. 2009. The Final-Over-Final Constraint and Predictions for Diachronic Change. *Toronto Working Papers in Linguistics (TWPL)*, Vol. 31, 1–17.

Biberauer, T., A. Holmberg, I. Roberts and M. Sheehan (eds.). 2010. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.

Bonta, S. 2010. Negombo Fishermen's Tamil: A Case of Indo-Aryan Contact-Induced Change in a Dravidian Dialect. *Anthropological Linguistics*. Vol. 52, 3/4, 310–343.

Brown, E. K. 2002. Generative Grammar. In K. Malmkjær (ed.), 171–192.

Brown, K. and S. Ogilvie (eds.). 2010. Introduction. In K. Brown and S. Ogilvie (eds.), xvii–xix.

Brown, K. and S. Ogilvie (eds.). 2010. *Concise Encyclopedia of Languages of the World*. Burlington: Elsevier Science.

Camacho, J. A. 2013. *Null Subjects*. Cambridge University Press: Cambridge.

Campbell, L. 2010. Areal Linguistics. In K. Brown and S. Ogilvie (eds.), 62–68.

Chandralal, D. 2010. *Sinhala*. The Netherlands and USA: John Benjamins.

Chomsky, N. 1971. Deep Structure, Surface Structure, and Semantic Interpretation. In L. Jakobovits and D. Steinberg (eds.). *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge: University Press, 183–216.

Chomsky, N. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chou, C.-T. T. and S. Hettiarachchi. 2012. Sinhala Case-marking, Modality and A-movement. Paper presented at GLOW in Asia IX, Mia University, Japan, 1–9. Available at http://faculty.human.mie-u.ac.jp/~glow_mie/IX_Proceedings_Poster/05Chou&Hettiarachchi.pdf.

Coates, W. A. 1972. Review of *Colloquial Sinhalese Clause Structures* by J. W. Gair. *Language*. Vol. 48, 2, 463–475.

Cole, M. 2009. Null Subjects: A Reanalysis of the Data. *Linguistics* 47.3, 559–587.

Cole, M. 2010. Thematic Null Subjects and Accessibility. *Studia Linguistica* 64(3), 271–320.

Comrie, B. 1989. *Language Universals and Linguistic Typology*. 2nd ed. Chicago: The University of Chicago Press.

Comrie, B. 2013a. Alignment of Case Marking of Full Noun Phrases. In M. S. Dryer and M. Haspelmath (eds.).

Comrie, B. 2013b. Alignment of Case Marking of Pronouns. In M. S. Dryer and M. Haspelmath (eds.).

Comrie, B. and T. Kuteva. 2013a. Relativization Strategies. In M. S. Dryer and M. Haspelmath (eds.).

Comrie, B. and T. Kuteva. 2013b. Relativization on Subjects. In M. S. Dryer and M. Haspelmath (eds.).

Comrie, B. and T. Kuteva. 2013c. Relativization on Obliques. In M. S. Dryer and M. Haspelmath (eds.).

Coperahewa, S. 2007. Language Contact and Linguistic Area: Sinhala – Tamil Contact Situation. *Journal of the Royal Asiatic Society of Sri Lanka*, New Series, Vol. 53, 133–152.

Cristofaro, S. 2010. Language Universals and Linguistic Knowledge. In J. J. Song (ed.), 227–249.

Cristofaro, S. 2013a. Purpose Clauses. In M. S. Dryer and M. Haspelmath (eds.).

Cristofaro, S. 2013b. ‘When’ Clauses. In M. S. Dryer and M. Haspelmath (eds.).

Cristofaro, S. 2013c. Reason Clauses. In M. S. Dryer and M. Haspelmath (eds.).

Cristofaro, S. 2013d. Utterance Complement Clauses. In M. S. Dryer and M. Haspelmath (eds.).

Croft, W. 2003. *Typology and Universals*. 2nd ed. Cambridge: Cambridge University Press.

Crystal, D. 2008. *A Dictionary of Linguistics and Phonetics*. 6th ed. USA, UK and Australia: Blackwell Publishing.

Dayal, V. and A. Mahajan (eds.). 2004. *Clause Structure in South Asian Languages*. Boston, Dordrecht, London: Kluwer Academic Publishers.

de Haan, F. 2013a. Semantic Distinctions of Evidentiality. In M. S. Dryer and M. Haspelmath (eds.).

de Haan, F. 2013b. Coding of Evidentiality. In M. S. Dryer and M. Haspelmath (eds.).

de Silva, N. 2001. Elu Langauge and Ancient Sri Lanka. Available at http://www.rootsweb.ancestry.com/~lkawgw/elu.html. Accessed on 24. 10. 2015.

Disanayaka, J. B. 1994. Sinhala. In R. E. Asher and J. M. Y. Simpson (eds.). *The Encyclopedia of Language and Linguistics*. Vol. 7. 1st ed. Oxford, New York, Seoul and Tokyo: Pergamon Press, 3949–3951.

Disanayaka, J. B. 2007. *Languages in Contact: Sinhala and Tamil*. Fifth Vesak Commemoration Lecture (delivered on May 04, 2007). Chennai, Tamil Nadu: Russian Center of Science and Culture.

Dryer, M. S. 1988. Object-verb Order and Adjective-noun Order: Dispelling a Myth. *Lingua. 74,* 185–217.

Dryer, M. S. 1992. The Greenbergian Word Order Correlations. *Language* 68, 81–138.

Dryer, M. S. 2013a. Order of Subject, Object and Verb. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013b. Order of Subject and Verb. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013c. Order of Object and Verb. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013d. Order of Adposition and Noun Phrase. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013e. Order of Genitive and Noun. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013f. Order of Adjective and Noun. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013g. Order of Demonstrative and Noun. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013h. Order of Numeral and Noun. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013i. Order of Degree Word and Adjective. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013j. Order of Relative Clause and Noun. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013k. Position of Polar Question Particles. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013l. Position of Interrogative Phrases in Content Questions. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013m. Order of Adverbial Subordinator and Clause. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013n. Relationship between the Order of Object and Verb and the Order of Adposition andNoun Phrase. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013o. Relationship between the Order of Object and Verb and the Order of Relative Clause and Noun. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013p. Relationship between the Order of Object and Verb and the Order of Adjective and Noun. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013q. Expression of Pronominal Subjects. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013r. Negative Morphemes. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013s. Order of Negative Morpheme and Verb. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013t. Position of Negative Morpheme with respect to Subject, Object, and Verb. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. 2013u. Polar Questions. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. (with O. D. Gensler). 2013. Order of Object, Oblique, and Verb. In M. S. Dryer and M. Haspelmath (eds.).

Dryer, M. S. and M. Haspelmath (eds.). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info. Accessed on 15. 10. 2015.

Ebert, K. 2010. South Asia as a Linguistic Area. In K. Brown and S. Ogilvie (eds.), 995–1001.

É.Kiss, K. 1998. Identificational Focus versus Information Focus. *Language* 7: 2, 245–273.

Emeneau, M. B. 1956. India as a Linguistic Area. *Language.* Vol. 32, 3–16.

Gair, J. W. 1980. Adaptation and Naturalization in a Linguistic Area: Sinhala Focused Sentences. *Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society*, 28–43. Revised version. Sinhala Focused Sentences: Naturalization of a Calque. In B. C. Lust (ed.), 155–169.

Gair, J. W. 1985. How Dravidianized Was Sinhala Phonology? Some Conclusions and Cautions. *Oceanic Linguistics Special Publications*. 20, For Gordon H. Fairbanks, 37–55. Also in B. C. Lust (ed.), 185–198.

Gair, J. W. 1998a. Sinhala, an Indo-Aryan Isolate. In B. C. Lust (ed.), 1–12.

Gair, J. W. 1998b. Colloquial Sinhala Inflectional Categories and Parts of Speech. In B. C. Lust (ed.), 13–24.

Gair, J. W. 1998c. Nonconfigurationality, Movement, and Sinhala Focus. In B. C. Lust (ed.), 50–64.

Gair, J. W. 1998d. Selections from “The Verb in Sinhala, with Some Preliminary Remarks on Dravidianization”. In B. C. Lust (ed.), 200–209.

Gair, J. W. 1998e. Syntactic Theory, AGR and Sinhala Diglossia. In B. C. Lust (ed.), 237–267.

Gair, J. W. 2005. Some Aspects of Finiteness and Control in Sinhala. In R. Singh and T. Bhattacharya (eds.). *The Yearbook of South Asian Languages and Linguistics*. Berlin and New York: Mouton De Gruyter, 116–143.

Gair, J. W. 2007. Sinhala. In J. Danesh and G. Cardona (eds.). *Indo-Aryan Languages*. Hoboken: Taylor & Francis, 847–904.

Gair, J. W. 2010. Sinhala. In K. Brown and S. Ogilvie (eds.), 964–968.

Gair, J. W. and W. S. Karunatillake. 2010. Lexical Pronouns and Anaphors in Sinhala. In B. C. Lust, K. Wali, J. W. Gair and K. V. Subbaro (eds.). *Lexical Anaphors and Pronouns in Selected South Asian Languages: a Principled Typology*. Berlin, New York: Mount de Gruyter, 715 – 773.

Gair, J. W. and J. C. Paolillo. 1988. Sinhala Non-verbal Sentences and Argument Structure. In V. Srivastav, J. W. Gair and K. Wali (eds.), 39–77.

Gair, J. W. and L. Sumangala. 1991. What to focus in Sinhala. *Proceedings of the Eighth Eastern States Conference on Linguistics*. Ithaca, NY: CLC Publications, Cornell University, 93–108.

Geiger, W. 1937. The Linguistic Character of Sinhalese. In *The Journal of the Royal Asiatic Society, Ceylon Branch.*Vol. 34. 90, 16–43.

Geiger, W. 1938. *A Grammar of the Sinhalese Language*. Colombo: The Royal Asiatic Society, Ceylon Branch.

GnanaPrakasar, S. 1937. The Dravidian Element in Sinhalese. *Anthropos,* Bd. 32, H. 1/2, 155–170.

Godakumbura, C. E. 1946. The Dravidian Element in Sinhalese. *Bulletin of Oriental and African Studies*. 11, 837–841.

Greenberg, J. H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In J. H. Greenberg (ed.). *Universals of Language*. London: Mit Press, 73–113.

Gumperz, J. J. and R. Wilson. 1971. Convergence and Creolization: a Case from the Indo-Aryan/Dravidian Border. In D. Hymes (ed.). *Pidginization and Creolization of Languages.* Cambridge: Cambridge University Press, 151–168.

Hagstrom, P. 2004. Particle Movement in Sinhala and Japanese. In V. Dayal and A. Mahajan (eds.), 227–252.

Haspelmath, M. 2013a. Ditransitive Constructions: the Verb ‘give’. In M. S. Dryer and M. Haspelmath (eds.).

Haspelmath, M. 2013b. Negative Indefinite Pronouns and Predicate Negation. In M. S. Dryer and M. Haspelmath (eds.).

Haspelmath, M. 2013c. Indefinite Pronouns. In M. S. Dryer and M. Haspelmath (eds.).

Haspelmath, M. 2013d. ‘Want’ Complement Subjects. In M. S. Dryer and M. Haspelmath (eds.).

Haspelmath, M., M. S. Dryer, D. Gil and B. Comrie. 2005a. Introduction. In M. Haspelmath, M. S. Dryer, D. Gil and B. Comrie. (eds.), 1–8.

Haspelmath, M., M. S. Dryer, D. Gil and B. Comrie (eds.). 2005b. *The World Atlas of Language Structures.* Oxford: Oxford University Press.

Heine, B. and T. Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Heine, B. and T. Kuteva. 2007. Identifying Instances of Contact-induced Grammatical Replication. Paper presented at the Symposium *Language Contact and the Dynamics of Language: Theory and Implications*, Max Planck Institute for Evolutionary Anthropology, Leipzig, 10 –13 May. Preliminary Version, 1–30. Available at http://www.unice.fr/ChaireIUF-Nicolai/Archives/Symposium/ Symposium_Textes/Heine-Kuteva_Leipzig07.pdf.

Heine, B. and T. Kuteva. 2008. Constraints on Contact-induced Linguistic Change. *Journal of Language Contact*, THEMA 2, 57–90.

Heine, B. and T. Kuteva. 2010. Contact and Grammaticalization. In R. Hickey (ed.), 86–105.

Heine, B. and T. Kuteva. 2011. The Areal Dimension of Grammaticalzation. In H. Narrog and B. Heine (eds.), 291–301.

Henadeerage, D. K. 1998. Anaphoric Binding in Colloquial Sinhala. In M. Butt and T. H. King (eds.). *Proceedings of the LFG98 Conference.* The University of Queensland, Brisbane: CSLI Publications.

Henadeerage, D. K. 2002. *Topics in Sinhala Syntax*. Unpublished PhD Thesis. The Australian National University.

Herring, S. C. 1993. Aspectogenesis in South Dravidian: On the Origin of the ‘Compound Continuative’ *KONTIRU*. In H. Aertsen and R. J. Jeffers. (eds.). *Historical Linguistics* 1989. Amsterdam: John Benjamins, 167–185.

Herring, S. C. and J. C. Paolillo. 1995. Focus Position in SOV Languages. In P. Downing and M. Noonan (eds.). *Word Order in Discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 163–198.

Hettiaratchi, D. E. 1959. The Languages of Ceylon. In S. Paranavithana (ed.). *The History of Ceylon Volume* 1, Colombo: Ceylon University board, 33–45.

Hickey, R. 2010. Language Contact: Reconsideration and Reassessment. In R. Hickey (ed.), 1–28.

Hickey, R. (ed.). 2010. *The Handbook of Language Contact*. UK: Wiley Blackwell, 1–28.

Holmberg, A. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36, 533–64.

Holmberg, A. 2010. Null Subject Parameters. In T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan (eds.), 88–124.

Holmberg, A. 2016. The Syntax of Yes and No. Oxford: Oxford University Press.

Holmberg, A. To appear. Linguistic Typology. In I. Roberts (ed.) *Handbook of Universal Grammar*. Cambridge: Cambridge University Press, 1–23.

Holmberg, A., A. Nayudu and M. Sheehan. 2009. Three Partial Null-subject Languages: a Comparison of Brazilian Portuguese, Finnish, and Marathi. *Studia Linguistica* 63, 59–97.

Holmberg, A. and O. Phimsawat. 2015. Generic Pronouns and phi-features: Evidence from Thai. *Newcastle and Northumbria Working Papers in Linguistics* 21.1, 55 –71.

Huang, J. 1984. On the Distribution and Reference of Empty Pronouns. *Linguistic Inquiry* 15, 531–574.

Jany, C. 2006. The Relationship between Case Marking and S, A and O in Spoken Sinhala. In R. Englebretson and C. Genetti (eds.). *Santara Barbara Papers in Linguistics.* Vol. 17, 68–84.

Jayaseelan, K. A. 1996. Question-word Movement to Focus and Scrambling in Malayalam. *Linguistic Analysis* 26/27, 63–83.

Jayaseelan K. A. 1999. Empty Pronouns in Dravidian. In K. A. Jayaseelan (ed.). *Parametric Studies in Malayalam Syntax*. New Delhi: Allied Publishers, 14–25.

Jayaseelan, K. A. 2001a. IP-internal Topic and Focus Phrases. *Studia Linguistica* 55.1, 39–75.

Jayaseelan, K. A. 2001b. Question and Question-word Incorporating Quantifiers in Malayalam. *Syntax* 4 (2), 63–93.

Jayaseelan, K. A. 2003. Question Words in Focus Positions. In J. Rooryck and P. Pica. (eds.). *Linguisic Variation Yearbook* 3. Amsterdam: John Benjamins, 69–99.

Jayaseelan, K. A. 2004. The Serial Verb Construction in Malayalam. In V. Dayal and A. Mahajan. (eds.), 67–91.

Jayaseelan, K. A. 2008. Topic, Focus and Adverb Positions in Clause Structure. *Nanzan Linguistics* 4, 43–68.

Jayaseelan, K. A. 2010. Stacking, Stranding, and Pied-piping: A Proposal about Word Order. *Syntax* 13(4), 298–330.

Jayaseelan, K. A. 2014. Coordination, Relativization and Finiteness in Dravidian. *Natural Language and Linguistic Theory*. 32. 191–211.

Jayaseelan, K. A. and R. Amritavalli. 2005. Scrambling in the Cleft Construction in Dravidian. In J. Sabel and M. Saito (eds.). *The Free Word Order Phenomenon: Its Syntactic Sources and Diversity*. Berlin: De Gruyter Mouton. 137–161.

Kishimoto, H. 2005. *Wh*-in-situ and Movement in Sinhala Questions. *Natural Language and Linguistic Theory* 23, 1–51.

Koptjevskaja-Tamm, M. 2011. Linguistic Typology and Language Contact. In J. J. Song. (ed.), 568–590.

Krishnamurti, B. 2003. *The Dravidian Languages*. Cambridge and New York: Cambridge University Press.

Lahiri, U. 1998. Focus and Negative Polarity in Hindi. *Natural Language Semantics* 6, 57–123.

Lehmann, T. 1989. *A Grammar of Modern Tamil*. India: Pondicherry Institute of Linguistics and Culture.

Lehmann, W. L. 1978. The Great Underlying Ground-Plans. In W. L. Lehmann. (ed.). *Syntactic Typology: Studies in the Phenomenology of Language*. Great Britain: The Harvester Press, 3–55.

Li, C. N. and S. A. Thompson. 1976. Subject and Topic: A New Typology of Language. In C. N. Li. (ed.). *Subject and Topic*. New York: Academic Press, Inc, 457–489.

Lucas, C. 2015. Contact-induced Language Change. In C. Bowern and B. Evans (eds.). *The Routledge Handbook of Historical Linguistics*. London: Routledge, 519–536. Final pre-publication version available at http://www.academia.edu/7619365/Contact-induced_language_change, 1–18.

Lust, B. C. (ed.). 1998. *Studies in South Asian Linguistics: Sinhala and Other South Asian Languages*. New York and Oxford: Oxford University Press.

Malmkjær, K. 2002a. Language Typology. In K. Malmkjær (ed.), 319–324.

Malmkjær, K. 2002b. Language Universals. In K. Malmkjær (ed.), 324–331.

Malmkjær, K. (ed.). 2002c. *The Linguistics Encyclopedia*. 2nd ed. London: Routledge.

Masica, C. P. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.

Maslova, E. and V. P. Nedjalkov. 2013. Reciprocal Constructions. In M. S. Dryer and M. Haspelmath (eds.).

Matras, Y. 2009. *Language Contact*. Cambridge: Cambridge University Press.

Matras, Y. 2010. Contact, Convergence and Typology. In R. Hickey. (ed.), 66–85.

Matras, Y. 2011. Grammaticalization and Language Contact. In H. Narrog and B. Heine (eds.), 279–290.

Matras, Y. and J. Sakel. 2007. Investigating the Mechanisms of Pattern-replication in Language Convergence. *Studies in Language* 31: 4, 829–865.

Miestamo, M. 2013a. Symmetric and Asymmetric Standard Negation. In M. S. Dryer and M. Haspelmath. (eds.).

Miestamo, M. 2013b. Subtypes of Asymmetric Standard Negation. In M. S. Dryer and M. Haspelmath. (eds.).

Modesto, M. 2008. Topic Prominence and Null Nubjects. In T. Biberauer. (ed.). *The Limits of Syntactic Variation*. Amsterdam: Benjamins, 375–410.

Moravcsik, E. A. 2011. Explaining Language Universals. In J. J. Song. (ed.), 69–89.

Munshi, S. 2010. Indo-Aryan Languages. In K. Brown and S. Ogilvie. (eds.), 522–528.

Myers-Scotton, C. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. UK and US: Oxford University Press.

Narrog, H. and B. Heine. (eds.). 2011. *The Oxford Handbook of Grammaticalization*, Oxford: Oxford University Press.

Nayudu, A. 2008. *Issues in the Syntax of Marathi: A Minimalist Approach*. PhD Dissertation. Durham: Durham University.

Noonan, M. 2010. Genetic Classification and Language Contact. In R. Hickey (ed.), 48–65.

Palmer, F. R. 2001. *Mood and Modality*. 2nd ed. Cambridge University Press.

Paolillo, J. C. 1994. The Co-development of Finiteness and Focus in Sinhala. In *Perspectives on Grammaticalization*. W. Pagliuca. (ed.). Amsterdam and Philadelphia: John Benjamins Publishing Company, 151–170.

Pereltsvaig, A. 2012. *Languages of the World*. Cambridge: Cambridge University Press.

Polinsky, M. 2013a. Antipassive Constructions. In M. S. Dryer and M. Haspelmath (eds.).

Polinsky, M. 2013b. Applicative Constructions. In M. S. Dryer and M. Haspelmath (eds.).

Portner, P. 2009. *Modality*. Oxford: Oxford University Press.

Ramchand, G. C. 2011. Licensing of Instrumental Case in Hindi/Urdu Causatives. *Nordlyd* 38, 49–85.

Rizzi, L. 1986. Null Objects in Italian and the Theory of *pro*. *Linguistic Inquiry* 17, 501–557.

Rizzi, L. 1997. The Fine Structure of the Left Periphery. In L. Haegeman (ed.). *Elements of Grammar*. Dordrecht: Kluwer, 281–337.

Roberts, I. 1997. *Comparative Syntax*. London, New York: Arnold.

Roberts, I. and A. Holmberg. 2010. Introduction: Parameters in Minimalist Theory. In T. Biberauer, A. Holmberg, I. Roberts, M. Sheehan (eds.), 1–57.

Ross, M. D. 1996. Contact-induced Change and the Comparataive Method: Cases from Papua New Guinea. In M. Durie and M. D. Ross. (eds). *The Comparative Method Reviewed: Regularity and Irregularity in Language Change.* New York: Oxford University Press, 180–217.

Ross, M. D. 1999. Exploring Metatypy: how does contact-induced typological change come about?. Available online athttp://www.researchgate.net/publication/251662587_Exploring_metatypy_how_does_contactinduced_typological_change_come_about, 1–32.

Ross, M. D. 2001. Contact-induced Change in Oceanic Languages in North-West Melanesia. In A. Y. Aikhenvald and R. M. W. Dixon. (eds.), 134–166.

Ross, M. D. 2003. Diagnosing Prehistoric Language Contact. In R. Hickey (ed.). *Motives for Language Change*. Cambridge: Cambridge University Press, 174–198.

Ross, M. D. 2006. Language Contact, Language Loss and Ethnosyntax: Reflections from the Pacific. Available at http://www.academia.edu/ 3404622/ Language_contact_language_loss_and_ethnosyntax_reflections_from_the_Pacific, 1–28.

Ross, M. D. 2007. Calquing and Metatypy. *Journal of Language Contact* – THEMA 1, 116– 143.

Saeed, J. I. 2003. *Semantics*. 3rd ed. Malden, MA: Blackwell Publishing.

Sarma, V. 1999. *Case, Agreement and Word Order: Issues in the Syntax and Acquisition of Tamil*. Unpublished Doctoral Dissertation, Cambridge, Massachusetts: Massachusetts Institute of Technology.

Sarma, V. 2003. Non-canonical Word Order: Topic and Focus in Adult and Child Tamil. In S. Karimi. (ed.). *Word Order and Scrambling*. USA, UK, Australia and Germany: Blackwell Publishing, 238–272.

Schiffman, H. F. 1999. The Role of Metaphor in the Grammaticalization of Aspect in Tamil. Paper presented in Panel on Spatial Metaphors in Language, American Anthropological Association Chicago. Available at file:///L:/Aspect,%20 Metaphor,%20and%20Variability%20in%20Tamil.html. Accessed on 14. 06. 2015.

Schiffman, H. F. 2004. The Tamil Case System. Available at http://ccat.sas.upenn.edu/~haroldfs/public/h_sch_9a.pdf, 302–313

Schiffman, H. F. 2010. Language Contact in South Asia. In R. Hickey (ed.), 738–756.

Selvanathan, N. 2012. Rightward-extraction and Focus in Tamil. Abstract. Paper presented at the South Asian Languages Workshop. Available at http://linguisticssouthasia.commons.yale.edu/files/2012/07/selvanathan_tamil.pdf.

Shibatani, M. 1999. Dative Subject Constructions Twenty-two Years Later. *Studies in the Linguistic Sciences*. Volume 29, 2, 45–76.

Siewierska, A. 2013a. Alignment of Verbal Person Marking. In M. S. Dryer and M. Haspelmath (eds.).

Siewierska, A. 2013b. Verbal Person Marking. In M. S. Dryer and M. Haspelmath (eds.).

Siewierska, A. 2013c. Third Person Zero of Verbal Person Marking. In M. S. Dryer and M. Haspelmath (eds.).

Siewierska, A. 2013d. Order of Person Markers on the Verb. In M. S. Dryer and M. Haspelmath (eds.).

Siewierska, A. 2013e. Passive Constructions. In M. S. Dryer and M. Haspelmath. (eds).

Simpson, A. and S. Syed (to appear). Blocking Effects of Higher Numerals in Bangla: a Phase-based Analysis. *Linguistic Inquiry*, 1–13.

Slade, B. 2013. Question-particles and Relative Clauses in the History of Sinhala with Comparison to Early and Modern Dravidian**.** In S. Chen and B. Slade (eds.). *Grammatica et verba / Glamour and verve — Studies in South Asian, Historical, and Indo-European Linguistics: A Festschrift in Honor of Professor Hans*

Henrich Hock on the Occasion of His 75th Birthday. Ann Arbor: Beech Stave Press: 245–268. Pre-print pdf. Available at http://ling.auf.net/ lingbuzz/002196, 1–23

Song, J. J. 2013a. Periphrastic Causative Constructions. In M. S. Dryer and M. Haspelmath (eds).

Song, J. J. 2013b. Nonperiphrastic Causative Constructions. In M. S. Dryer and M. Haspelmath. (eds).

Southworth, F. C. 1971. Detecting prior Creolization: an Analysis of the Historical Origin of Marathi. In D. Hymes (ed.). *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press, 255–273.

Srivastav, V., J. W. Gair and K. Wali (eds.). *Cornell Working Papers in Linguistics 8*. Ithaca, NY: Department of Modern Languages and Linguistics, Cornell University.

Stassen, L. 2013a. Predicative Possession. In M. S. Dryer and M. Haspelmath (eds.).

Stassen, L. 2013b. Predicative Adjectives. In M. S. Dryer and M. Haspelmath (eds.).

Stassen, L. 2013c. Nominal and Locational Predication. In M. S. Dryer and M. Haspelmath (eds.).

Stassen, L. 2013d. Zero Copula for Predicate Nominals. In M. S. Dryer and M. Haspelmath (eds.).

Stassen, L. 2013e. Comparative Constructions. In M. S. Dryer and M. Haspelmath (eds.).

Steever, S. B. 1988. *The Serial Verb Formation in the Dravidian Languages*. Delhi: Motilal Banarsidass.

Subbarao, K. V. 2010. Dravidian Languages: a SyntacticTypology. Special lecture delivered at the seminar on *The Linguistic and Genetic Architecture of Dravidian* at Dravidian University, Kuppam, India, 22nd–24th February 2010. Available at http://www.researchgate.net/publication/273480927, 1–60.

Subbarao, K. V. 2012. *South Asian Languages: A Syntactic Typology*. Cambridge: Cambridge University Press.

Takahashi, D. 2013. Argument Ellipsis in Japanese and Malayalam. *Nanzan Linguistics* 9, 173–192.

Tallerman, M. 2015. *Understanding Syntax*. 4th ed. London and New York: Routledge Taylor and Francis Group.

Thomason, S. G. 2000. Linguistic Areas and Language History. In D. Gilbers, J. A. Nerbonne and J. Schaeken (eds). *Studies in Slavic and General Linguistics.* Vol.

28, *Languages in Contact*. Amsterdam - Atlanta, GA: Rodopi, 311–327. Also available at http://www-personal.umich.edu/~thomason/papers/areas.pdf, 1–14.

Thomason, S. G. 2001a. *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.

Thomason, S. G. 2001b. Contact-induced Typological Change. In M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds.). *Language typology and language universals, Sprachtypologieund sprachliche Universalien: An international handbook.* Berlin and New York: Walter de Gruyter, 1640–1648. Available also at http://www-personal.umich.edu/~thomason/papers/typech.pdf, 1–11.

Thomason, S. G. 2003. Contact as a Source of Language Change. In B. D. Joseph and R. D. Janda (eds.). *The Handbook of Historical Linguistics.* USA, UK, and Australia: Blackwell Publishing Ltd., 687–712.

Thomason, S. G. and T. S. Kaufman. 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley, Los Angeles and Oxford: University of California Press.

Tomioka, S. 2003. The Semantics of Japanese Null Pronouns and Its Cross-linguistic Implications. In K. Schwabe and S. Winkler (eds.). *The Interfaces Deriving and Interpreting Omitted Structures*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 321–339.

Trask, R. L. 2007. *Key Concepts in Language and Linguistics*. 2nd ed. London, New York: Routledge.

Velupillai, V. 2012. *An Introduction to Linguistic Typology*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Winford, D. 2003. *An Introduction to Contact Linguistics*. USA, UK and Germany: Blackwell.

Yadurajan, K. S. 1988. Binding Theory and Reflexives in Dravidian. In V. Srivastav, J. W. Gair and K. Wali (eds.), 181–203.

Zwicky, A. M. and G. K. Pullum. 1983. Cliticization vs. IInflection: English *N'T*. *Language*. Vol. 59, 3, 502–513.