Santa Barbara
Papers in Linguistics

Volume 17

Proceedings from the Workshop on Sinhala Linguistics
June 3-4, 2005

Robert Englebretson and Carol Genetti, Editors

Dedication

This volume is dedicated to Oshan Fernando and Nissanka S. Wickremasinghe.

Contents

Abbreviations

1	First person		HES	Hesitation particle
2	Second person		IMP	Imperative
3	Third person		IMPF	Inperfective
A	Animal		INAN	Inanimate
ABL	Ablative		IND	Indefinite
ACC	Accusative		INF	Infinitive
ADJ	Adjective		INST	Instrumental
AGENT	Agent		INVOL	Involitive
ANIM	Animate		LOC	Locative
ASS	Assertive		M	Masculine
AUX	Auxiliary		NEG	Negative
CAUS	Causative		NOM	Nominalizer
CL	Classifier		NPST	Non-past
COMP	Complementizer		p	Plural
CONC	Concessive		PRED	Predicate
COND	Conditional		PRES	Present
CONJ	Conjunction		PRT	Particle
CONV	Converb		PRTMP	Prior temporal (verb form)
DAT	Dative		PROX	Proximal
DEF	Definite		PST	Past
DEM	Demonstrative		PL	Plural
DIST	Distal		PPL	Participle
EMPH	Emphatic		Q	Question particle
EPIST	Epistemic		QUOT	Quotative
EQ	Equative		REDUP	Reduplicative
EX	Existential		REFL	Reflexive
F	Feminine		REL	Relative
FOC	Focus		s	Singular
FUT	Future		SG	Singular
GEN	Genitive/Possessive		VIS	Visual
GOAL	Goal		VOL	Volitive

Preface

Robert Englebretson, Rice University

Carol Genetti, UCSB

The 12 working papers in this volume comprise original student research on specific aspects of spoken colloquial Sinhala. These papers were originally presented at the Workshop on Sinhala Linguistics, held June 3-4, 2005 at the University of California Santa Barbara. This conference, organized by graduate students at UCSB, represented the culmination and collaboration of two courses in field methods during the 2004-2005 academic year, one led by Carol Genetti at UCSB, and the other led by Robert Englebretson at Rice University. The workshop gave Rice and UCSB field methods students the opportunity to interact with one another, to publicly present their original research, and to receive invaluable feedback from John Paolillo (Indiana University) who also gave the keynote address.

Since each of the papers in this volume focuses on a specific aspect of Sinhala grammar, we shall begin by presenting a brief general overview of Sinhala for the benefit of readers who may be unfamiliar with this language and its background. For a more thorough grammatical description, see Gair and Paolillo (1997) *inter alia*.

Sinhala (also referred to as Singhala, Singhalese, and Sinhalese) is spoken natively by approximately 13 million speakers, primarily in the country of Sri Lanka. It is a member of the Indo-Aryan language family, and is genetically related to Hindi, Urdu, Bengali, Gujarati, and the other New Indo-Aryan languages. The exact position of Sinhala within Indo-Aryan has been a matter of debate (see Masica 1999:446-463). According to Gair and Paolillo (1997:1) together with Dhivehi (Maldivian), it forms a separate branch within Indo-Aryan.

Sinhala is one of two national languages of Sri Lanka, the other being Tamil, a member of the Dravidian language family. The ethnic Tamil and Sinhalese have been in close contact for over two millennia, so Sinhala language structures have developed under the influence of Tamil language contact. Gair and Paolillo (1997:2) note that Tamil influence is especially evident in Sinhala syntax, citing the left-branching structure of the language and the pervasive focused sentence constructions.

Sinhala is composed of two quite distinct varieties, the formal written variety and the colloquial spoken variety. The two varieties differ markedly in their core grammatical structures. They exist in a diglossic relationship (De Silva 1974, 1976, Gair 1968, 1986, 1992, Paolillo 1991, 1997). Complex codeswitching and code mixing of Sinhala and English is also common among educated Sinhalese (cf. Abeywickrama 2004). The current volume focuses exclusively on colloquial Sinhala, especially on a dialect spoken just to the west of the capital city of Colombo, as represented in the speech of our two Field Methods language consultants.

Each paper in this volume brings to fruition a specific research project undertaken by individual Field Methods students at Rice and UCSB. These contributions address aspects of colloquial Sinhala at all levels of linguistic structure—from phonetics to discourse and everything in between. Due to the close theoretical affinity between Rice and UCSB Linguistics, and the emphasis which each department places on primary data, these papers are unified in presenting an approach based on functional, cognitive, and typological perspectives. Taken

together, this volume offers an overview of relevant theoretical issues in functional linguistics as observed in colloquial Sinhala.

We shall now turn to a brief summary of each of the papers. Contributions in this volume are arranged alphabetically by the author's last name; however, for purposes of the present summary, we will discuss them topically, in terms of the general subfield of linguistics which they represent.

The contribution by Carlos Nash (UCSB) deals specifically with Sinhala phonetics and phonology. The basic phonemic inventory of Sinhala (cf. Gair and Paolillo 1997:3-4) comprises 27 consonants and 12 vowels. The consonant inventory includes a contrast among dental and retroflex obstruents, as well as a typologically-rare series of prenasalized stops. Vowel length is phonemic, and thus the 12 distinct vocalic phonemes consist of six pairs of long and short vowels. There is additionally a complex interaction among syllable types, weights, and stress, which is the general topic of Nash's paper, with particular focus on stress in Sinhala verbs. Nash explores the role of intensity and duration as the key acoustic parameters in determining syllable stress, and provides an equation to model these findings. He then develops a constraint-based phonological account based on the acoustic results and presents them within the framework of Optimality Theory.

Five of the papers in the volume address Sinhala morphology and word classes. The contributions by Garland, Henderson, and Jany address aspects of nominal morphology, while those by Hilpert and Taylor concern Sinhala verbs. Sinhala nominal morphology includes the marking of definiteness, number, and case, and is generally organized based on animacy. Sinhala verb morphology is particularly complex (cf. Gair and Paolillo 1997:23-28). Verb roots fall into several inflectional classes. There are simple, causative, and involitive root forms for each class. Each root form can take a myriad of TAM and participial suffixes. In addition to main verbs, there is also a unique grammatical category of quasi-verbs which function as the predicates of clauses but which generally do not inflect with the typical Sinhala verb morphology.

Jennifer Garland (UCSB) explores the complex expression of Sinhala nominal morphology (definiteness, number, and case-marking) as combinations of affixes, clitics, and postpositions. She demonstrates that the traditional morphological typology of synthesis and fusion does not adequately account for the observed levels of structure. Garland claims that the Sinhala system is best accounted for by recognizing the interdependence of 'phonological word' and 'grammatical word' boundaries.

Mara Henderson (UCSB) also deals with Sinhala nouns, examining the morphosyntax and semantics of 'specific-general noun sequences' (SGNs); namely, constructions consisting of a specific head noun followed by a general classificatory noun, e.g. *kehel geḍi* 'bananas' (lit. 'banana fruit'). Henderson argues that nominal classification in Sinhala lies on the typological continuum between lexico-grammatical (classifiers and measure terms) and lexical (gender/noun-class markers). Sinhala SGNs do not fit neatly into one type or the other, and they sometimes display mismatches based on semantics and morphology; thus, Henderson suggests Sinhala may illustrate a type of classification system not previously described in the literature.

Carmen Jany's (UCSB) contribution takes on the thorny issue of the interaction between grammatical relations and case-marking of Sinhala nouns. Jany illustrates that morphological case in Sinhala is not directly assigned simply based on S, A, and O roles; rather, a conspiracy of lexical and semantic factors co-occur to contribute to the case-marking of a nominal argument.

These factors include definiteness of the argument, volitivity of the verb, and in some cases even the semantics of the entire clause.

Chris Taylor (Rice) analyzes the functions of the Sinhala conjunctive participle—the form of the verb usually marked by the suffix -*la*. Taylor shows that this inflectional verb form indicates perfect aspect in some contexts, but in other contexts it functions as a prototypical South-Asian converb (cf. Haspelmath and König 1995, Genetti 2005 *inter alia*). Taylor argues that these two seemingly disparate functions are actually semantically related based on event construal, and that 'event sequencing' and 'recapitulation' provide the basis for a unified account of this multifunctional verb form.

Martin Hilpert (Rice) addresses the question of auxiliaries in Sinhala. Based on grammaticization theory (Hopper and Traugott 1993), Hilpert identifies several Sinhala forms which are typically good candidates for auxiliation cross-linguistically—the development of a word from a lexical source to a grammatical auxiliary. Using synchronic distributional evidence, Hilpert concludes that a number of forms exist in Sinhala which are justifiable as auxiliaries on morphological, syntactic, and semantic grounds. He also concludes that the category of quasi-verb (Gair and Paolillo 1997:26) can be fruitfully analyzed as consisting of two sub-classes: epistemic elements and stance elements.

You-Jing Lin (UCSB) presents a Cognitive Linguistic analysis of Sinhala spatial postpositions. Namely, she provides a case-study of how Sinhala uses vertical postpositions to encode horizontal spatial relationships. Lin presents the results of an experiment which was devised to assess the extent in which horizontal relationships among objects are conceived of in vertical terms. She then proposes an analysis based on image schemas, and discusses the role of perspective (the 'route perspective' versus the 'survey perspective') which appears to motivate when this transformation can take place, and when this transformation is blocked.

Valerie Sultan (UCSB) explores the information-structuring function of adverbial clauses in Sinhala focus constructions. Focus constructions are highly grammaticalized in Sinhala and occur frequently in our discourse data. While focus constructions are generally regarded in the typological literature as being a means of profiling referents of noun phrases, the Sinhala focus construction can also be used to profile propositions expressed by adverbial clauses. Sultan demonstrates that the conditions under which adverbial clauses are focused are the same as those under which noun phrases are focused; focused elements are those which refer to entities or events that are either new in the discourse or contradictory to the supposed beliefs of the hearer. This paper thus presents an overview of both adverbial clauses and focus constructions in Sinhala, then examines features of focused adverbial clauses in detail.

Each of the four remaining papers in the volume explicitly addresses a classic issue in linguistic typology relevant to the role of meaning in shaping grammar. The construction types addressed in these four contributions are: locationals, causatives, relative clauses, and object complements. While each of these general construction types has indeed already received substantial treatment in cross-linguistic typological literature, to our knowledge these are the first published papers to explore them within colloquial Sinhala specifically. The results of these four contributions support the larger typological findings presented in previous research, and provide a successful investigation of general typological principles as observed at work in colloquial Sinhala.

Salomé Gutierrez (UCSB) presents a study of Sinhala existential and possessive clauses, describing each as a subtype of locational constructions. Following Lyons (1968) and Clark

(1978), Gutierrez demonstrates the relatedness of the syntax and semantics of these Sinhala construction types, which thus supports the validity of the previous cross-linguistic observations.

Danielle Mathieu-Reeves (Rice) examines Sinhala causative constructions in terms of the inverse relation between grammatical complexity and semantic directness. She finds that grammatical simplicity corresponds with semantic directness, while greater complexity corresponds with semantic indirectness. Her contribution therefore serves to support Haiman's (1983) 'complexity continuum', and to demonstrate how this general iconic principle is manifest in Sinhala specifically.

Ben Walker (Rice) provides a comprehensive description of Sinhala relative clauses in terms of traditional relative clause typology (cf. Keenan 1985). Sinhala relative clauses are pre-head, are indicated by special relative verb forms, and the relativized noun phrase is gapped. Walker examines the role of case marking to disambiguate the gapped NP. He also observes that Sinhala relativizes on all positions of the 'accessibility hierarchy' (Keenan and Comrie 1977, 1979).

Ben Wheeler (Rice) uses a corpus of textual and elicited data to investigate Givón's (1980) 'binding hierarchy' for complementation. Wheeler finds that in general, Sinhala conforms to the binding hierarchy as expected: CTP's which are verbs of utterance, cognition or epistemicity tend to take less-integrated clausal complements, while implicative or modal CTP's tend to take complements that are more tightly bound into the main clause. Interestingly, while the general predictions of the binding hierarchy are confirmed, Wheeler also finds a few instances of complements which do not occur as expected. Wheeler's contribution thus provides strong empirical support for the binding hierarchy, and affirms that the hierarchy should be understood as a general typological tendency rather than as an absolute universal.

Three additional papers were presented at the workshop which are not included in this volume. Anne-Marie Hartenstein (Rice) raised the question of subjecthood in Sinhala; she outlined the results of traditional morphological and syntactic tests, in order to assess whether the grammatical category 'subject' is relevant for Sinhala grammar. Priya Abeywickrama (UCLA), who is herself a native speaker of Sinhala, discussed the language repertoires of Sinhala-English bilinguals, which she situated in terms of codeswitching and code mixing. A version of her talk also appears as Abeywickrama 2004. Finally, in his keynote address, John Paolillo (Indiana University) brought together the divergent subdisciplines of computational and field linguistics, using an electronic corpus of Sinhala texts to explore the distribution of phonemic and grammatical categories. We would like to acknowledge the unique and important contributions of each of these three talks, and we regret that they are not able to be published in this working papers volume alongside the other papers.

This volume and the workshop on which it is based owe their success to a number of individuals and institutions. First, we would like to thank each of the student participants for their contributions, and for their interest and hard work throughout the field methods courses. We especially recognize the UCSB graduate students who organized the workshop in the midst of final exams, and those who hosted the visiting students from Rice. We especially wish to thank John Paolillo for his excellent keynote, and for his expertise in Sinhala linguistics which he graciously contributed in the discussion of the student papers. Each of the contributions to this volume has benefited immensely from his input, and we appreciate his willingness to engage students in discussion during and after the workshop.

For direct financial support of the workshop, we gratefully acknowledge the following organizations: the UCSB Graduate Division; the UCSB Graduate Student Association; the

REFERENCES

ABEYWICKRAMA, PRIYANVADA. 2004. Do we codeswitch or code mix in Sri Lanka?. Paper presented at SALA 23 (South Asian Language Analysis): Austin, TX.

CLARK, EVE V. 1978. Locationals: Existential, locative, and possessive constructions. *Universals of human language*, ed. by Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik, 4.85-126. Stanford: Stanford University Press.

DE SILVA, M.W.S., 1974. Convergence in diglossia: the Singhalese situation. *International Journal of Dravidian Linguistics* 3.1:60-91.

DE SILVA, M.W.S. 1976. *Diglossia and literacy*. Mysore: Central Institute of Indian Languages.

GAIR, JAMES W. 1968. Sinhalese diglossia. *Anthropological Linguistics* 10.8.1-15.

GAIR, JAMES W. 1986. Sinhala diglossia revisited, or diglossia dies hard. *South Asian Languages: Structure Convergence and Diglossia*, ed. by Bh. Krishnamurti, Colin P. Masica, and Anjani K. Sinha, 322-336. Delhi: Motilal Banarsidass.

GAIR, JAMES W. 1992. AGR, INFL, case and Sinhala diglossia or: can linguistic theory find a home in variety? *Studies in Dravidian and general linguistics*, ed. by B. Lakshmi Bai and B. Ramakrishna Reddy, 448-467. Hyderabad: Centre for Advanced Study in Linguistics, Osmania University.

GAIR, JAMES W. and JOHN C. PAOLILLO. 1997. *Sinhala*. Munich: Lincom Europa.

GENETTI, CAROL. 2005. The participle construction of Dolakha Newar: Syntactic Implications of an Asian Converb. *Studies in Language* 29.1.35-87.

GIVÓN, TALMY. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4.3.333-377.

HAIMAN, JOHN. 1983. Iconic and economic motivation. *Language* 59.781-819.

HASPELMATH, MARTIN. and EKKEHARD KÖNIG (eds.) 1995. *Converbs in cross-linguistic perspective*. Berlin: Mouton de Gruyter.

HOPPER, PAUL and ELIZABETH CLOSS TRAUGOTT. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

KEENAN, EDWARD L. 1985. Relative Clauses. *Language typology and syntactic description*, ed. by Timothy Shopen, 2.141-170. New York: Cambridge University Press.

KEENAN, EDWARD L. and BERNARD COMRIE. 1977. NP accessibility and Universal Grammar. *Linguistic Inquiry* 8.63-100.

KEENAN, EDWARD L. and BERNARD COMRIE. 1979. Data on the noun phrase accessibility hierarchy. *Language* 55.333-351.

LYONS, JOHN. 1968. *Introduction to theoretical linguistics*. London: Cambridge University Press.

MASICA, COLIN P. 1999. *The Indo-Aryan languages*. [Cambridge Language Surveys.] Cambridge: Cambridge University Press.

PAOLILLO, JOHN C. 1991. Sinhala diglossia and the theory of government and binding. *Southwest Journal of Linguistics* 10.1.41-59.

PAOLILLO, JOHN C. 1997. Sinhala diglossic variation: continuous or discrete? *Language in Society* 26.2.269-296.

MORPHOLOGICAL TYPOLOGY AND THE COMPLEXITY OF NOMINAL MORPHOLOGY IN SINHALA

Jennifer Garland

University of California, Santa Barbara

1. Introduction. Sinhala, an Indo-Aryan language spoken in Sri Lanka by approximately 13 million people, has a complicated system of nominal morphology. Number marking on nouns in the nominative case is based on a series of twelve noun classes partially determined by animacy. The marking of definiteness and case on nouns is simpler in some respects because the shape of these markers are determined only by number and animacy without reference to the noun classes that are apparent in the system of number marking. However, in other respects the case marking paradigm is more complicated than the number marking system in that is includes both clitics and postpositions. So in order to adequately describe the case marking of nouns, it is necessary to recognize three levels of structure (affix, clitic, and postposition) as number, animacy, and definiteness interact with case. One of the traditional morphological typology measures, the index of fusion, can capture some of this structural complexity, but the result is unsatisfying in that the language is placed somewhere toward the fusional end of the continuum. I argue that the concepts of phonological and grammatical word categories offer an alternative way of deconstructing the notion of fusion which captures the structural complexity of Sinhala with a more precise level of detail.

1.1. Morphological Typology

THE TRADITIONAL VIEW: FUSION AND SYNTHESIS. The problem of classifying languages according to the familiar morphological typology of isolating, agglutinating, fusional, and polysynthetic has occupied linguists for many years (Greenberg 1960, Sapir 1921, *inter alia*). This involves determining a language's place along two continua: isolating – synthetic and fusional – agglutinating. The isolating – synthetic continuum focuses on the number of morphemes per word (an isolating language having, ideally, one morpheme per word and a synthetic language having many). The fusional – agglutinating continuum focuses on the extent to which there are clear boundaries between morphemes within a word (a fusional language lacks clear boundaries, while an agglutinating language has them). The analysis in this paper will focus on the measures connected to the index of fusion. According to Comrie (1981), the two main measures for determining the level of agglutination versus fusion are invariance of the morphemes and the segmentability of the morphemes. The closer a language is to the agglutinating end of the continuum, the more invariant and easily segmentable the morphemes will be. Languages closer to the fusional end will have morphemes with more morphophonemic variation and less segmentability.

Many have pointed out that these are ideal types that represent points along a continuum and that no real language completely fits one category or the other, since for example, a completely fusional language would be entirely suppletive (e.g. Comrie 1981). Other objections to the adequacy of this typology have also been raised, including the difficulty of dealing with the root-like lexical affixes of some North American languages (Mithun 1997) and the need to explore connections between this typology and the grammaticization of meaning as well as form (Bybee 1997).

However, these questions have been addressed largely in languages which clearly use only affixes in their inflectional morphology. No comment has been made in the literature I have reviewed that addresses languages which seem to involve more than one structural level in the marking of inflectional categories. I believe Sinhala presents special challenges in that case marking appears to involve affixes, clitics, and postpositions in a complicated pattern affected by number, definiteness, and animacy. That is, there is no way to avoid having affixes, clitics, and postpositions all present in the case marking paradigm for this language, not as alternative ways of marking the same relations (as, for example in the English genitive, which can be expressed with a preposition *of* or a clitic *='s*), but as the only way of marking these relations in different parts of the paradigm. Describing the degree of fusion without recognizing that there are several levels of structure involved does not provide enough precision of analysis and is ultimately unsatisfying.

PHONOLOGICAL VERSUS GRAMMATICAL WORDS. Just as the traditional categories of morphological typology are recognized as ideal types that no real language fits neatly, the very category of *word* has been recognized as problematic as well. Dixon and Aikhenvald (2002) explore the phonological and grammatical criteria for wordhood, which often yield different results within a language (so that a phonological word is not always identical to a grammatical word). By their definition, a *phonological word* is a "phonological unit larger than the syllable...which has at least one (and generally more than one) phonological defining property" (13) based on segmental features (such as internal syllabic structures or word boundary phenomena), prosodic features (such as stress assignment or vowel harmony), or phonological rules (rules which apply only within the word or across word boundaries). A *grammatical word*, on the other hand, "consists of a number of grammatical elements which: a) always occur together..., b) occur in a fixed order, [and] c) have a conventionalized coherence and meaning" (19). Grammatical words, then, may consist of part of one, exactly one, or more than one phonological word (and vice versa).

Crucial to the analysis presented in this paper are the categories of affix, clitic, and postposition. By the definitions presented above, affixes are neither phonological words nor grammatical words. They are phonologically bound to the stem, taking part in word-internal phonological processes, and they display cohesion with the noun stem grammatically (nothing can intervene between the stem and the affix). Like affixes, clitics are not phonological words. They are bound to the noun stem and take part in word-internal phonological processes. Clitics are, however, grammatical words. they do not have the same cohesion with the noun stem that affixes have (other elements my intervene). Finally, postpositions are both phonological and grammatical words. They do not participate in word-internal phonological processes with the noun, and they may show instead word-boundary phenomena. They have a 'conventionalized coherence and meaning' of their own, and like clitics, they do not show cohesion with the noun. In section 4, I make use of these definitions and the analytical tools of phonological versus grammatical words in trying to capture the patterns of Sinhala nominal morphology while also paying tribute to its structural complexity.

2. NOMINAL MORPHOLOGY IN SINHALA. The analysis presented here involves number, definiteness, and case marking on nouns. The distinctions made by Sinhala morphology in these three categories are as follows:

Number: Singular, Plural
Definiteness: Definite, Indefinite
Case: Nominative (unmarked), Accusative, Dative, Genitive/Locative,
                 Instrumental/Ablative

Number and definiteness will be examined separately before the discussion of the case marking system, which necessarily involves both number and definiteness along with case.

2.1. NUMBER MARKING ON NOUNS IN THE WORLD'S LANGUAGES. Number marking in Sinhala consists of a binary distinction between singular and plural for count nouns. In English and many other Indo-European languages, the singular is unmarked and the plural carries some marking. There are many languages, however, that mark the singular rather than the plural or mark both the singular and plural morphologically. Corbett (2000:156) provides the following summary of the three possible systems:

Type A: base   *versus* plural
Type B: singulative   *versus* base
Type C: singulative   *versus*   plural

Following Dimmendaal (2000), I refer to Corbett's Type A as *plural*, Type B as *singulative*, and Type C as *replacive.* While many languages use only one of these systems, it is possible, though less common, for languages to use more than one system for different kinds of nouns (as Dimmendaal claims is common among Nilo-Saharan languages). As demonstrated in the analysis below, Sinhala uses all three.

For some languages that use singulative, plural, and replacive morphology to mark number, animacy has been found to be useful in determining which nouns take part in each system (see Dimmendaal 2000 on number marking in Nilo-Saharan languages). Animacy is a relevant category for many processes in language, including case marking, verb agreement and number marking (Comrie 1981) and has been noted as a salient category in both Indo-Aryan languages (Cardona 1990, Masica 1991) (which includes Sinhala) and Dravidian languages (Steever 1990) (including Tamil, a neighboring language of Sinhala). Thus it should be no surprise that animacy seems to play a role in determining which nouns in Sinhala fall into the various classes. Further, the division of nouns into several classes with different number marking patterns is common in Indo-Aryan languages, although Masica (1991) notes that Sinhala "presents an exceedingly complex picture" (228).

2.2. NUMBER MARKING IN NOMINATIVE CASE IN SINHALA. The seemingly simple picture of singular/plural marking by suffixes on nouns in Sinhala is complicated by the rather large number of noun classes (twelve , including seven animate classes and five inanimate classes). These noun classes cannot be predicted based on semantics or phonology. It is further complicated by the fact that some of the classes show a singulative marking pattern, some show a plural marking pattern, and some show a replacive pattern.

The singulative, plural, and replacive patterns are dealt with in the subsections below. For each general pattern, the classes of count nouns that fall under the general pattern are outlined, and any obvious semantic patterns are discussed.

SINGULATIVE PATTERNS. There are two groups of nouns that show a singulative pattern in Sinhala. One group is made up of animate nouns and the other is made up of inanimate nouns. As can be seen in Table 1, both groups use the stem for the plural form and add a vowel suffix to form the singular. Animate nouns form the singular by adding the suffix -*a*, while inanimate nouns add the suffix –*ə*.

Noun Class	Animate/Inanimate	Singular Forms	Plural Forms	English Gloss
A1	Animate	(stem + -a)	(stem)	
		kumbi-ya	*kumbi*	'ant'
		harək-a	*harak*	'cow'
		maalu-wa	*maalu*	'fish'
		laməy-a	*lamai*	'child'
I1	Inanimate	(stem + - ə)	(stem)	
		raum-ə	*raum, rauŋ*	'circle'
		æs-ə	*æs*	'eye'
		huləǐg-ə	*hulaŋ*	'wind'
		taruw-ə	*taru*	'star'
		suli-yə	*suli*	'current'
		liňd-ə	*liŋ*	'well'

TABLE 1. Examples of singulative patterns

All of the nouns in class A1 are animate, and all of the nouns in class I1 are inanimate. There is a slight trend for the nouns in both classes to be items usually occurring in groups or pairs (such as cows, ants, horns, and stars), though the trend is not absolute (counterexamples include circle, cave, and desert) and seems to be stronger for the nouns in the animate class.

PLURAL PATTERNS. There are three plural patterns in Sinhala, all of which apply to animate nouns. These nouns use the stem for the singular and add a suffix to form the plural. As can be seen in Table 2, the three plural suffixes are –*la*, -*n*, and –*wəru*.

Noun Class	Animate/Inanimate	Singular Forms	Plural Forms	English Gloss
A2	Animate	(stem)	(stem + -la)	
		taata	*taata-la*	'father'
		aaci	*aaci-la*	'grandmother'
		duwə	*duwa-la*	'daughter'
		raaləhaami	*raaləhaami-la*	'police officer'
		ætinni	*ætinni-la*	'female elephant'
A3[1]	Animate	(stem)	(stem + -n)	
		iiri	*iirii-n*	'sow'
		birində	*birinda-n*	'wife'
		ætinni	*ætinnii-n*	'female elephant'

---

[1] Note: -n occurs on other plural nouns as a case marker, but the consultant claims that there is no other way to pluralize these nouns and that the –n does not indicate a different case.

A4	Animate	(stem)	(stem + -wəru)	
		duwə	*duu-wəru*	'daughter'
		mawə	*mau-wəru*	'mother'
		piya	*piya-wəru*	'father'

TABLE 2. Examples of nouns with plural patterns

The nouns in class A2 are all human except *ætinni* 'female elephant,' which has an alternate plural form in class A3. The human terms are all kin terms and professions. The nouns in class A3 are all female, but there are very few examples, so it is unclear whether the generalization would hold across more examples. The nouns in class A4 are human, and possibly carry a respect connotation. The terms for 'mother' and 'father' in this class are considered more formal than the terms for 'mother' and 'father' that belong to class A2.

REPLACIVE PATTERNS. There are seven more patterns for number marking on Sinhala count nouns, all of which are replacive. Three of these patterns operate on animate nouns, while four operate on inanimate nouns. Table 3 shows that some of the patterns partially overlap. The suffix *-a* is used for the singular in two of the three animate noun groups, and *-ə* marks the singular for all four groups of inanimate nouns. The suffix *-u* is used in two of the groups of animate noun plurals and one of the groups of inanimate noun plurals.

Noun Class	Animate/Inanimate	Singular Forms	Plural Forms	English Gloss
A5	Animate	(stem + -a)	(stem + gem + -u)	
		walah-a	*walass-u*	'bear'
		makər-a	*makər-u*	'dragon'
		put-aa	*putt-u*	'son'
		leen-a	*leenn-u*	'squirrel'
		riləw-a	*rila-u*	'rhesus monkey'
		hiwəl-a	*hiwall-u*	'wolf'
		næædææ-ya	*næædææ-yu*	'relative'
A6	Animate	(stem + -a)	(stem + -o)	
		kurull-a	*kurull-o*	'bird'
		waluur-a	*waluur-o*	'boar'
		mu-wa	*mu-wo*	'deer'
		puus-a	*puus-o*	'cat'
		ukun-a	*ukun-o*	'louse'
		gowi-ya	*gowi-yo*	'farmer'
A7[2]	Animate	(stem + -i)	(stem + -u)	
		gææn-i	*gææn-u*	'woman'

---

[2] Only one token was found for this 'pattern.' Although the consultant was not able to provide another example, he felt that it was a pattern rather than an isolated irregular form. This pattern is therefore tentative at best.

I2	Inanimate[3]	(stem + -ə) *walaakul-ə* *deedunn-ə* *akun-ə* *kand-ə* *nœw-ə*	(stem + dgm + -u) *walaakul-u* *deedun-u* *akun-u* *kaňd-u* *nœ-u*	'cloud' 'rainbow' 'lightning' 'mountain, hill' 'ship'
I3	Inanimate	(stem + -ə) *raṭ-ə* *lookə-yə* *œl-ə*	(stem + -əwal) *raṭ-əwal* *look-əwal* *œl-əwal*	'country, nation' 'world' 'stream, brook'
I4	Inanimate	(stem + gem + -ə) *wœss-ə* *pinn-ə* *diyə œll-ə*	(stem + -i) *wœs-i, wœh-i* *pin-i* *diyə œl-i*	'rain' 'dew' 'waterfall'
I5[4]	Inanimate	(stem + -ə) *gaňg-ə*	(stem + -aa) *gaňg-aa*	'river'

TABLE 3. Examples of nouns with replacive patterns[5]

Several of these patterns involve changes to the stem when the plural or singular suffix is added. In class A5, stem-final stops, nasals, fricatives and laterals geminate before the plural suffix (as in 'squirrel' *leena/leennu*), and [h] becomes [s] before gemination (as in 'bear' *walaha/walassu*). In class I2, stem-final geminate consonants become single when the plural suffix is added (see 'rainbow' *deedunnə/deedunu*), and stem-final nasal+stop sequences become prenasalized stops (see 'mountain, hill' *kandə/kaňdu*). A comparison of 'rainbow,' which has a geminate [n] in the singular and a single consonant in the plural, with 'lightning' *akunə/akunu* shows that the geminate is part of the stem and that the process involves degemination in the plural rather than gemination in the singular. In class I4, stem-final stops, nasals, fricatives, or laterals geminate when the singular suffix is added (for example, 'dew' *pinnə/pini*), and [h] goes to [s] before gemination (see 'rain' *wœssə/wœhi*), as in other patterns. This pattern seems to involve gemination of the stem-final consonant in the singular rather than degemination in the plural because the term for 'waterfall' *diyə œllə/diyə œli* contains the word for 'stream,' which is *œlə* (with two possible plural forms *œlə* or *œləwal*).

The nouns in classes A5, A6, and A7 are all animate, and the nouns in classes I2, I3, I4, and I5 are inanimate. Also, the nouns in class I4 have peculiar meanings for some forms, not so much singular and plural as marking individuation. For these nouns, the singular denotes the general substance or concept, while the plural calls attention to some individual parts or pieces (for example, the singular form of 'dew' *pinnə* would be used to indicate that there is dew on the ground, but the plural *pini* would be used to bring attention to the drops of dew in the environment).

---

[3] One noun in this pattern *dostərə/dostəru* 'doctor' (SG/PL) is animate, but the rest are inanimate.
[4] See footnote 3.
[5] gem = gemination of stem-final consonant, dgm = degemination of stem-final CC

2.3. DEFINITENESS

DEFINITENESS MARKING ON COUNT NOUNS. Singular count nouns in Sinhala are marked for indefiniteness with a clitic, but plural nouns are not marked (unless they are followed by a quantifier or numeral, which may take indefinite marking), as can be seen in the examples below. The status of the indefinite marker as a clitic is established in section 2.3.2.

(1) lamǝy-a   maawǝ   dækka
     child-SG.DEF   1SG.ACC   see.PST
     'The child saw me.'

(2) lamǝy=ek   maawǝ   dækka
     child=SG.IND   1SG.ACC   see.PST
     'A child saw me.'

(3) lamai   maawǝ   dækka
     child.PL   1SG.ACC   see.PST
     'The children saw me.' or 'Children saw me.'

(4) maŋ   gaŋ̌g-ǝ   dækka
     1SG   river-SG.DEF   see.PST
     'I saw the river.'

(5) maŋ   gaŋ̌g=ak   dækka
     1SG   river=SG.IND   see.PST
     'I saw a river.'

(6) maŋ   gaŋ̌g-aa   dækka
     1SG   river-PL   see.PST
     'I saw the rivers.' or 'I saw rivers.'

Based on these patterns, it appears that count nouns in Sinhala have three basic forms: definite singular, indefinite singular, and plural. The plural marker (when there is one) is clearly a suffix rather than a clitic, as can bee seen from the fact that it appears only on nouns and appears on nouns when they are not the last item in the noun phrase (see example 9 below). The plural form varies depending on which of the twelve classes the noun belongs to (some nouns use the stem for the plural, while others have *-la, -n, -wǝru, -u, -o, -ǝwal, -i,* or *–aa* suffixes). The singular definite suffix also depends on the class of the noun (again, some nouns use the stem for the singular definite form, while others have *-a, -ǝ,* or *–i* suffixes). The singular definite marker behaves much like the plural marker and will be considered a suffix. The singular indefinite clitic appears to be more consistent, with only two variants, *=ek* for animate nouns and *=ak* for inanimate nouns. Table 4 provides a representative set of examples.

Noun class	Singular definite	Singular indefinite	Plural	English gloss
Animate	*laməy-a*	*laməy=ek*	*lamai*	'child'
	birind-ə	*birind=ak*	*birind-an*	'wife'
	duwə[6]	*duw=ek*	*duwə-la,* *duu-wəru*	'daughter'
	noona	*noona kenek*	*noona-la*	'wife'
	dostərə	*dostərə kenek*	*dostərə-la, dostərə-wəru*	'doctor'
	minih-a	*minih=ek*	*miniss-u*	'man'
	buuru-wa	*buuru=wek*	*buuru-wo*	'donkey'
	ætinni	*ætinni=yek*	*ætinni-in*	'f. elephant'
Inanimate	*meesə-yə*	*meesə=yak*	*meesə*	'table'
	gal-ə	*gal=ak*	*gal*	'stone'
	raṭə	*raṭ=ak*	*raṭə-wal*	'country, nation'
	dawəs-ə	*dawəs=ak*	*dawas*	'day'
	pot-ə	*pot=ak*	*pot*	'book'
	gaŋg-ə	*gaŋg=ak*	*gaŋg-aa*	'river'

TABLE 4. Number and definiteness marking on animate and inanimate nouns

Note that the table also includes a number of exceptions in the animate class. Two of the nouns referring to humans are made indefinite through the use of an apparent classifier, *kenek* 'people', which does not have a definite form. This is also true of at least one other noun, *rajjəkenek* 'king'. Note that these nouns are part of the small number of nouns in Sinhala (all animate) which use the noun stem as the singular form. This small collection suggests that there may be a respectful connotation to this classifier, but further investigation would be required before making such a generalization. The use of this classifier provides yet another example of how the singular definite and plural suffixes behave differently from the indefinite clitic. Although these nouns require the *kenek* classifier to take the indefinite marker, the singular definite and plural markers can attach directly to the noun itself.

In addition, *birində* 'wife' appears to take the inanimate suffix rather than the animate. Masica (1991:248) points out that the inanimate indefinite marker is also used for a few feminine nouns. It is not used on all feminine nouns, nor is it predictable from the *ə*-final stem, as shown by 'daughter' *duwə/duwek*.

One thing that this data suggests is that, for animate nouns at least, singular nouns are formed by adding either the singular definite suffix or the indefinite clitic appropriate to the noun class (rather than by adding first the singular suffix and then the indefinite clitic after it). The vowel alternation between the *–a* endings on many animate singular definite nouns and the *=ek* endings on animate singular indefinite nouns, suggests that the *=ek* clitic is added to the noun stem directly, replacing the singular definite suffix rather than being added after it.

---

[6] The singular definite form of 'daughter' is difficult to analyze because its alternate plural forms point to different possible interpretations. The *–la* plural form suggests that the *ə* in the singular form is part of the stem, but the *–weru* form suggests that the stem might be *duw-* with the *–ə* functioning as the singular marker.

The similar alternation between *–ə* and *=ak* for inanimates, however, might be explained by the tendency for alternation between [ə] and [a] based on syllable structure in Sinhala, with [a] appearing in closed syllables and [ə] in open syllables (Gair and Paolillo, 1997). Thus, the vowel alternation between the *–ə* singular definite inanimate suffix and the *=ak* inanimate indefinite clitic may be predicted by phonological rule. The indefinite clitic for inanimate nouns, therefore, could be viewed either as *=ak,* following the same pattern as the animate nouns (attaching to the noun stem and taking the place of the definite suffix), or as *=k,* which is added after the singular definite suffix and triggers the vowel change from [ə] to [a] by phonological rule.

DEFINITENESS MARKING ON QUANTIFIERS AND NUMERALS. Mass nouns and plural nouns can be marked as definite or indefinite by using a quantifier or numeral. The plural form of the noun is used, followed by a numeral or quantifier, which can be marked with the indefinite clitic. Table 4 shows some quantifiers and numerals in their definite and indefinite forms. Numerals and some quantifiers have different forms to accompany animate and inanimate nouns.
   The sentences below illustrate the use of definite and indefinite quantifiers and numerals.

(7) maŋ   mas   tikə   dækka
     1SG   meat   some.DEF   see.PST
     'I saw some (specific) meat.'

(8) maŋ   mas   tikak   dækka
     1SG   meat   some.IND   see.PST
     'I saw some (unspecified) meat.'

(9) maŋ   gaŋ̆gaa   kiipəyak   dækka
     1SG   river.PL   a.few.IND   see.PST
     'I saw some (unspecified) rivers.'

(10) maŋ   gaŋ̆gaa   tunak   dækka
      1SG   river.PL   three.IND   see.PST
      'I saw three rivers.'

(11) maŋ   gaŋ̆gaa   tunə   dækka
      1SG   river.PL   three.DEF   see.PST
      'I saw the three rivers.'

Animacy in the noun requires the quantifier or numeral to use the *–denə* marker, and once again *=ek* is used for animate indefinites and *=ak* is used for inanimate indefinites.

(12) maŋ   lamai   kiipədenekwə   dækka
      1SG   child.PL   a.few.ANIM.IND.ACC   see.PST
      'I saw some (unspecified) children.'

(13) maŋ   lamai   tundenek   dækka
      1SG   child.PL   three.ANIM.IND   see.PST
      'I saw three children.'

(14) maŋ   lamai   tundenə   dækka
      1SG   child.PL   three.ANIM.DEF   see.PST
      'I saw the three children.'

The status of the indefinite marker as a clitic is shown by the fact that it attaches to the last item in the noun phrase, so that it is the quantifier or numeral, and not the noun, that is marked for indefiniteness in the examples above. Crucially, it must be the last item in the noun phrase, as shown by the use of *saaməharə* 'some' in example 15 below.

(15) a. saaməharə   taata-la
          some.DEF   father.PL
      b. taata-la   saaməharek
          father.PL   some.ANIM.IND
      c. *saaməharek   taata-la
          some.ANIM.IND   father.PL

Although the quantifier *saaməharə* may be positioned before or after the noun, it can only be marked for indefiniteness when it follows the noun. Although the singular definite marker also seems to appear on quantifiers, it can appear on a quantifier before the noun (as seen in example 15). This combined with the fact that it patterns with the plural suffixes according to the twelve noun classes while the indefiniteness clitics pattern only according to the animacy of the noun leads me to analyze the singular definite as a suffix rather than a clitic.

   2.4. OVERVIEW OF CASE MARKING.   Sinhala marks noun phrases as accusative, dative, locative/genitive, and instrumental/ablative using clitics and postpositions, as shown in Table 5 (= indicates a clitic). The status of these markers as clitics and postpositions is demonstrated below.

Case	Singular Definite		Singular Indefinite		Plural	
	Animate	Inanimate	Animate	Inanimate	Animate	Inanimate
ACC[7]	=wə	--	=wə	--	=wə	--
DAT	=ţə	=ţə	=ek(o)ţə[8]	=əkəţə	=ţə[9]	wələţə
LOC[10]	--	=e	--	=əkə	--	wələ
GEN	=ge	=e	=ekge	=əkə	=ge	wələ

---

[7] Accusative case is marked only on animate nouns and appears to be optional in at least some instances.

[8] Most animate nouns appear to take =*ekţə*. However, *duwə* 'daughter' takes =*ekoţə*.

[9] V in all animate plural oblique forms represents a, i, or u, depending on the singular/plural pattern of the noun and the shape of the noun stem.

[10] Though the locative and genitive clitics are identical for all inanimate categories, the genitive clitic cannot be used with a locative meaning on animate nouns. For example, to express the equivalent of 'The fly landed on the donkey,' the consultant uses the following:

mæssa   buuruwage   æŋgee   wæhuwa
fly.SG.ANM.DEF   donkey.SG.ANM.DEF.GEN   body. SG.INAN.DEF.LOC   land.PST
Lit.: 'The fly landed on the donkey's body.'

INST[11]	--	=eŋ/iŋ	--	=əkiŋ	--	wəliŋ
ABL	=geŋ	=eŋ/iŋ	=ekgeŋ	=əkiŋ	=geŋ	wəliŋ

TABLE 5. Summary of case markers

CASE MARKERS AS CLITICS (AND SOMETIMES POSTPOSITIONS) IN SINHALA. Case markers, like indefinite markers, occur on the final element in the noun phrase rather than specifically on the noun, which shows their status as clitics, as shown in (16) below.

(16) a ohu   lamaiŋgeŋ   losindərə   gatta
         3SM   child.PL.ANIM.ABL   candy   take.PST
         'He took candy from the children.'

      b ohu   lamai   kiipədenekgeŋ   losindərə   gatta
         3SM   child.PL   a.few.AN.IND.ABL   candy   take.PST
         'He took candy from a few children.'

      c ohu   kootu   wəliŋ   buuruwaṭə   gæhuwa
         3SM   stick.PL   PL.INAN.INST   donkey.SG.ANM.DEF.DAT   hit.PST
         'He hit the donkey with sticks.'

      d ohu   kootu   kiipəyəkiŋ   buuruwaṭə   gæhuwa
         3SM   stick.PL   a.few.INAN.IND.INST   donkey.SG.ANIM.DEF.DAT   hit.PST
         'He hit the donkey with a few sticks.'

      e ohu   kootu   wəliŋ   buuruwo   kiipədenekoṭə   gæhuwa
         3SM   stick.PL   PL.INAN.INST   donkey.PL.ANIM   a.few.ANIM.IND.DAT   hit.PST
         'He hit a few donkeys with sticks.'

The examples above show the case marker attaching to the quantifier 'a few' when it follows the noun, illustrating that it is the last item in the noun phrase, rather than the noun itself, that receives case marking. For singular nouns and animate plurals, the case markers are phonologically bound to the word they attach to.

The case markers for plural inanimates, however, are not phonologically bound and therefore resemble case marking postpositions rather than clitics. I use the plural inanimate instrumental/ablative marker *wəliŋ* to illustrate this in the examples below. The phoneme /w/ has various allophones based on position within the word. In word initial position, it is pronounced [v], while word internally, it is pronounced [w] following a consonant. The /w/ in the first sentence in (16) is pronounced [v], which supports the argument that *wəliŋ* is a separate phonological word, hence a postposition rather than a clitic.

---

[11] Similar to the situation for locative in the previous note, the ablative clitic cannot be used with an instrumental meaning on animate nouns. For example, to express the equivalent of 'He pulled the cart with donkeys,' the consultant uses the following:
ohu   buuruwo   lauwa   karatte   add-a
3SM   donkey.PL.ANM   using   cart.SG.INAN.DEF   pull-PST
'He pulled the cart using donkeys.'

(17) a   æya   pææŋ wəliŋ   pinture   ænda
          3SF   pen.PL PL.INAN.INST   picture.SG.INAN.DEF   draw.PST
          'She drew the picture with pens.'

      b   æya   pænsal   saha creyon   wəliŋ
          3SF   pencil.PL   and crayon.PL   PL.INAN.INST
          pinture   ænda
          picture.SG.INAN.DEF   draw.PST
          'She drew the picture with pencils and crayons.'

      c   æya   pænsal   dek=əkiŋ   saha   creyon
          3SF   pencil.PL two=IND.INST   and   crayon.PL
          tun=əkiŋ   pinture   ænda
          three=IND.INST   picture.SG.INAN.DEF   draw.PST
          'She drew the picture with two pencils and three crayons.'

In addition, the second two sentences in the example show that *wəliŋ* can have scope over a conjoined noun phrase (*pænsal saha creyon)* while the clitics on the numerals in the third sentence must be repeated in each of the conjoined noun phrases. One further argument for the status of the plural inanimate case markers can be made based on the way in which the various case markers attach to the nouns. This argument will be addressed below, once the pattern of attachment to noun stems has been discussed.

HOW CASE MARKERS ATTACH TO NOUN STEMS. The accusative and dative markers attach to both singular and plural forms of nouns (Tables 6 and 7 below). These case markers, like those for oblique cases, attach to a special form of the plural noun ending in *-Vn* (Masica 1991 identifies this as a vestigial general oblique marking from Old Indo-Aryan). Given this identification along with the consultant's identification of the *–Vn* as 'another plural,' I am considering it a plural suffix for the purpose of describing the attachment of case markers to the nouns. The case markers shown in Tables 6 and 7 are clitics, coming after the singular or plural affix. The /w/ of the accusative marker (Table 6) is pronounced as [w] after a consonant, consistent with its status as phonologically bound to the noun. The case for =ṭə (Table 7) being phonologically bound (and therefore a clitic) can be made through the fact that the final nasal of the plural oblique suffix *–Vn* is realized as [n] rather than the usual word-final realization of all nasals as [ŋ].

Gloss	Accusative Singular	Nominative Singular	Accusative Plural	Nominative Plural
'child'	*laməy-a=wə*	*laməy-a*	*lama-iŋ=wə*	*lamai*
'farmer'	*gowi-ya=wə*	*gowi-ya*	*gowi-yaŋ=wə*	*gowi-yo*

TABLE 6. Accusative markers attach to singular or plural form

Gloss	Dative Singular	Nominative Singular	Dative Plural	Nominative Plural
'child'	*laməy-a=ʈə*	*laməy-a*	*lama-in=ʈə*	*lamai*
'farmer'	*gowi-ya=ʈə*	*gowi-ya*	*gowi-yan=ʈə*	*gowi-yo*
'cow'	*harək=ʈə*	*harək-ə*	*harək-un=ʈə*	*harak*
'head'	*olu-wə=ʈə*	*olu-wə*	*olu=ʈə*	*olu*
'school'	*iskoole=ʈə* (*iskoolə-yə=ʈə*)	*iskoole* (*iskoolə-yə*)	*iskoolə=ʈə*	*iskoolə*

TABLE 7. Dative markers attach to singular or plural form

The locative and genitive markers are syncretic, as are the instrumental and ablative markers. Both the LOC/GEN and ABL/INST markers are attached to the singular or plural form for animate nouns (see Table 8), but they are attached to the noun stem for inanimates (see Table 9).

Gloss	LOC/GEN Singular	ABL/INST Singular	Direct Singular	LOC/GEN Plural	ABL/INST Plural	Direct Plural
'child'	*lamə-ya=ge*	*lamə-ya=geŋ*	*lamə-ya*	*lama-iŋ=ge*	*lama-iŋ=geŋ*	*lamai*
'dog'	*ball-a=ge*	*ball-a=geŋ*	*ball-a*	*ball-aŋ=ge*	*ball-aŋ=geŋ*	*ball-o*
'man'	*minih-a=ge*	*minih-a=geŋ*	*minih-a*	*miniss-uŋ=ge*	*miniss-uŋ=geŋ*	*miniss-u*

TABLE 8. LOC/GEN and ABL/INST markers on animate nouns

Gloss	LOC/GEN Singular	ABL/INST Singular	Direct Singular	LOC/GEN Plural	ABL/INST Plural	Direct Plural
'head'	*olu=we*	*olu=weŋ*	*olu-wə*	*olu wələ*	*olu wəliŋ*	*olu*
'knife'	*pihi=ye*	*pihi=yeŋ*	*pihi-yə*	*pihi wələ*	*pihi wəliŋ*	*pihi*
'table'	*mees=e* (*meesə=ye*)	*meesə=yeŋ*	*mees-e* (*meesə-yə*)	*meesə wələ*	*meesə wəliŋ*	*meesə*

TABLE 9. LOC/GEN and ABL/INST markers on inanimate nouns

Table 9 shows that the case markers attach to the noun stem rather than the singular form (the plural form and the noun stem are identical in these examples). If the case markers were attached to the singular form of the noun, for example, 'on the head' would be **oluwəye* instead of the attested form *oluwe*. Further evidence can be seen in some less typical examples for the inanimate plural case markers. The noun stem form is often the plural form, since many inanimate nouns in Sinhala are part of the singulative pattern of singular/plural marking in the direct case. The examples in the table above show this most typical case, but the rare case of 'country' *raʈə* (which is *raʈəwal* in the plural but *raʈə wəliŋ* in the instrumental) shows that the noun stem is used, even if it is singular.

However, if the noun is part of the replacive singular/plural marking pattern in the direct case, so that the noun root is bound and cannot appear without a singular or plural marking, the case marker follows the plural form, as is shown by 'stick' *kooʈə*, which is *kooʈu* in the plural and *kooʈu weliŋ* in the plural instrumental. This shows that it is the plural form specifically,

and not the noun stem, that is followed by the case marker for locative/genitive and instrumental/ablative.

   To return now to the status of the inanimate plural case markers as free postpositions rather than phonologically bound clitics, the example above provides further evidence for this by illustrating the fact that the markers must follow a free form of the noun. In other words, unlike the clitics which can attach to a bound noun stem (e.g. 'with the stick' *koot=eŋ*, in which *koot-* is a bound form), the postposition cannot follow a bound form (so that *weliŋ* only follows free plural forms such as *kootu* 'sticks') .

   Having provided an overview of the case marking forms, I now move on to discuss the extent to which these forms should be considered agglutinating or fusional, according to the traditional definitions in morphological typology.

   3. ANALYSIS IN TERMS OF TRADITIONAL MORPHOLOGICAL TYPOLOGY. The traditional measures of morphological typology are the indices of synthesis and fusion, as mentioned in the literature review. The analysis in this section deals with the measures of fusion rather than synthesis both to limit the scope of the paper and because the measures of fusion seem more suited to the analysis of the types of structures involved than measures of synthesis. Each of the cases will be analyzed separately on the basis of segmentability and invariance of the case marking morphemes in the subsections that follow, since the behavior of each is slightly different.

   3.1. ACCUSATIVE MARKERS. In all the markers for accusative case, the *=wə* portion of the marker remains invariant (though it is preceded by other material in the singular indefinite and plural forms). Table 10 shows the accusative forms with examples, using *buuruwa* 'donkey' (inanimate nouns do not take the accusative marker in Sinhala).

Case	Singular Definite	Singular Indefinite	Plural
	Animate (*buuruwa*)	Animate (*buuruwek*)	Animate (*buuruwo*)
ACC	=wə (*buuru-wa=wə*)	=ekwə (*buuru=wekwə*)	=wə (*buuru-waŋ=wə*)

TABLE 10. Accusative case markers with examples

   The accusative marker is easily segmentable from the number marking (so *buuruwawə* is easily segmented into the noun root *buuru* 'donkey,' the singular marker *–wa,* and the dative marker *=wə*). For singular indefinite,   *=wə* is preceded by *=ek,* which is the indefinite marker for animate nouns, so it is clearly possible to segment the markers. For plurals, *=wə* is preceded by *–Vn,* the plural general oblique identified above, so it is again clearly segmentable by assigning the representation of number to *–Vn* and the representation of accusative case to *=wə*. Thus, the accusative marker appears to be a straightforward example of an agglutinative pattern.

   3.2. DATIVE MARKERS. In all the markers for dative case, the *=ṭə* portion of the marker remains invariant (though it is preceded by other material in all but the singular definite forms). Table 15 shows the dative forms with examples, using *buuruwa* 'donkey' for animate forms and *pæænə* 'bread' for inanimate forms.

Case	Singular Definite		Singular Indefinite		Plural	
	Animate (*buuruwa*)	Inanimate (*pæænə*)	Animate (*buuruwek*)	Inanimate (*pæænak*)	Animate (*buuruwo*)	Inanimate (*pæææŋ*)
DAT	=ṭə (*buuru-wa=ṭə*)	=ṭə (*pææn-ə=ṭə*)	=ek(o)ṭə (*buuru=wekṭə*)	=əkəṭə (*pææn=əkəṭə*)	–Vn=ṭə (*buuru-wan=ṭə*)	wələṭə (*pæææŋ wələṭə*)

TABLE 11. Dative case markers with examples

Since the singular definite dative marker attaches to the singular form of the verb for both animate and inanimate, it is easily segmentable from the number marking (so *buuruwaṭə* is easily segmented into the noun root *buuru* 'donkey,' the singular marker *–wa,* and the dative marker *=ṭə*). For singular indefinite, *-ṭə* is preceded by *–əkə* or *–eko* depending on animacy, so it is tempting to segment the markers and say that *-əkə /-eko* represents indefiniteness and animacy, although these forms differ slightly from the indefinite animate and inanimate forms used in direct cases (*=ek* and *=ak*, respectively), while *=ṭə* represents dative case. For plurals, *=ṭə* is preceded by *–Vn* for animate and *wələ* for inanimate, so it is again tempting to segment by assigning the representation of number and animacy to *–Vn/wələ* and the representation of dative case to *=ṭə*. However, *wələ* alone is the LOC/GEN marker. It clearly does not represent that here. In this case, it seems that the segmentability of *wələṭə* is at best ambiguous. On the other hand, *-Vn* is a plural ending used with animate nouns in all the oblique cases, as noted above in section 2.3.2, and it is consistent throughout the other cases, so this seems segmentable. In summary, while the dative marker seems segmentable and has the invariant *=ṭə* form all the way through, it is not equally segmentable in all combinations of animacy, definiteness, and number.

3.3. LOCATIVE/GENITIVE MARKERS. In terms of variability, the LOC/GEN marker is much less stable throughout the paradigm, to the extent that there are forms that share no phonemes between them (e.g. *=e* for singular definite inanimate nouns and *wələ* for plural inanimate nouns). Table 12 shows the LOC/GEN forms with examples, using *buuruwa* 'donkey' for animate forms and *pæænə* 'bread' for inanimate forms.

Case	Singular Definite		Singular Indefinite		Plural	
	Animate (*buuruwa*)	Inanimate (*pæænə*)	Animate (*buuruwek*)	Inanimate (*pæænak*)	Animate (*buuruwo*)	Inanimate (*pæææŋ*)
LOC/GEN	=ge (*buuruwage*)	=e (*pææne*)	=ekge (*buuruwekge*)	=əkə (*pæænəkə*)	=Vŋge (*buuruwaŋge*)	wələ (*pæææŋ wələ*)

TABLE 12. LOC/GEN case markers with examples

The apparent segmentability of the LOC/GEN markers varies by animacy. The singular definite markers *=e* and *=ge* seem segmentable into *–e* for LOC/GEN and *–g-* for animate. The singular indefinite form supports this segmentation for animate nouns, since *=ekge* segments nicely into *=ek*, representing singular indefinite animate, and *=ge* segmented as above. In this analysis, animacy is represented twice (by *–e-* in *=ek* and *–g-* in *=ge*). The singular indefinite inanimate marker *=əkə* is also segmentable to an extent. The *=ək* clearly represents indefinite (with the change of [a] to [ə] due to syllable structure), but *–ə* is a less satisfying candidate to represent LOC/GEN since it would be homophonous with the singular definite suffix for many inanimate nouns and does not resemble the marker for LOC/GEN in the singular definite. This

is similar to the case of *wɔlɔṭɔ* in the previous section in that the form appears segmentable, but the segmentation produces homophony within the paradigm (*wɔlɔ* represents LOC/GEN in some forms and plural inanimate in others, and *–ɔ* represents LOC/GEN is some forms and singular inanimate definite in others). The plurals again support segmentability more easily for animate than inanimate. The animate marker *-Vŋge* is clearly segmentable into the animate plural oblique *-Vn* (with assimilation of the nasal to following velar), the animate *–g-* and the LOC/GEN *-e*. The inanimate marker *wɔlɔ* could be segmented into *wɔl-* for inanimate plural and *-ɔ* for LOC/GEN, but the segmentation is less certain than for the animate. The overall picture for LOC/GEN shows although both animate and inanimate are segmentable, the segmentation of the animate forms is clearer and the forms less variable than for the inanimate.

3.4. INSTRUMENTAL/ABLATIVE MARKERS. In terms of invariance, the INST/ABL *-eŋ/-iŋ* is more like the dative than the LOC/GEN, as it is consistent throughout[12] (though it may be preceded by other material). Table 13 shows the INST/ABL forms with examples, using *buuruwa* 'donkey' for animate forms and *pæænɔ* 'bread' for inanimate forms.

Case	Singular Definite		Singular Indefinite		Plural	
	Animate (*buuruwa*)	Inanimate (*pæænɔ*)	Animate (*buuruwek*)	Inanimate (*pæænak*)	Animate (*buuruwo*)	Inanimate (*pææŋ*)
INST/ABL	=geŋ (*buuruwageŋ*)	=eŋ/iŋ (*pææneŋ*)	=ekgeŋ (*buuruwekgeŋ*)	=ɔkiŋ (*pæænɔkiŋ*)	=Vŋgeŋ (*buuruwaŋgeŋ*)	wɔliŋ (*pææŋ wɔliŋ*)

TABLE 13. INST/ABL case markers with examples

In terms of segmentability, the situation for the INST/ABL is very similar to the LOC/GEN. The singular definite *=eŋ* for inanimates and *=geŋ* for animates seem segmentable into *-eŋ* for INST/ABL and *–g-* for animate. The singular indefinite marker supports this segmentation for animates, since *=ekgeŋ* uses same *=geŋ* preceded by *=ek*, the animate singular indefinite marker used in the nominative case. The inanimate indefinite also seems segmentable, with the direct case *=ak* changing to *=ɔk* (due to the previously mentioned alternation between [ɔ] and [a]). The plural animate also appears segmentable with the by now familiar *–Vn* animate plural oblique marker followed by *=geŋ* (the segmentation of which is detailed above). The plural inanimate is segmentable into *wɔl-* for plural inanimate (as in the LOC/GEN forms) and ABL/INST *=iŋ*, The overall segmentability of INST/ABL markers is fairly clear for all forms.

The traditional measures of fusion begin to capture a pattern in which there is a difference between cases and between the animate and inanimate nouns within each case, but the ultimate result is to place Sinhala toward the fusional end of the continuum (but not at the extreme end). The details are left unaccounted for, a problem which the analysis in the next section takes a step towards rectifying.

4. AN ANALYSIS USING PHONOLOGICAL AND GRAMMATICAL WORD CATEGORIES. The structural complexity of nominal morphology in Sinhala can be better captured by making use of the grammatical word versus phonological word distinction. The match, or mismatch, between

---

[12] The variability of the vowel between *i* and *e* does not appear to be predictable by phonologicl rule or noun class. For example *potɔ* 'book' and *atɔ* 'hand' are both inanimate singulative nouns and similar in phonology, but one takes the *–iŋ* and the other takes *-eŋ* (*poteŋ* 'with the book' and *atiŋ* 'with/from the hand.')

grammatical and phonological word boundaries can help to show the different levels of structure involved, as shown in Table 14 below. A free noun form or postposition counts as both a grammatical word and a phonological word, a clitic counts as a grammatical word but not a phonological word, and a bound noun stem plus a clitic counts as 1.5 grammatical words. The whole-number values for free noun forms, postpositions, and clitics are taken from Dixon and Aikhenvald, but the decision to assign the value 1.5 to a bound noun stem plus a clitic is my own. I use this simply as a shorthand for capturing the fact that the clitic attaches to a form that does not stand on its own, and thus cannot be considered a whole grammatical word. Since the clitic carries its own status as a full grammatical word, I use the .5 designation for the bound noun stem. For example, *koot-ə* 'stick' consists of a stem plus singular definite affix and therefore counts as one phonological and grammatical word. The indefinite form *koot=ak* consists of a bound noun stem plus the indefinite clitic and therefore counts as one phonological word and 1.5 grammatical words. The plural instrumental form *koot-u wəliŋ* consists of a bound noun stem plus the plural affix (one grammatical and phonological word) and the postposition (also a grammatical and phonological word) and therefore counts as 2 grammatical and two phonological words.

Case	Singular Definite				Singular Indefinite				Plural			
	Animate		Inanimate		Animate		Inanimate		Animate		Inanimate	
	Gw	Pw	Gw	Pw	Gw	Pw	Gw	Pw	Gw	Pw	Gw	Pw
NOM	1	1	1	1	1.5	1	1.5	1	1	1	1	1
ACC	2	1	--	--	2	1	--	--	2	1	--	--
DAT	2	1	2	1	2	1	2	1	2	1	2	2
LOC	--	--	1.5	1	--	--	1.5	1	--	--	2	2
GEN	2	1	1.5	1	2	1	1.5	1	2	1	2	2
INST	--	--	1.5	1	--	--	1.5	1	--	--	2	2
ABL	2	1	1.5	1	2	1	1.5	1	2	1	2	2

TABLE 14[13]. Number of grammatical and phonological words in case forms

Analyzing the forms by grammatical and phonological word categories captures some of the structural complexity of the system, and it points to some of the patterns of difference between animate and inanimate nouns and between the different cases. In examining Table 14, we can note that the accusative and dative markers consistently result in forms that consist of two grammatical words and one phonological word across animacy and number (with the exception of plural inanimate, which has two grammatical and two phonological words in the dative). Animate nouns in general also exhibit this pattern (with the exception of the nominative, which consists of 1.5 grammatical words and one phonological word in the indefinite and one grammatical and one phonological word in the definite). For inanimate nouns, the overall pattern is 1.5 grammatical words and one phonological word for singular (with the exception of dative, which has two grammatical words) and two grammatical words and two phonological words for plural.

This analysis of case forms into grammatical and phonological words captures some of the structural complexity that is missed by the traditional analysis in terms of fusion, but it also confirms the differences between animate and inanimate forms noted in that analysis (in

---

[13] For the purposes of this table, Gw denotes grammatical word and Pw stands for phonological word.

which the degree of fusion seemed slightly greater for inanimates than for animates). The analysis by phonological/grammatical word perhaps helps clarify why the segmentability of animate and inanimate forms is different in the traditional analysis. The degree of fusion in the forms is reflected to a certain extent in the number of grammatical words. The forms which have 1.5 grammatical words (mostly inanimates) are generally more difficult to segment than those with two grammatical words (mostly animates). Also, those cases (accusative and dative) which were most easily segmentable and invariable in the traditional analysis are the same cases that display the most consistent structure across animacy and number in the phonological/grammatical word analysis.

5. CONCLUSION: STRUCTURAL COMPLEXITY IS MISSED BY THE TRADITIONAL ANALYSIS. The analysis above indicates that Sinhala nominal morphology would be appropriately classified as fusional, although not at the most extreme end of the scale. However, this classification does not give a very clear picture of the structure of nominal morphology in Sinhala because it fails to address the use of different kinds of structures (clitics and postpositions) to mark case and ignores the complex ways that the case markers are attached to the noun stem. The analysis by grammatical and phonological words revealed a distinct difference in the structure of animate and inanimate nouns, as well as differences based on number and differences between cases. This analysis confirms similar differences in the degree of fusion present along the same lines of animacy, number, and case. Combining the two analyses yields a clearer and more detailed picture of the structural complexity of Sinhala nominal morphology and its connections to categories of animacy and number.

REFERENCES

BYBEE, JOAN. 1997. Semantic aspects of morphological typology. *Essays on Language Function and Language Type: Dedicated to T. Givón*, ed.by Joan Bybee, John Haiman and Sandra Thompson, 25-37. Philadelphia: John Benjamins.

CARDONA, GEORGE. 1990. Indo-Aryan Languages. *The major languages of South Asia, the Middle East and Africa*, ed.by Bernard Comrie. London: Routledge.

COMRIE, BERNARD. 1981. *Language universals and linguistic typology: Syntax and morphology.* University of Chicago Press, Chicago.

CORBETT, GREVILLE G. 2000. *Number.* Cambridge University Press, Cambridge.

DIMMENDAAL, GERRIT J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42.214-61.

DIXON, R. M. W. and AIKHENVALD, ALEXANDRA Y. 2002. Word: a typological framework. *Word: a cross linguistic typology*, ed.by R. M. W. Dixon and Alexandra Y. Aikhenvald, 1-41. Cambridge: Cambridge University Press.

GAIR, JAMES W. and PAOLILLO, JOHN C. 1997. *Sinhala.* Lincom Europa, Munich.

GREENBERG, JOSEPH H. 1960. A quantitative approach to the morphological typology of language. *International Journal of American Linguistics* 26.178-94.

MASICA, COLIN P. 1991. *The Indo-Aryan Languages.* Cambridge University Press, Cambridge.

MITHUN, MARIANNE. 1997. Lexical affixes and morphological typology. *Essays on language function and language type: Dedicated to T. Givón*, ed.by Joan Bybee, John Haiman and Sandra Thompson, 357-71. Philadelphia: John Benjamins.

Sapir, Edward. 1921. *Language, an introduction to the study of speech.* Harcourt, Brace and Company, New York.

Steever, Sanford B. 1990. Tamil and the Dravidian languages. *The major langauges of South Asia, the Middle East and Africa*, ed.by Bernard Comrie. London: Routledge.

University of California, Santa Barbara
Department of Linguistics
3607 South Hall
Santa Barbara, CA 93106
jng@umail.ucsb.edu

MORPHOSYNTACTIC EXPRESSIONS OF POSSESSION AND EXISTENCE IN SINHALA

SALOME GUTIERREZ MORALES

University of California, Santa Barbara

1. INTRODUCTION. The main goal of this short paper is the study of the morphosyntactic relation of possessive and existential in Sinhala.[1] The motivation to study and analyze these kinds of constructions is mainly that they have been considered to be locational constructions since they are strongly related, not only because they share morphosyntactic behavior, but also because they are locational in origin (Clark 1968). The foundation of this kind of study comes from Lyons (1968), who noticed that the existential function of the verb 'to be' in English could not take place without a locative or temporal complement. Therefore, he argues that, the existential construction is strongly related to the locative construction. Furthermore, he also points out that there is the same parallelism between locative and possessive constructions. The only difference is that the possessive construction varies in word order because the topic should always be an animate noun. Therefore, what we are going to pursue in this paper is to discover how locational constructions are built in Sinhala and how are they related.

Interestingly, Sinhala shows the two kinds of possessive constructions that are found in the world's languages. That is, this language has possessive noun phrases and possessive clauses. However, these constructions do not behave similarly for the following reasons:

- *a)* Possessive NPs only occur with a genitive morpheme that attaches to a personal pronoun or to a noun.
- *b)* Possessive clauses utilize two different lexical verbs predicate: tiyenəwa 'exist' and innəwa 'exist'. The use of one or the other verb is determined by the animacy of the possessed noun. However, whenever the negative morpheme næ is incorporated into the possessive clause, not only is the verb no longer required, but the animacy distinction is also neutralized.
- c) The morphosyntactic characteristics expressed on possessive clauses (see b) are relevant as well in existential clauses because they behave similarly.
- d) It is noteworthy to say that what triggers the selection of the verb is semantically the animacy but grammatically is the Copula-S argument. Therefore, it is evident that possessive and existential constructions are closely related semantically and morphosyntactically.

Lyon (1968) was the first person who notices the similarity of locational constructions. Then Clark (1978) found the same behavior of these constructions in many languages in the world. Nevertheless, what makes Sinhala a very interesting language with respect to this issue is that this language uses two different verbs in locational constructions, based on the animacy of the object or thing being located in the clause. It is this general behavior of locational constructions based on animacy that is the focus of this short paper.

The organization of this paper is as follow: First, we present the behavior of possessive noun phrases and possessive clauses so as to show that possessors in possessive NPs take the

---

[1] Sinhala is a native language spoken in Sri Lanka. The dialect that we are going to use in here is that one from the city of Marutua, which belongs to the Columbo district. All linguistics information came from Wiroshana Nuwanpriya Oshan Fernando, who has been our consultant since September of 2004.

genitive/locative *–ge*, while in possessive clauses, the possessor takes the dative/locative *-ṭə* Second, we focus our attention on the existential construction to show that it has the same structure as a possessive clause. Third, we discuss the negation construction of locationals so as to show that even in this case, locational constructions behave alike since the predicate (*tiyenəwa* or *innəwa*) is neutralized. Finally, we present our conclusion, which is that study of the Sinhala locational constructions allow for the expansion of typological knowledge of these kinds of clauses.

2. Possessives.

2.1. Possessive Pronouns and NPs. Sinhala does not have special possessive pronoun forms since all of the possessives are built through the use of the personal pronoun root plus a genitive suffix.

PERSONAL PRONOUNS		POSSESSIVE PRONOUN	
1	*maa*[2]	1-GEN	*ma-gee*
2	*oyaa*	2-GEN	*oyaa-ge*
3M	*ohu*	3M-GEN	*ohu-ge*
3F	*æyə*	3F-GEN	*æyə-ge*
3A	*uu*	3A-GEN	*u-gee*
1PL	*api*	1PL-GEN	*ap-ee*
2PL	*oogollo*	2PL-GEN	*oogollaŋ-ge*
3PL	*eegollo*	3PL-GEN	*eegollaŋ-ge*

Paying attention to the possessive forms, we can see that the genitive suffix has three forms when it is added to the personal pronoun to make possessive pronoun forms. They are: *-ge*, *-gee*, and *-ee*. Each allomorph is restricted by syllable structure since it is realized as *-gee* only when the personal pronoun root has a monosyllabic open syllable form (see 1-GEN and 3A-GEN). On the other hand, *-ee* takes place when the root of the personal pronoun has a closed monosyllabic form, i.e. when it ends in a consonant (see 1PL-GEN). Finally, we have *-ge* when the root of the personal pronoun has more than one syllable as can be seen from the paradigm above. A summary of the behavior of the genitive with personal pronouns is shown in figure 1 below.

PERSONAL PRONOUN STEM	GENITIVE FORM
Monosyllabic open syllable	*-gee*
Monosyllabic close syllable	*-ee*
More than one syllable	*-ge*

Figure 1. Genitive form allomorphs with personal pronouns

The genitive suffix is very productive as it is used to make possessive noun phrases as well. As is generally accepted, possessive noun phrases are all of those that contain a possessor and a possessed. The possessor could be a pronoun or a noun as is showed in (1)-(6).

---

[2] In Sinhala first personal pronoun has three allomorphs. They are: Maa, maŋ, and mamə. Generally speaking the allomorph *mamə* is the one that is most used in this language. However, when the genitive suffix is added *mamə* is not use for the speakers but *maa*.

(1) ma-gee   amma   ma-gee   bat
     1-GEN   mother   1-GEN   rice
     'my mother'   'my rice'

(2) oyaa-ge   amma   oyaa-ge   bat
     2-GEN   mother   2-GEN   rice
     'your mother'   'your rice'

(3) ap-ee   amma   ap-ee   bat
     1PL-GEN   mother   1PL-GEN   rice
     'our mother'   'our rice'

(4) Mary-ge   mahatteyə
     Mary-GEN   husband
     'Mary's husband'

(5) gæænu   laməya-ge   sapattu
     girl   child-GEN   shoes
     'the girl's shoes'

(6) Nimal-ge   gedərə
     Nimal-GEN house
     'Nimal's house'

Our examples above demonstrate that in Sinhala there is no difference between a possessor expressed by a lexical noun phrase and one expressed by a pronominal noun phrase. Both of them behave alike because they take the same genitive suffix to indicate possession. Consequently, grammatically there is no strong distinction between a pronoun and a lexical noun in a possessive noun phrase. In addition to this, it is also possible in Sinhala to include a genitive noun phrase in a clause to overtly mark a possessive relationship, as it is show in example (7)-(12).

(7) [ma-gee   taatta]$_{NP}$ [dustərə kenek]$_{NP}$
     1-GEN   father   doctor person
     'My father is a doctor.'

(8) [ma-gee   balla]$_{NP}$ [lediŋ]$_{NP}$
     1-GEN   dog   sick
     'My dog is sick.'

(9) [meekə]$_{NP}$ [ma-gee   potə]$_{NP}$
     this   1-GEN   book
     'This is my book.'

(10) [ma-gee   taatta]$_{NP}$ pænn-a
     1-GEN   father   jump-PST
     'My father jumped.'

(11) [oya-gee   baba]$_{NP}$ hinaawun-a
     2-GEN   baby   smile-PST
     'Your baby smiled.'

(12) [ohu-gee   assəya]$_{NP}$ diuw-a
     3M-GEN   horse   ran-PST
     'His horse ran.'

Possessive clauses with existential verbs behave totally differently. As we will discuss below.

2.2. Possessive Clauses. Possessive clauses in Sinhala are built through the use of two existential verbs. This kind of construction not only behaves totally differently from possessive NPs, but it is also very interesting for the following reasons:
  a) The possessor does not take the genitive suffix *–ge* as is required in possessive NPs. Instead, it takes the dative case marker *-ṭə.*
  b) There are two existential verbs. They are: *tiyenəwa* and *innəwa.* The use of *tiyenəwa* or *innəwa* is determined by the animacy of the possessed. If the possessed is inanimate, the verb *tiyenəwa* is used; if the possessed is animate, the verb *innəwa* is used. Examples (13)-(17) show the use of *tiyenəwa.*

(13) [laməya-ṭə]_{NP}   [sellaŋbaduw-ak]_{NP}   tiye-nəwa
      child-DAT   play.thing-IND   exist-IMPF
      'The child has a toy.'

(14) [miniha-ṭə]_{NP}   [pihiy-ak]_{NP}   tiye-nəwa
      man-DAT   knife-IND   exist-IMPF
      'The man has a knife.'

(15) [ohu-ṭə]_{NP}   [gey-ak]_{NP}   tiye-nəwa
      3M-DAT   house-IND   exist-IMPF
      'He has a house.'

(16) [æyə-ṭə]_{NP}   [hungak salli]_{NP}   tibun-a
      3F-DAT   much money   exist-PST
      'She had a great deal of money.'

(17) [ma-ṭə]   Salli   tiye-nəwa
      1-DAT   money   exist-IMPF
      'I have money.'   (Gair and Paolillo 1997:66)

As we can learn from the data above, the possessive relation for inanimate possessed entities is indicated with the lexical verb *tiyenəwa* 'exist'. The following data (examples 18-23) show that with animate possessed entities, we have *innəwa* 'exist' as a lexical verb predicate, instead.

(18) [æyə-ṭə]_{NP}   [muv-ek]_{NP}   in-nəwa
      3F-DAT   deer-IND   exist-IMPF
      'She has a deer.'

(19) [laməya-ṭə]_{NP} [amm-ek]_{NP}   in- nəwa
      child-DAT   mother-IND   exist-IMPF
      'The child has a mother.'

(20) [muvaam-ṭə]_{NP}   [patiy-ek]_{NP}   in-nəwa
      deer-DAT   baby-IND   exist-IMPF
      'The deer has a baby.'

(21) [ma-ṭə]_{NP}   [puttu]_{NP}   innə-way
      1-DAT   son-PL   exist-QUOT
      'I have sons.'   (Gair and Paolillo 1997:66)

(22) [laməya-ṭə]ₙₚ [ball-ek]ₙₚ   hitiy-a
      child-DAT   dog-IND   exist-PST
      'The child had a dog.'

(23) [mahadænəmutta-ṭə]ₙₚ   [gooləyo pas-denek-ut]ₙₚ   hitiy-a
      mahad□n«mutta-DAT   follower.PL five-people-also   exist-PST
      'Mahad□n«mutta also had five followers.'

From these examples, we learn that Sinhala uses two existential verbs for possessive constructions. In addition, it is very clear that in this language the form *tiyenəwa* 'exist' is used when the possessed is inanimate, whereas if it is animate, speakers use the form *innəwa* 'exist' as a lexical verb predicate. Therefore, it is noteworthy that the verb agrees in animacy with the unmarked NP. That is, with the possessed NP since both of them show a semantic correlation. Therefore we can argue that in possessive clause construction with existential verbs, there is the following relationship:

Inanimate possessed ↔↔Inanimate verb

Animate possessed ↔↔Animate Verb

Looking at the structure of both kinds of possessive clauses (animate and inanimate) we can see that even though the possessor takes the dative/locative suffix *-ṭə*, it is actually working as an argument NP, the reason for what it has the dative/locative suffix is that it is actually indicating a locative relationship. Therefore, the noun that takes this suffix is also an argument of the clause. Consequently, *tiyenəwa* and *innəwa* require two arguments in possessive constructions. In addition, what is interesting is that the possessor always refers to an entity that is animate because if the possessor is inanimated, it takes the genitive case marker *–ge* and automatically turns to an existential construction rather than a possessive construction. Therefore, we can say that possessive clause construction support Carmen's finding (2005) in the sense that only NP's with animate referents can take the dative case marker. So, this kind of construction has the following basic syntactic structure shown in figure two below.

Possessor + -ṭə	Possessed	Predicate
Noun or pronoun	Noun	*tiyenəwa* or *innəwa*

Figure 2. Syntactic structure of possessive clause with *tiyenəwa* and *innəwa*

Gair and Paolillo (1997), points out that the genitive suffix *–ge* is actually a genitive/locative since it is used as a locative suffix as well. As we just demonstrated above, the dative case marker *-ṭə* works in a similar fashion since it also functions as a locative suffix when indicating possession. Therefore, we believe that the noun with the *-ṭə* marked is actually acting as possessor. Nevertheless, the possessor NP has a locative property, which is intriguing since Lyons (1968) and Clark (1978) point out that possessive, existential and locative constructions are not only strongly related but are also locative in origin. Therefore, it would be of interest to see if this claim applies into the Sinhala locational constructions as well. If it is the case, then, we should expect that existential and locative construction would have not only behavior similar to possessive clauses but also the same verb distinction along animacy lines. That is, the use of *tiyenəwa* or *innəwa* should be determined by the animacy of the located NP.

In order to see if it is the case in Sinhala, we begin with a discussion of existential constructions.

3. EXISTENTIAL CONSTRUCTION. Existential constructions follow a similar pattern to that observed in possessive constructions, since the locative goes before the subject. The main difference is that in this case, the language uses the genitive/locative suffix –*ge* as is in (24)-(27).

(24)   [pingaan-e]~NP~   [pihiy-ak]~NP~   tiye-nəwa
       plate-GEN/LOC   knife-IND   exist-IMPF
       'There is a knife on the plate.'
(25)   [vaaldiy-e]~NP~   [kaasiy-ak]~NP~   tiye-nəwa
       pail-GEN/LOC   coin-IND   exist-IMPF
       'There is a coin in the pail.'
(26)   [vattur-e]~NP~   [maaluv-ek]~NP~ in-nəwa
       water-GEN/LOC   fish-IND   exist-IMPF
       'There is a fish in the water.'
(27)   [gaal-e]~NP~   [harək-ak]~NP~   in-nəwa
       pen-GEN/LOC   cow-IND   exist-IMPF
       'There is a cow in the pen.'

Examples 24-27 above suggest that existential constructions have much in common with possessive constructions. The only main difference is that the latter uses the dative case marker -*ʈə*, while the former use the genetive/locative marker –*ge* (expressed as –*e*). In spite of this difference, the selection of *tiyenəwa* and *innəwa* as a lexical verb in existential constructions depend on the animacy of the referent whose existence is referred to in the sentence. Therefore, there is no doubt that existentials use a structure parallel to the structure of possessives. This structure is of the form shown in figure three below.

Nominal-GEN/LOC	Nominal	Predicate
Noun + -*e*	Noun	*tiyenəwa* and *innəwa*

Table 3. Syntactic structure of existential clauses

Even though, existentials can be related to possessives, Lyons (1968) and Clark (1978) both point out that existential constructions have a very strong relationship to locative constructions since they always requires locative complementation, giving them a locative function. Both of these authors emphasize that what distinguishes existentials from locatives; is mainly the word order of the subject (S) and the Locative (LOC) in both kind of sentences. According to my data clauses 24-27 above can also be order in the way below:

(24a)   [pihiy-ak]~NP~ [piŋgaan-e]~NP~   tiye-nəwa
       knife-IND   plate-GEN/LOC   exist-IMPF
       'A knife is on the plate.'
(25a)   [kaasiy-ak]~NP~   [vaaldiy-e]~NP~   tiye-nəwa
       coin-IND   pail-GEN/LOC   exist-IMPF
       'A coin is in the pail.'

(26a)　　[maaluv-ek]_{NP} [vattur-e]_{NP}　　　in-nəwa
　　　　　fish-IND　　　　water-GEN/LOC　　exist-IMPF
　　　　　'A fish is in the water.'

(27a)　　[harək-ak]_{NP}　[gaal-e]_{NP}　　　　in-nəwa
　　　　　cow-IND　　　pen-GEN/LOC　　exist-IMPF
　　　　　'A cow is in the pen.'

In each one of the examples above, the S has moved to the most left position. So, if we follow Lyons (1968) and Clark's (1978) statement, those clauses should be interpreted as locative constructions since the S precedes the location. However, another alternative analysis would be to take both constructions as two instantiation of a single existential construction. They use the same verbs, and Sinhala has a flexible word order permitting the permutations. Therefore, this alternation is allowed in the language. The two realizations of the same existential construction can then be seen as resulting from the information structure in the clause. That is, if the S is new information, it becomes the focus in the clause and it is posted close to the verb, while the topic (the locative NP) goes before the S (examples 24-27). When the S is the topical or known information and the locative NP is the focus or new information, the S is posted clause-initially and the locative NP gets close to the verb. (examples 24a-27a). In essence the word order clarifies the topic-focus pattern of the information of the clause. These alternations can be summarized as in figure three below.

Topic/Known Info	Focus/New Info	Predicate
Nominal-GEN/LOC	S	*tiyenəwa* and *innəwa*
S	Nominal-GEN/LOC	*tiyenəwa* and *innəwa*

Figure 3. Word order alternation in existential clauses

As with possessive constructional, animacy is clearly the key parameter motivating the choice of the two copular verbs. Nevertheless, if we take a look at the syntactic role, we can see the picture below:

Possessive clauses:　　　Copula-Complement　　　Copula-S　Verb
　　　　　　　　　　　　　　NP-DAT

Existential clauses:　　　Copula-Complement　　　Copula-S　Verb
　　　　　　　　　　　　　　NP-LOC

There is no doubt that both of the clauses are copular constructions since they have the same syntactic structure. The possessed argument in a possessive clause and the located argument in an existential is the copular subject, which form a coherent grammatical class as they are all in the nominative case and they also share the property of determining which of the copular verbs to use. On the other hand, the copula-complement is in the dative case when it is related to the possessor because any time it is related to the location, it takes the GEN/LOC.

Negative constructions confirm the relationship among locational constructions because whenever the negative morpheme *nææ* is incorporated in any of the two kinds of clauses, not only is the verb no longer required by the predicate but the animate distinction is also neutralized.

4. NEGATION IN LOCATIONAL CONSTRUCTIONS. As is common in many languages (Clark 1978), locational constructions (possessive and existential in this case) can also be negated. In Sinhala, the negation of such clauses is very interesting for the following reasons:

  a) When the negative morpheme *næ̃æ* is used in a locational clause, *tiyenəwa* and *innəwa* is no longer required for the clause. Therefore, the two kinds of lexical verbs simply do not appear in this context.

  b) Since the clause does not take any existential verb, the distinction between animate and inanimate is neutralized by the negation.

(28) Mæ̃æri-ṭə   maaluv-ek   næ̃æ
      Mary-DAT   fish-IND   NEG
      'Mary does not have a fish.'

(29) æyə-ṭə   hænd-ak   næ̃æ
      3F-DAT   spoon-IND   NEG
      'She does not have a spoon.'

(30) Daruwa-ṭə   amma   næ̃æ
      child-DAT   mother   NEG
      'The child did not have mother.'   (Gair and Paolillo 1997:62)

(31) Vatur-e   maaluv-ek   næ̃æ
      water-GEN   fish-IND   NEG
      'There is no fish in the water.'

(32) Koop-e   hænd-ak   næ̃æ
      cup-GEN   spoon-IND   NEG
      'There is no spoon in the cup.'

(33) Maaluv-ek   vatur-e   næ̃æ
      fish-IND   water-GEN   NEG
      'The fish is not in the water.'

(34) Hænd-ak   koop-e   næ̃æ
      spoon-IND   cup-GEN   NEG
      'The spoon is not in the cup.'

As is shown above, *næ̃æ* shows up only at the end of the clause just as verbs typically do. Therefore, in this particular case, the semantic meaning of *næ̃æ* is something like 'does not exist'. (that is, X does not exist in Y). Although, *næ̃æ* has verbal properties,[3] it is actually not a verb but a quasi-verb since it cannot be inflected for case or for tense (Gair 1970:38). Nevertheless, for the current study, what is of interest is that this morpheme is working as the predicate of the negative locational clause.

5. CONCLUSION. Sinhala has possessive noun phrases and possessive clauses. However, these constructions do not behave completely alike. Possessive noun phrases always take the genitive suffix *–ge*, while possessive clauses are constructed by two existential verbs: *tiyenəwa* or *innæwa*. However, these verbs are restricted semantically since speakers use *tiyenəwa* when the possessed is inanimate and *innəwa* when the possessed is animate. Nevertheless,

---

[3] It has inflectional possibilities, such as *nætte* 'emphatic', *nætot* 'conditional, and *nætat* 'concessive'.

grammatically speaking what triggers the selection of the verb is the Copula-S. Interestingly, the same restriction is applied in existential constructions. Consequently, there is no doubt that Sinhala treats locational constructions in the same way. The negative construction of these locational clauses (possessive and existential) also confirms the relationship between them because whenever the negative quasi-verb n□□ is incorporated, the clause does not require either of the existential verbs for locational clauses (*tiyenəwa* or *innəwa*). Furthermore, the distinction between animacy is neutralized due to the fact that the animacy of the nominal does not influence the predicate. The relation of possessive and existential clause has been testified in many other languages. Nevertheless, what makes Sinhala an interesting language in this respect is the use of two different existential verbs based on the animacy of the Copula-S argument. In summary, Sinhala provides more information about the general behavior of locational constructions in the languages in the world and therefore it allows for expansion of typological characteristics of such kind of constructions.

REFERENCES

Asher, R.E. 1985. *Tamil.* Australia: Crom Helm descriptive grammars.

Clark, Eve V. 1978. Existential, locative, and possessive construction. *Universals of human language*, ed. by Joseph H. Greenberg, 4.85-126. Stanford: Stanford University Press.

Gair, James W. 1970. *Colloquial Sinhalese clause structures.* Paris: Mouton.

Gair, James W. and John C. Paolillo. 1997. *Sinhala.* Germany: Lincom Europa.

Jany, Carmen. 2005. The relationship between case marking and syntactic role in spoken Sinhala. Paper presented at the Sinhala Workshop, University of California, Santa Barbara.

Lyons, John. 1968. *Introduction to Theoretical Linguistics.* Great Britain: Cambridge University Press.

Payne, Thomas E. 1997. Describing Morphosyntax. *A guide for field linguists.* United Kingdom: Cambridge University Press.

University of California, Santa Barbara
Department of Linguistics
3607 South Hall
Santa Barbara, CA 93106
Salome@umail.ucsb.edu

BETWEEN LEXICAL AND LEXICO-GRAMMATICAL CLASSIFICATION: NOMINAL CLASSIFICATION IN SINHALA

MARA HENDERSON

University of California, Santa Barbara

1. INTRODUCTION. This paper investigates a small set of specific-general noun sequences (SGNs) and their role as a system of nominal classification in Sinhala. Systems of nominal classification have typically been described in terms of three sub-types: 1) lexical systems (class terms and measure terms), 2) lexico-grammatical systems (classifiers), and 3) grammatical systems (noun class markers and gender) (Grinevald and Seifart 2004:261). These three subtypes can be seen as occupying positions on a typological continuum as well as reflecting a diachronic pattern of language change from class terms or measure terms to noun class markers of agreement or gender as illustrated by Figure 1, below.

| | Lexical (class/measure terms) | Lexico-grammatical (classifiers) | Grammatical (noun class markers/gender) |
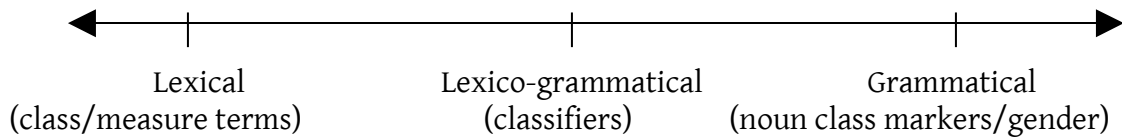
FIGURE 1. Systems of nominal classification (Grinevald and Seifart 2004:261)

Each of these systems and the patterns by which they may be distinguished from each other are elucidated in section 2. For the purposes of this paper, I focus on differentiating semantic and morphosyntactic patters of lexical and lexico-grammatical systems of nominal classification in an attempt to explicate the role of the general noun of SGNs in Sinhala.

Sinhala is an Indo-Aryan language spoken primarily in Sri Lanka (Ethnologue 2004). According to statistics from 1993, approximately 72 percent of the population of Sri Lanka are native Sinhala speakers (Ethnologue 2004). The Sri Lankan language community is primarily made up of Sinhala, English, and Tamil speakers. Sinhala functions as the language of most domains (i.e. government, marketplace, Buddhist temples), while English functions as the lingua franca in private business and education. Sinhala has been described as possessing one system of nominal classification; namely a system of animacy and honorific marking on numerals (Gair and Paolillo 1997:22). Gair and Paolillo (1997:22) describe this as a gender system--that is, it is a system of agreement between the noun and the numeral which quantifies it. A partial reconstruction of that system is as follows:

		Inanimate		Animate	
	Stem	Definite	Indefinite	Definite	Indefinite
one	ek	ekə	ekak	ekenaa	ek**kenek**/**kenek**
two	de	deka	dekak	de**nna**	de**nnek**
three	tun	tunə	tunak	tun**denaa**	tun**denek**
four	hatərə	hatərə	hatərak	hatərə**denaa**	hatərə**denek**

TABLE 1. Numeral gender system in Sinhala (Gair and Paolillo 1997:22)

In this system numerals are obligatorily marked for animacy and definiteness. Here animacy is more aptly described as humanness as the animate forms only appear with humans and not other animate beings like animals. The forms for 'one' have complex patterns of use in Sinhala, which I do not discuss further other than to mention that they are likely a result of a contrast between 'one' and general singular indefiniteness. Other systems of nominal classification, specifically lexical or lexico-grammatical systems have not been described for Sinhala. From a typological perspective, although Sri Lanka is not a linguistic area identified as possessing nominal classification systems, Aikhenvald (2003:77-78, 121-122) does include languages spoken in southern India in her typological study of noun class markers and numeral classifiers. Furthermore, Emeneau (1956:10) identifies Indo-Aryan as the historical source of nominal classification (specifically, noun class markers, measure terms, and numeral classifiers) in Dravidian and Munda languages of India. The presence of lexical systems of nominal classification is perhaps not all that surprising, as they are common crosslinguistically. However, lexico-grammatical systems are far more restricted--that is, they are typically described as an areal phenomenon with a high concentration in Southeast Asia. In fact, Emeneau (1956:16) notes the possibility of Southeast Asia as the source of classifiers in Indo-Aryan. It is, therefore, not too surprising to find a lexico-grammatical system of nominal classification emerge from a study of Sinhala. The evidence provided in this paper suggests that a system of nominal classification not unlike a classifier system indeed exists in a limited semantic domain of language use in Sinhala.

2. NOMINAL CLASSIFICATION SYSTEMS. The terminology used to discuss nominal classification systems typologically is not consistent in the literature. In particular, researchers tend to use the term CLASSIFIER to describe both lexico-grammatical and grammatical systems of classification. For the purposes of this paper, I use the term NOMINAL CLASSIFICATION as a broad cover term to mean a system through which language or language users mark nouns based on categories, which would include class terms, measure terms, classifiers, and noun class markers. CLASS TERMS occur as part of endocentric nominal compounds in which the class term is taken from a higher position in the taxonomy than the other element in the compound, which specifies the type (DeLancey 1986:440). In English, for example, *snake* functions as a class term in compounds like *rattlesnake, king snake,* and *grass snake* where *snake* denotes the basic category and *rattle, king*, and *grass* denote the type of snake (DeLancey 1986:440). MEASURE TERMS are terms that denote a quantity of the entity they modify. In English, for example, *pound* functions as a measure term in phrases like *a pound of butter, a pound of sugar,* and *a pound of oranges* (Grinevald and Seifart 2004:261). CLASSIFIERS are defined broadly as "morphemes which occur 'in surface structures under specifiable conditions', denote 'some salient perceived or imputed characteristics of the entity to which an associated noun refers' (Allan 1977:285), and are restricted to particular constructions types known as 'classifier constructions'" (Aikhenvald 2003:13). Aikhenvald defines CLASSIFIERS CONSTRUCTIONS as "morphosyntactic units...which require the presence of a particular kind of a morpheme, the choice of which is dictated by the semantic characteristics of the referent of the head of a noun phrase" (2003:13). This definition of classifiers is decidedly broad to include a full range of classifier types, however, only two classifier types (numeral and noun) are of particular salience for this discussion of Sinhala nominal classification. NOUN CLASSIFIERS are morphemes in classifier constructions that appear in the noun phrase, typically next to the noun, and typically denote generic semantic characteristics of the noun they categorize, such as, men, women, plants and

animals (Grinevald and Siefart 2004:262-263). Dixon (1982:1992 ff. in Aikhenvald 2003:2) provides an example of a noun classifier in Yidiny:

(1)   bama   waguja
            CL:PERSON   man
      'a man'

NUMERAL CLASSIFIERS too are morphemes in classifier constructions that appear in the noun phrase, however these classifiers occur in numeral phrases and typically denote characteristics of the physical shape of the entity they categorize, such as, 1D long-rigid, 2D flat-flexible, 3D round. Rehg (1981:130 in Grinevald and Seifart 2004:262) provides the following examples of a numeral classifiers in Ponapean:

(2)   tuhke rioapwoat
            tree   two.CL:LONG
            'two trees'
 (3)   pwihk riemen
            pig   two.CL:ANIMATE
            'two pigs'

These examples illustrate the diversity in semantic denotation of numeral classifiers. Although prototypically numeral classifiers denote physical properties such as shape, they have also been found to denote animacy. NOUN CLASS MARKERS (aka, NOUN CLASSES, CLASS MARKERS, GENDER, CONCORDIAL CLASSIFIERS) are "an obligatory grammatical system where each noun chooses one from a small number of possibilities" (Dixon 1986:105). Aikhenvald further states that they are "grammatical agreement classes, based on such core semantic characteristics as animacy, sex, or humanness" (2003:1). An example of noun class markers is found in Portuguese (Aikhenvald 2003:2):

(4)   o   menin-o   bonit-o
            ART.MASC.SG   child-MASC.SG   beautiful-MASC.SG
      'the beautiful boy'

This example clearly illustrates the agreement function of noun class markers through the masculine singular form that appears on the article, noun, and the adjective in the noun phrase.

    Many of these nominal classification systems share semantic and morphosyntactic characteristics which makes distinguishing them from each other rather difficult. The following table is an attempt to clarify the characteristics of each system that may distinguish them from each other.

	Class terms (CT)	Measure terms (MT)	Classifiers (CL)	Noun class markers (CM)
Size	*open; restricted*	*open; restricted*	open, varies; large	small finite set
Distribution	?	?	1N:>1CL some Ns may not take a CL	1N:1CM all
Boundedness	lexically bound; may be but, often not independent nouns	*varies; free form, affix*	free form; occurs in the same NP as the N it qualifies; not independent noun/independent noun	closed grammatical system; affixes, GW, clitics
Scope	*noun*	*noun*	never any reference outside the NP	marking is never entirely within the noun word
Semantics	consistent, circumscribed, hyponym	*provide the measure for a specified quantity*	animacy, shape, functional, generic	animacy, sex, humanness
Inter-speaker Variation	?	?	use varies across registers or styles	little variation between speakers

TABLE 2. Distinguishing nominal classification systems[1]

The following elucidates Table 2.

(a)   Size refers to whether or not it is an open or closed class and in general the number of terms typically found in these kinds of systems cross-linguistically.   An explicit discussion of size for class terms and measure terms was not found, however, I believe that it is safe to state that these are open classes, but typically restricted.   Dixon (1986:106) describes classifiers as typically quite large crosslinguistically (50-400), although there are languages such as Indonesian which have very small sets of classifiers (7)[2].   This is additionally, highlighted by Aikhenvald (2003:81) in her discussion of noun classifiers, she states that the size of the inventory may vary crosslinguistically from a small closed set to a large open set.

(b)   Distribution refers to which nouns in the language take the classifying morpheme. Information was not found regarding class terms and measure terms, however, Dixon (1986) and Aikhenvald (2003) provide some typical characteristics of classifiers and noun class markers.   Dixon (1986:106) provides that typically in languages with classifiers systems not all nouns take classifiers.   Nouns that do not take classifiers are typically mass

---

[1]The information compiled for this table was taken from Dixon (1986:105-108) and Aikhenvald (2003), which explicates the distinctions between noun class markers and classifiers, and DeLancey (1986:440-444), which discusses some differences between class markers and classifiers. The information in italics is my own educated guess and question marks indicate that this information was not found and is therefore unknown.

[2] Thank you to Robert Englebretson for pointing out this particular example.

nouns, time units, and some of the most frequent nouns (Dixon 1986:106). Additionally, a single noun typically is able to take more than one classifier with a resulting change in meaning (Aikhenvald 2003:81, 98, Dixon 1986:106, Greenberg 1972:8). Noun class markers, on the other hand, classify ALL the nouns in a language--that is, there is a 1 to 1 ration of nouns to noun class markers and their distribution is fixed (Aikhenvald 2003: 21, Dixon 1986:106).

(c)   Boundedness refers to the classifying morpheme's realization as bound or free. By definition class terms are lexically bound and may function as independent nouns in other contexts, though they often do not (DeLancey 1986:439). However, these properties are again best described as tendencies since class terms may occur as independent nouns and are not obligatory in all cases (DeLancey 1986:439). The realization of measure terms crosslinguistically is bound or free. Dixon claims that "noun classifiers are always separate lexemes, which may be included with a noun in certain syntactic environments" (1986:105). However, according to Aikhenvald noun classifiers may appear as clitics or nominal affixes via grammaticization or phonological reduction processes (2003:91, 101). Noun class markers typically emerge as affixes, grammatical words, or clitics (Dixon 1986:106). They are often fused with other grammatical morphemes such as definiteness, case, or number (Dixon 1986:106). In Delancey's work on Tai class terms and classifiers, he alludes to a prototypical property of classifiers, which states that classifiers would not function as independent nouns or part of compounds (1986:439). However, Greenberg states "in the majority of instances, the classifier is itself a noun with its own lexical meaning and may, in fact, have its own classifier when it functions as the head of a noun phrase" (1972:7). Conflicting findings on the function of class terms and classifiers as independent nouns within classifying languages provides evidence that a crosslinguistic explanation of the terms' ability to operate as independent nouns is not a defining feature and therefore should not be heavily weighted in distinguishing it from other nominal classification systems.

(d)   Scope refers to the classifying morpheme's domain--that is, the noun itself, the noun phrase, or outside of the noun phrase. I believe that it is safe to state that the scope of class terms and measure terms is the noun. Classifiers are specifically distinguished from noun class markers based on scope. According to Dixon (1986:106-107) noun class marking is "never entirely within the noun word" rather other elements in the sentence are obligatorily marked with the same marker, while classifiers are never referenced outside of the noun phrase (Dixon 1986:107, Aikhenvald 2003:81, 98).

(e)   Semantically, there is some degree of overlap between the categories. Class terms are described generally as semantically consistent and circumscribed (DeLancey 1986:441). Furthermore their relationship with the entity they categorize tends to be taxonomic. Measure terms can be simply described as denoting the measure for a specified quantity. Classifiers range semantically from animacy, shape, generic, or function (i.e. clothing, transport, food), often depending on the type of classifier (Aikhenvald 2003:1-2, Grinevald and Seifart 2004:263-264). Noun class markers tend to denote "such core semantic characteristics as animacy, sex, and humanness" (Aikhenvald 2003:1).

(f)   Finally, Dixon (1986:107) also notes a tendency for inter-speaker variation in the use of classifiers based on register or style shifts, while little variation between speakers is found in systems of categorization described as noun class marker systems.

It is clear from the details discussed above that SGNs in Sinhala are not noun class markers.

The general noun of SGNs in Sinhala do not distribute obligatorily in all cases or with all nouns, nor do they reference any other element of the sentence beyond the noun phrase. In these ways, they are clearly not operating as agreement systems and therefore not noun class markers. However, they do display characteristics typical of class terms, measure terms, noun classifiers and numeral classifiers. Because SGNs exhibit properties of lexical and lexico-grammatical systems of classification it is important to this analysis to focus on the properties that distinguish these system types, namely their realization and semantic relationship to the entity they categorize. Specifically, class terms are lexically bound and measure terms and classifiers vary in boundedness crosslinguistically. Therefore, if it can be established that the general noun of SGNs in Sinhala are not bound, then they can be distinguished from class terms. To determine this, I focus on properties of wordhood, obligatoriness, and anaphoric reference of the general term of SGNs in Sinhala (§4). Although there is some overlap, as discussed above, the semantic relationships between the classifying term and the entity they categorize differ for prototypical class terms, measure terms and classifiers. These prototypical patterns are discussed in relation to the semantic patterns of SGNs in Sinhala to aid in the analytical distinction between them (§4).

While wordhood, obligatoriness, and anaphoric reference provide important features that could distinguish lexical from lexico-grammatical systems, lexico-grammatical systems themselves can be further subdivided into types of classifiers. Most important to the analysis here is the distinction between noun classifiers and numeral classifiers. Therefore, some discussion of the properties which distinguish the two subtypes of relevant classifiers is necessary before moving on to the specifics of SGNs in Sinhala.

Noun classifiers and numeral classifiers share several of the same properties: a) they appear in the noun phrase, b) their selection is based on semantic properties of the entity they categorize, c) their level of grammaticization varies, d) they are characterized as open lexical classes, e) there is evidence of inter-speaker variation, f) some nouns do not take the classifier, while others may vary the classifier with a resulting change in meaning, g) they are typically realized as free forms, and h) they may be used for anaphoric reference (Aikhenvald 2003:81, 98, Greenberg 1972:6). The distinguishing property is their specific location within the noun phrase and their tendency toward types of semantic categorization. As previously mentioned, noun classifiers typically denote generic semantic categories, while numeral classifiers typically denote animacy or physical properties (i.e. size, shape, structure) (Aikhenvald 2003:98, Grinevald and Seifart 2004:262-263). Numeral classifiers more specifically occur in quantifying expressions and numeral noun phrases (Aikhenvald 2003:98). Greenberg (1972) further points to the individuating function of numeral classifiers, a function that has not be claimed for noun classifiers. He states

[I]n the usual classifier language...classifiable nouns in their isolated form, that is when not accompanied by a classifier or a plural marker, are like collectives in their semantic non-specification of number and in their avoidance of a direct number construction. The classifier is an INDIVIDUALIZER which performs the same function as a singulative derivational affix in languages with the collective/singulative opposition (Greenberg 1972:26, emphasis added)

This individuating function along with the tendencies of a semantic denotation of shape and the syntactic distribution in numeral phrases distinguishes numeral classifiers from the more generic non-individuating noun classifiers.

The implications for an analysis of SGNs as members of a nominal classification system based on semantic properties and relationships between elements of SGNs are discussed further after a brief description of the overall patterns of SGNs in Sinhala in section 3, below.

3. Description of SGNs in Sinhala. A set of nine terms that appear to fulfill a classificatory function have been identified through the examination of elicited sentences from one Sinhala speaker. The nine terms are as follows in Table 3.

geḍi	fruit-like.thing.PL	*palu*	section.PL
mal	flower.PL	*æṭə*	seed.PL
karal	pod-like.thing.PL	*kææli*	piece.PL
peti	flat.thing.PL	*kæṭə*	block.PL
alə	potato.PL		

Table 3. Identified classificatory terms in Sinhala

These terms appear in noun phrases following a more specific noun which they classify as in the following example.

(1) **hatu**   **mal**   narakwelaa
  mushroom   **flower.PL**   rotten
  'The mushrooms are rotten.'

3.1. Semantic Properties.
Semantic Domains. SGNs have so far appeared largely in the domain of food, but they also occur with plants, medicine, and other small objects such as dice and beads. Some examples of their semantic distribution is as follows:

dehi geḍi	'limes'	*boonci karal*	'green beans'
vambotu geḍi	'eggplants'	*behet karal*	'capsules'
keek geḍi	'cakes' (whole)		
paaŋ geḍi	'loaf of bread'	*dehi æṭə*	'lime seeds'
		wii æṭə	'rice grains'
mannel mal	'blue lotuses'	*pabəlu æṭə*	'beads'
hatu mal	'mushrooms'		
kehel mal	'banana stalks'	*pipiŋña kææli*	'chopped cucumbers'
		bætəri kææli	'batteries'
daadu kæṭə	'dice'		
ais kæṭə	'ice cubes'		

dodaŋ palu	'orange segments'	*roosə peti*	'rose petals'
suduluunu palu	'garlic cloves'	*behet peti*	'tablets'
		maalu peti	'fish fillets'

The SGNs above show that *geḍi* may occur with fruit, vegetables, and whole breads of a certain shape. Although *geḍi* is often thought of as meaning fruit, evidence shows that this term may be used with other items that possess properties often thought of as pertaining to fruit, but not necessarily only fruit. According to the consultant, a more precise semantic description of *geḍi* would be a small, hard, fruit-like things. The terms *mal* and *æṭə* appear with all types of things they denote (i.e. flowers and seeds), but also appears with things that are judged to come in the form of flowers (e.g. stalks of bananas and mushrooms) or seeds (e.g. beads). The terms *kæṭə, palu*, and *karal* have so far been found with only a few terms. It is unclear what their full semantic distribution is, however, they appear to denote the shape of the items they categorize. The last two terms *peti* and *kææli* have a much wider semantic distribution. In many cases, *kææli* denotes a changed state as in the example above (*pipiŋña kææli* 'chopped cucumbers') or *darə kææli* 'chopped wood'. It may even appear with mass nouns as in *harakmas kææli* 'beef pieces'. However, it also appears with small items that come in sets, such as batteries (*bætəri kææli*) or magnets (*kandaŋ kææli*), in which case the term is used to individuate a single or number of items from the set. A similar case arises from the patterns of *peti*, which may denote a change of state to a mass noun as in *harakmas peti* 'slices of beef' or *maalu peti* 'fillets of fish'. Like *kææli*, it also appears with items that do not undergo a change in state, but rather occur in groups, such as medicine tablets, *behet peti*, and flower petals, *mal peti.*

SEMANTIC RELATIONSHIPS BETWEEN ELEMENTS. The SGNs in this study fall into three types of semantic relationships: 1) hypernym-hyponym taxonomic relationships, 2) noun + shape/physical property denotational relationships, and 3) noun + quanification relationships. The following examples illustrate these findings.

Hypernym-hyponym taxonomic relationships
(2)	*arəliyə mal*	'frangipani flowers'
(3)	*vattakka æṭə*	'pumpkin seeds'
(4)	*batələ alə*	'yam potatoes'

Noun + shape/physical property denotational relationships
(5)	*goowə geḍi*	'cabbages' (lit. cabbage small.hard.fruit-like.things)
(6)	*boonci karal*	'long beans' (lit. bean long.pod-like.things)
(7)	*daadu kæṭə*	'dice' (lit. die blocks)

Noun + quantification relationships
(8)	*dehi palu*	'lime sections'
(9)	*kukulmas kææli*	'pieces of chicken'
(10)	*maalu peti*	'fillet of fish'

It should be noted here that these examples are representative of the semantic tendency of the relationship between the elements for each of the nine classifying terms; however, a few instances have been found that pattern outside of a single term's general tendency (e.g. *kehel*

mal 'banana stalks', *hatu mal* 'mushrooms', *pabəlu ǽʈə* 'beads' *gammiris ǽʈə* 'peppercorns' *wii ǽʈə* 'rice grains' *kaju ǽʈə* 'cashews'). Because in an overwhelming majority of cases, *ǽʈə* and *mal* appeared in SGNs that could be categorized as having a taxonomic relationship with the specific noun they classified, I attribute the occurrence of these few terms to semantic extension.

INDIVIDUATION VS. NON-INDIVIDUATION. These classificatory terms appear with some but not all count nouns in the semantic domains previously mentioned. The most striking semantic motivation for the presence or absence of the general noun is the degree of individuation of the referent.

(11) **sudu-luunu** suddə-kara-nnə amaarui
   white-onion peel-do-INF   difficult
   'Garlic cloves are hard to peel.'

(12) mee **sudu-luunu** **(palu)**   suddə-kara-nnə amaarui
   these **white-onion** **section.PL** peel-do-INF   difficult
   'These garlic cloves are hard to peel.'

(13) maŋ laŋgə **sudu-luunu** **paluwak**   tiye-n´wa
   1SG   near **white-onion** **section.SG.IND** exist-IMPF
   'I have a clove of garlic.'

(14) **kǽrət** tiye-nəwa=də
   carrot exist-IMPF=Q
   'Do you have carrots?'

(15) maŋ laŋgə **kǽrət** **(alə)**   dahayak tiye-nəwa
   1SG   near **carrot** **potato.PL** ten.IND exist-IMPF
   'I have ten carrots.'

In examples 11 and 14 the referent is non-individuated. In these cases, the general term is dispreferred. However, when specifying a set or a number of the referent, the general term may appear as in examples 12, 13 and 15. Examples 12 and 15 show that the general term in these cases are not entirely required. Further, although the general term in example 15 is judged to be optional, my consultant states that he prefers that the general term appear in constructions such as this one.

VARIATIONS. The general nouns of the SGNs may vary with one specific noun with a resulting change in meaning as in the following example.

(16) kehel geḍi   kehel mal
   banana fruit.PL   banana flower.PL
   'bananas'   'banana stalks'

However, there does appear to be a default general term for each specific noun. This is evidenced by the patterns of obligatoriness--that is, while the default term may be optional in some cases, if the specific noun occurs with a general noun other than the default term it becomes obligatory. For example, in example 16 above, *geḍi* is the default term for bananas. It is therefore optional in some cases. However, *mal* is not the default term for bananas; it is

therefore, required in all cases--that is, *kehel* alone may not mean banana stalks, it may only mean bananas.

In sum, SGNs in Sinhala occur in a small semantic domain, primarily with count nouns and only with mass nouns when the mass noun undergoes a change of state that produces countable pieces. Three types of semantic relationships between elements emerge: 1) taxonomic, 2) the general term denotes the shape or physical property of the referent, or 3) the general term quantifies the referent. Further, the general terms serve to individuate a set or number of referents. The general nouns may alternate with a single noun with a change in meaning and one general term operates as the default term for a specific noun. Before moving on to the implications of these patterns for the analysis of SGNs as a system of nominal classification in Sinhala, I discuss the morphosyntactic patterns of SGNs in Sinhala.

3.2. MORPHOSYNTACTIC PATTERNS

MORPHOLOGICAL MARKING. In most cases, nominal morphology, such as case, definiteness, number or question marking, may only appear on the general noun, which in these cases are obligatorily present in the noun phrase as in the following examples.

(17)   mæssa   **miris**   **karələkə**   wæhuw-a
       fly.SG   chili.pepper   pod.like.thing.SG.IND.LOC   land-PST
       'The fly landed on a chili pepper.'

(18)   *mæssa   **miris-yəkə**   wæhuw-a
       fly.SG   chili.pepper-SG.IND.LOC   land-PST
       'The fly landed on a chili pepper.'

(19)   mona   **paaŋ**   **geḍiyə=də**   narakwelaa   tiye-nne
       which   bread   fruit.like.thing.SG.DEF=Q   rotten   exist-FOC.NPST
       'Which loaf of bread is rotten?'

(20)   *mona   **paaŋə=də**   narakwelaa   tiye-nne
       which   bread.SG.DEF=Q   rotten   exist-FOC.NPST
       'Which loaf of bread is rotten?'

However, some specific lexical items have been found to carry the nominal morphology, such as pineapples or mangos. In these cases, the classifying term may be omitted.

(21)   mee   **annaasi**   **geḍi-yə**   pæniraha-i
       1PROX   pineapple   fruit.like.thing-SG.DEF   sweet-PRED
       'This pineapple is sweet.'

(22)   mee   **annaasi-yə**   pæniraha-i
       1PROX   pineapple-SG.DEF   sweet-PRED
       'This pineapple is sweet.'

Most of the examples elicited involving nominal morphology required the presence of the general noun. Examples such as 22 above were rare.

CONSTITUENCY. SGNs display two patterns of constituency--that is, they are cohesive and they move as a unit. As of yet, SGNs always appear together without any intervening lexical or

grammatical formatives. Furthermore, SGNs move as a unit as illustrated in the following examples.

(23)   maŋ   laŋgə   **dehi**   **geḍi**   tiye-nəwa
      1SG   near   **lime**   **fruit.PL**   exist-IMPF
      'I have limes.'

(24)   **dehi**   **geḍi**   maŋ   laŋgə   tiye-nəwa
      lime   **fruit.PL**   1SG   near   exist-IMPF
      'I have limes.'

The specific and general nouns may not be separated throughout the phrase and no lexical or grammatical formatives have been found to intervene between them. These patterns of cohesiveness and movement provide evidence for the analysis of SGNs as a lexical constituent.

    Related to constituency is the general terms' ability to operate as independent nouns. Although a couple of these nouns were judged to not operate as independent nouns (*palu* and *geḍi*), they have been found outside of SGN contexts in the following examples.

(25)   atə   tiyenne   janeele   dakunu   **paluwe**
      hand.SG.DEF   exist-FOC.NPST   window.SG.DEF   right   section.LOC
      'The hand is on the right section of the window.'

(26)   **alə**   wələ   **æṭə**   nææ
      potato.PL   seed.PL   NEG
      'Potatoes don't have seeds.'

(27)   wiiduruwə   **kææli**   wələṭə   kædun-a
      glass.SG.DEF   piece.PL   break-PST
      'The glass broke into pieces.'

(28)   **mal**   lasənai
      flower.PL   beautiful-PRED
      'The flowers are beautiful.'

(29)   annaasi   kiyanne   **geḍiyak**
      pineapple   known.as   fruit-like.thing.IND
      'The thing known as a pineapple is a small, hard fruit like thing.'

According to my consultant, the nouns *alə*, *æṭə*, *kææli*, and *mal* are easily identifiable as independent nouns. However, according to my consultant, *palu* and *geḍi* are not typically thought of as independent nouns although, as example 25 and 29 show *palu* may be used in a possessive construction that individuates the thing possessed (the window's section) and *geḍi* may be found independently when talking specifically about the category. The other three terms *peti*, *kæṭə*, and *karal* are judged as unable to operate as independent nouns and no data has been found to the contrary.

    ANAPHORIC REFERENCE. SGNs also display patterns of anaphoric reference. SGNs may be anaphorically referenced by inanimate pronouns, such as *eekə* or *eeva*, or by the general noun in individuating contexts as illustrated by the following examples.

(30)   **peera**   **(geḍi)**   narakwelaa
       guava   (fruit-like.things)   rotten
       'The guavas are rotten.'

(31)   **eeva**   narakwelaa
       3p.ɪɴᴀɴ   rotten
       'They're rotten.'

(32)   **mee**   narakwelaa
       these   fruit-like.things   rotten
       'These are rotten'

(33)   ?***geḍi**   narakwelaa
       3p.ɪɴᴀɴ   rotten
       'They're rotten.'

Example 31 illustrates the pronominalization of guavas from example 30 using the third person inanimate pronoun, *eeva*. Although, this is the most preferred form, the consultant also provided the example given in 32 as an alternative employing the general term of the SGN in 30 to refer to a specific set of guavas. Here the context is more individuated and therefore, the anaphoric function, illustrated in the gloss, of the general term is judged acceptable. Finally, example 33 illustrates a case in which the consultant judged the use of the classifying term without an individuating deictic as highly dispreferred. However, a couple of examples were obtained in which the general term of an SGN could be function anaphorically without an individuating deictic, as illustrated below.

(34)   waṇḍura   **kehel**   **malə**   uḍə-ṭə   pænn-a
       monkey   **banana**   **flower. sɢ.ᴅᴇꜰ**   top-ᴅᴀᴛ   jump-ᴘsᴛ
       'The monkey jumped on the stalk of bananas.'

(35)   **eeka**   uḍə-ṭə   waṇḍura   pænn-a
       3s.ɪɴᴀɴ   top-ᴅᴀᴛ   monkey   jump-ᴘsᴛ
       'The monkey jumped on it.'

(36)   **malə**   uḍə-ṭə   waṇḍura   pænn-a
       3s.ɪɴᴀɴ   top-ᴅᴀᴛ   monkey   jump-ᴘsᴛ
       'The monkey jumped on it.'

Although anaphoric reference itself is not a test for constituency, example 36 displays a pattern similar to the pronominalization in 35, a classic test for constituency. Here the general term of the SGN functions as a pronominal, as indicated in the gloss. This pattern provides further evidence for analyzing SGNs as lexical units. Though it must be noted that this is pattern emerged only with a select few examples and is not representative of the patterns of SGNs more generally as shown in examples 31-33.

    The morphosyntactic patterns discussed above demonstrate that 1) in the vast majority of cases nominal morphology may not appear on the specific noun, rather the general term is required in cases where nominal morphology must occur on the noun or noun phrase, 2) SGNs operate as lexical constituents based on patterns of cohesiveness and movement. Additionally, the general term was shown to function anaphorically for the SGN in individuating contexts.

4. Sɪɴʜᴀʟᴀ'ꜱ Sʏꜱᴛᴇᴍ ᴏꜰ Nᴏᴍɪɴᴀʟ Cʟᴀꜱꜱɪꜰɪᴄᴀᴛɪᴏɴ. Many of these properties of SGNs in Sinhala are properties of nominal classification systems as discussed in section 2 above. In this section, I aim to describe how Sinhala's nominal classification system fits into the larger framework of nominal classification systems described in the literature. I begin by discussing how the patterns of SGNs implicate an analysis of them as lexical or lexico-grammatical systems. I conclude this section with a discussion of the semantic properties of SGNs that implicate their placement among subtypes of lexico-grammatical classification systems.

4.1. Dɪꜱᴛɪɴɢᴜɪꜱʜɪɴɢ Lᴇxɪᴄᴀʟ ꜰʀᴏᴍ Lᴇxɪᴄᴏ-Gʀᴀᴍᴍᴀᴛɪᴄᴀʟ Sʏꜱᴛᴇᴍꜱ ɪɴ Sɪɴʜᴀʟᴀ. Class terms and measure terms, though both lexical systems, possess quite different defining characteristics. As a result, I discuss these separately in relation to classifiers. Differentiating class terms from classifiers is accomplished by comparing SGNs and compounds in Sinhala based on three morphosyntactic patterns: 1) wordhood, 2) obligatoriness, and 3) anaphoric reference. However, measure terms are best differentiated from classifiers based on the semantic function of individuation.

Cʟᴀꜱꜱ Tᴇʀᴍꜱ ᴠꜱ. Cʟᴀꜱꜱɪꜰɪᴇʀꜱ. As discussed in section 2, class terms are by definition part of compounds, and classifiers are prototypically separate lexemes (Dixon 1986:105). With this in mind, the first step for distinguishing class terms from classifiers is to determine SGN's status as a word. The fact that SGNs operate as a single unit and that nothing has been found to intervene between the specific and general nouns, complicates their differentiation from compounds. However, by examining language internal patterns of wordhood, obligatoriness of elements, and patterns of anaphoric reference, we may contrast SGNs with compounds in Sinhala.

For this part of the analysis, I focus on the contrast between a few general terms of SGNs (*æṭə*, *karal*, *mal*, *alə*) and a few compounds that use the same general terms as the second element in the compound. These compounds are: *muŋæṭə* 'mung beans', *mæækaral* 'Chinese long beans', *innələ* 'potato' (particular kind), *muhudumal* 'coral'.

Wᴏʀᴅʜᴏᴏᴅ. According to Dixon and Aikhenvald, a compound is a single grammatical word made up of one or more phonological words (Dixon and Aikhenvald 2002:19). It therefore may be useful to investigate phonological and grammatical criteria for wordhood for both SGNs and compounds in Sinhala. If SGNs are shown to operate as more than one grammatical word, then we can confidently say that the general noun is not part of a compound and therefore not a class term. However, if SGNs are found to operate as a single grammatical word then other methods of distinguishing them from compounds must be explored. The general nouns would not automatically be discounted from being classifiers since, as Aikhenvald points out, classifiers may emerge in various stages of grammaticization and therefore do not always appear as separate words. In this case, more evidence would be needed to assert their status as classifiers.

Dixon and Aikhenvald provide a set of crosslinguistic criteria for determining grammatical and phonological wordhood (2002:19-21). Grammatical words are identified as having the following universal criteria: a) cohesiveness (the elements always occur together, b) a fixed order, and c) a conventionalized and coherent meaning (2002:19). Universal criteria for phonological words are not as easily explicated. However, Dixon and Aikhenvald point to stress, phonotactics, and phonology the primary areas where distinct patterns may be found

for language internal criteria for phonological wordhood. In the case of Sinhala, although there are some interesting isolated cases where phonological criteria points to one analysis over the other, there is no overwhelming evidence that phonological wordhood is crucial to the distinction between compounds and SGNs.

While the criteria for grammatical wordhood could potentially provide key evidence for the classification of the general nouns of Sinhala's SGNs as class terms or classifiers, the morphosyntactic patterns of SGNs and compounds are remarkably similar. As I have shown above, SGNs occur in a fixed order (specific noun followed by general noun), when both elements are required they always occur together as evidenced by their operation as a syntactic unit that may not be separated by other lexical or grammatical formatives, and they have conventionalized and coherent meanings as a unit as evidence by change in meaning accompanying variations in general terms with specific nouns. Therefore, according to Dixon and Aikhenvald's universal criteria, SGNs qualify as a single grammatical word. Furthermore, as expected the lexical items identified as compounds in Sinhala for this study also conform to this set of criteria. The following examples illustrate the similarity in grammatical wordhood status.

(37) **muhudumal**   lasənai
     coral.PL   beautiful
     'The coral is beautiful.'
(38) **hatu**   **mal**   narakwelaa
     mushroom   **flower.**PL   rotten
     'The mushrooms are rotten.'

In example 37, the order of the elements in the compound *muhudumal* may not be reversed, nor may they be separated by other grammatical or lexical formatives. Furthermore, the elements as a unit have a coherent and conventional meaning. In these ways, the SGN, *hatu mal* in example 38 is similar to the compound in example 37. The similarity in the patterns of wordhood between SGNs and compounds leads to the investigation of other patterns which may distinguish SGNs from compounds.

Obligatoriness. The second morphosyntactic pattern that may provide evidence for the analysis of SGNs in Sinhala is patterns of obligatoriness of the elements of SGNs and compounds. Although the obligatoriness of the elements of compounds varies cross-linguistically, we may expect that the conditions for the omission of elements to be fairly restricted since the elements are by definition lexically bound. Classifiers, on the other hand, have been identified as being optional in many languages (Greenberg 1972:6). Therefore, if the general nouns of SGNs are classifiers rather than class terms, we would expect patterns of obligatoriness to be more restrictive for compounds.

The general nouns of SGNs are optional in all cases in which the specific noun is not being individuated or is otherwise morphologically unmarked. However, the elements of compounds are for the most part not optional in Sinhala. The following examples illustrate the distinct patterns of SGNs and compounds.

(39)   oyaa   mæækaral   wikunə-nəwa=də
       2SG   Chinese.long.bean.PL   sell-IMPF=Q
       'Do you sell Chinese long beans?'

(40)   oyaa   miris   wikunə-nəwa=də
       2SG   chili.pepper sell-IMPF=Q
       'Do you sell chili peppers?'

(41)   oyaa   muŋæʈə   wikunə-nəwa=də
       2SG   mung.seed.PL   sell-IMPF=Q
       'Do you sell mung beans?'

(42)   oyaa   pabəlu   (æʈə)   wikunə-nəwa=də
       2SG   bead   seed.PL   sell-IMPF=Q
       'Do you sell beads?'

(43)   oyaa   innələ   wikunə-nəwa=də
       2SG   potato.PL   sell-IMPF=Q
       'Do you sell potatoes?' (a particular type)

(44)   oyaa   kærət   wikunə-nəwa=də
       2SG   carrot   sell-IMPF=Q
       'Do you sell beads?'

Examples 39, 41, and 43 demonstrate that the general term in these compounds is obligatory even when the referent is non-individuated. As examples 40, 42, and 44 illustrate, the general term as a part of SGN constructions is either dispreferred or optional in cases when the referent is non-individuated.

ANAPHORIC REFERENCE. Patterns of anaphoric reference may point to an analysis of the general terms as classifiers or class terms. While classifiers have been described typologically as having an anaphoric function, class terms have not. Since I have already shown that the general term of SGNs functions anaphorically in context, I investigate the patterns of compounds in this regard. Both compounds and SGNs may be anaphorically referenced by the standard inanimate pronouns. However, unlike the second/general element of compounds, the general nouns of SGNs may be used anaphorically in context. The following examples illustrate the distinct patterns of compounds in Sinhala.

(45)   eeva   narakwelaa
       3p.INAN   rotten
       'They're rotten.' (Chinese long beans)

(46)   *karal   narakwelaa
       3p.INAN   rotten
       'They're rotten.' (Chinese long beans)

(47)   eekə   lasənai
        3s.ɪɴᴀɴ   beautiful-ᴘʀᴇᴅ
        'It's beautiful.' (coral)

(48)   mal   lasənai
        3s.ɪɴᴀɴ   beautiful-ᴘʀᴇᴅ
        *'It's beautiful' (coral)
        'The flowers are beautiful.'

Examples 46-48 demonstrate the finding that the second element of these compounds may not be used anaphorically to refer to the entity denoted by the compound.[3] Example 48 further illustrates that while the sentence may be grammatical, the omission of an element of a compound may simply alter the meaning of the sentence. So, that even in context, sentence 48 would not make sense in reference to the term *muhudəmal* 'coral' as it would mean 'The flowers are beautiful' not 'It's beautiful' (coral).

    The evidence presented suggests that SGNs do not operate as compounds in Sinhala even though they function as a single grammatical word. The patterns of obligatoriness and anaphoric reference are clearly different in the examples found. While the general term of SGNs is optional or dispreferred in non-individuating contexts and optional when morphologically unmarked, both elements of compounds are required in the same contexts. Further, while the general term of SGNs may be used anaphorically in individuating contexts, neither element of the compound may be employed in the same fashion. Since class terms are by definition compounds, these patterns clearly distinguish SGNs from class terms.

    Mᴇᴀsᴜʀᴇ Tᴇʀᴍs ᴠs. Cʟᴀssɪꜰɪᴇʀs. Measure terms and classifiers, particularly numeral classifiers, can be difficult to distinguish from each other. As the precursors of numeral classifiers, measure terms perform similar functions and often occur in the same syntactic position. One key difference is that measure terms typically occur with mass nouns. Additionally, Greenberg (1972:9) describes the case for Khmer in which classifiers are optional as a general rule except in instances in which the classifier is functioning as a measure term with mass nouns. Furthermore, as mentioned in section 2, numeral classifiers have been identified as having an individuating function. Both of these characteristics are relevant to the patterns of SGNs in Sinhala.

    Most SGNs collected may not occur with mass nouns. However, two (*kææli, peti*) SGNs have been found to occur with mass nouns. This suggests that these two terms are best categorized as measure terms. However, characteristics of at least a couple of the examples of *kææli* displayed characteristics more suggestive of numeral classifiers--that is, it was used to individuate items from a set. The following examples illustrate their patterns as measure terms and as numeral classifiers.

(49)   darə   kææli   dahayak
        firewood   piece.ᴘʟ   ten.ɪɴᴅ
        'ten pieces of firewood'
(50)   maalu   peti   dahayak
        fish   flat.thing.ᴘʟ   ten.ɪɴᴅ

---

[3] However, it must be noted that very few noun-noun compounds were found and their patterns more generally in this regard were not found in the existing literature.

'ten fillets of fish'

(51)   bætəri   (kææli)   dahayak
         battery   piece.PL   ten.IND
         'ten batteries'

In examples 49 and 50, *kææli* and *peti* function to provide the unit by which the mass nouns *darə* and *maalu* may be quantified. It should be further noted that *kææli* and *peti* are not optional in these examples. In example 50, however, *kææli* is optional and functions to individuate a number of batteries from the set that batteries usually come in. Only a very small number of items were found with the term *peti* all of which pattern more like example 50 than 51. Additionally, the term *palu* has been found to occur with only a couple of items that are best described as items whose parts constitute a countable whole (oranges, garlic). In these cases, *palu* also patterns like examples 49 and 50, above, and therefore, it is best characterized tentatively as a measure term.[4] The issue of individuation is revisited in the following section as a characteristic which aids in locating the position of Sinhala nominal classification within the lexico-grammatical system.

4.2. DISTINGUISHING BETWEEN SUBTYPES OF LEXICO-GRAMMATICAL SYSTEMS IN SINHALA. As previously discussed the two primary differences between noun classifiers and numeral classifiers is their location within the noun phrase and individuation. While noun classifiers are found next to the noun in noun phrases and typically do not function to individuate the referent, numeral classifiers are typically found in numeral or quantifying phrases and are used in individuating contexts. While the distinction between the classifier's location in the noun phrase sounds clearly distinguishable, it is not so clear cut. In fact, as Greenberg points out, "in many languages the classifiers are not compulsory even for the restricted set of nouns that have them" (1972:6). An, in fact this is the case for SGNs in Sinhala. While SGNs are sometimes preferred in numeral phrases, they are typically not mandatory unless otherwise morphologically marked. Therefore, at first glance, they may appear to behave more like noun classifiers, however, their semantic properties are more suggestive of numeral classifiers. That is, they are used for individuation and for the remaining six terms (*geḍi, karal, kæṭə, æṭə, mal, alə*) the semantic relationships between elements are not clearly taxonomic, but rather the general terms carry some information about the shape or form of the specific nouns they accompany. Furthermore, their semantic consistency comes more from the properties of shape or form than their taxonomic relationship. The emergence of shape as a device for categorization is a typical semantic feature of numeral classifiers. However, in the case of Sinhala, the it is clear that it is not *only* shape that categorization relies upon. In this way, the general nouns of SGNs do not emerge as semantically prototypical numeral classifiers. However, along with the individuating function these terms emerge as more numeral than noun classifier-like.

5. CONCLUSION. The patterns described for SGNs in Sinhala suggest the presence of both lexical and lexico-grammatical systems of nominal classification. Three of the nine general terms of SGNs (*kææli* 'pieces', *peti* 'flat.things', and *palu* 'sections') investigated in this paper displayed characteristics more typical of measure terms (quantification in non-individuated contexts, use with mass nouns) while also showing signs of lexico-grammatical systems (individuation, denotation of shape). Of the remaining six classifying terms two (*æṭə* 'seed', *mal* 'flowers') are

---

[4]More evidence on the distribution of this particular term is needed.

more characteristic of class terms semantically since in most cases the semantic relationship between the elements is taxonomic; however, they displayed morphosyntactic patterns more characteristic of classifiers than of compounds based on language internal patterns of endocentric nominal compounding (obligatoriness of elements, anaphoric reference). Two of the remaining four classifying terms (*karal* 'pod-like.things', *alə* 'root.vegetables') also proved to pattern more like classifiers based on obligatoriness of elements and anaphoric reference. Additionally, these two terms exhibit semantic properties typical of classifiers (denotation of shape/physical properties, individuation). The remaining two classifying terms (*geḍi* 'small.hard.fruit-like.things', *kæṭə* 'block-like.things') display properties characteristic of classifiers, both semantically (denotation of shape, individuation) and morphosyntactically (anaphoric reference in individuated contexts, obligatoriness of elements). These classificatory terms in Sinhala may best be described as residing synchronically on the continuum of noun classification systems between class and measure terms and classifiers. The following figure attempts to demonstrate how we may locate Sinhala's SGNs among the systems of nominal classification.
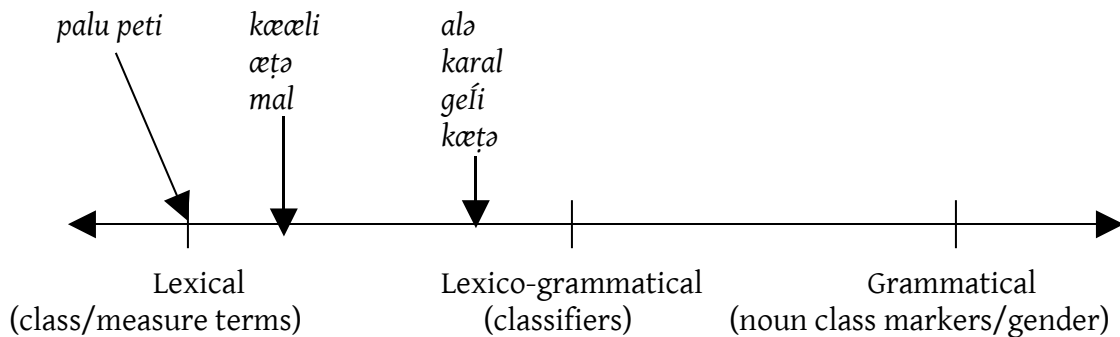
FIGURE 2. Sinhala's system of nominal classification

Although, several of Sinhala's SGNs pattern like lexico-grammatical systems, they are not the best exemplars of classifiers crosslinguistically. Furthermore, the semantic and morphosyntactic evidence suggests that the distinction between classifier subtypes too, may not be so clear cut.

As we have seen even among those whose characteristics are most suggestive of classifiers, the patterns are not prototypical of noun or numeral classifiers. The patterns described here suggest that those terms most like classifiers conform closest to the semantic properties of numeral classifiers, while less representative of numeral classifiers morphosyntactically. The following figure attempts to clarify the classification of these terms in Sinhala.
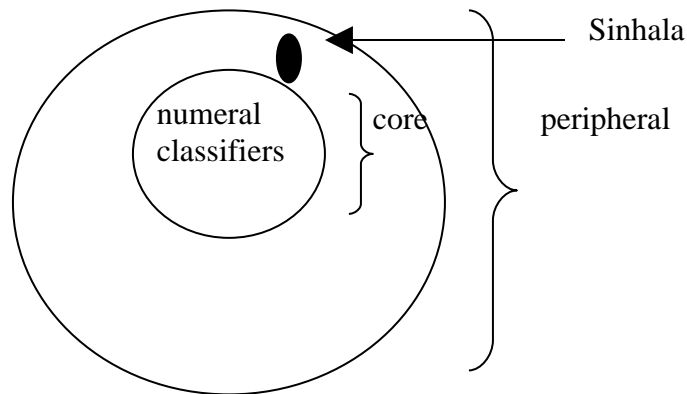
FIGURE 3. Classifier system in Sinhala

This figure attempts to illustrate the location of some SGNs in Sinhala as peripheral members of the category numeral classifiers. Here the inner circle represents the class of numeral classifiers that display the most prototypical characteristics of numeral classifiers. While the outer circle, within which I have placed Sinhala, represents the class of classifiers that do not possess clearly core characteristics or do not pattern systematically in the way those in the core class do, throughout the language.

This paper contributes to the growing body of literature on spoken Sinhala by exploring a small and until now underdescribed aspect of the grammatical system of Sinhala. It further contributes to typological work on Indo-Aryan languages by providing evidence for another system of nominal classification not yet described. There is much left to explore in the semantic and morphosyntactic patterns of nominal classification in Sinhala. This research would be much enhanced by an investigation in to naturally occurring discourse patterns among speakers of Sinhala in Sri Lanka.

REFERENCES

AIKHENVALD, ALEXANDRA Y. 2003. *Classifiers: A typology of noun categorization devices*, ch. 3, 81-97. Oxford: Oxford University Press.

ALLAN, K. 1977. Classifiers. *Language* 53.284-310.

DELANCEY, SCOTT. 1986. Toward a History of Tai Classifiers Systems. *Typological Studies in Language: Noun Classes and Categorization*, vol. 7, ed. by Colette Craig. 437-452. Amsterdam: John Benjamins Publishing Company.

DIXON, R.M.W. 1986. Noun Classes and Noun Classification in Typological Perspective. *Typological Studies in Language: Noun Classes and Categorization*, vol. 7, ed. by Colette Craig. 105-112. Amsterdam: John Benjamins Publishing Company.

DIXON, R.M.W. and ALEXANDRA Y. AIKHENVALD. 2002. Word: A typological framework. *Word: A cross-linguistic typology*, ed. by R. M. W. Dixon and Alexandra Y. Aikhenvald, 1-41. Cambridge: Cambridge University Press.

EMENEAU, M. B. 1956. India as a Linguistic Area. *Language* 32.1.3-16.

GAIR, JAMES W. and JOHN C. PAOLILLO. 1997. *Sinhala.* München: Lincom Europa.

GREENBERG, JOSEPH H. 1972. Numeral Classifiers and Substantival Number: Problems in the Genesis of a Linguistic Type. *Working Papers on Language Universals* 9.1-39.

GRINEVALD, COLETTE, and FRANK SEIFART. 2004. Noun Classes in African and Amazonian languages: Towards a comparison. *Linguistic Typology* 8.243-285.

University of California, Santa Barbara
Department of Linguistics
3607 South Hall
Santa Barbara, CA 93106
mhenderson@umail.ucsb.edu

THE LEXICAL CATEGORY *AUXILIARY* IN SINHALA[1]

MARTIN HILPERT

Rice University

1. INTRODUCTION. This paper discusses whether there are elements in colloquial Sinhala that can be appropriately labeled AUXILIARY verbs, and what evidence there is to motivate such a label. While auxiliaries are thought of as a nearly universal lexical category (Steele 1978), the term is not mentioned in standard works on Sinhala (Gair & Paolillo 1997, Gair 1998), which warrants a closer examination of the existing data.

Auxiliaries have been studied under various aspects in a multitude of theoretical frameworks (Heine 1993). Accordingly, there is more than one definition of the term *auxiliary*. The present study adopts the framework of GRAMMATICIZATION THEORY (Heine & Traugott 1991, Hopper & Traugott 1993) and a definition of auxiliary that presupposes some assumptions of that theory. While it will be argued that grammaticization theory provides a fertile ground for an analysis of the Sinhala data, it needs to be pointed out that by the same token, the theory is subject to modification or even falsification in the event of anomalies (Kuhn 1970) in the observed data.

With Heine (1993:70), I take an auxiliary to be 'a linguistic item covering some range of uses along the Verb-to-TAM chain'. To explicate this definition, a frequent, cross-linguistically attested development is that main verbs over time develop into grammatical markers. This development happens gradually, so that main verbs shed some of their lexical meaning and acquire grammatical meaning concerning tense, modality, or aspect, and thus change into auxiliaries. Auxiliaries may grammaticize even further, reduce in form, and ultimately change into affixes. Elements occupying the middle ground of the continuum from main verb to affix can be called auxiliaries. This view acknowledges the fact that it is impossible to cross-linguistically define auxiliaries in terms of necessary and sufficient criteria. It also does not make the claim that auxiliaries are a universal cross-linguistic category. Rather, it makes room for empirical data to decide whether there are elements that exist somewhere along the Verb-to-TAM chain, what their lexical sources are, and how far they are along in the process of grammaticization.

The present study uses functional and formal criteria to heuristically arrive at a set of possible candidates for auxiliary status, which are then analyzed in terms of syntactic behavior, morphology, and grammatical function. The database for this study consists of 15 texts that were collected from two consultants in 2004 and 2005, class notes from that time, and additional elicitation data.

A hallmark of auxiliaries is that they take verbal complements that are not fully finite (Bolinger 1980:297). As there is an infinitive verb form in Sinhala, this means that elements co-occurring with an infinitive complement may qualify as auxiliaries. Finiteness in Sinhala is a

matter of degree. For this reason, elements which are not maximally finite verbal complements should also be considered. Table 1 gives an overview of Sinhala elements that take non-finite verbal or clausal complements, and thus form the object of investigation for the present study.

Element	English gloss	Function	Complement types
yannə	go	FUTURE	INF
dennə	give	PERMISSIVE	INF
patan gannə	start take	INCEPTIVE	INF
wennə	become	EQUATIVE	CLAUSE
næhæ	not	NEGATION	FOCUS PHRASE, AUX
bæhæ	impossibly	EPISTEMIC	INF, AUX
æti	definitely	EPISTEMIC	INF, VERB PHRASE, AUX
puluwaŋ	possibly	EPISTEMIC	INF, AUX, CLAUSE
kæməti	like	AFFECTION	INF, CLAUSE
kanəgatu	sorry	REGRET	INF
bayə	afraid	FEAR	INF, CLAUSE
oone	need	DESIRE	INF

TABLE 1. Sinhala elements taking infinitive or not fully finite verbal complements

Table 1 presents a provisional classification into VERBAL ELEMENTS (yannə, dennə, etc.), EPISTEMIC ELEMENTS (bæhæ, næhæ, etc.), and STANCE ELEMENTS (kæməti, etc.). The first category is motivated by morphological form while the two others are based on semantics. All elements will be analyzed in terms of distribution across different construction types, difference in morphology from regular main verbs, the semantics of their lexical sources, and their grammatical function. All of these are indicators that either allow a placement of an element on the Verb-to-TAM chain, and hence are suggestive of auxiliary status, or characterize the element as belonging to a different category. All considered evidence is synchronic. Since the Verb-to-TAM chain is an inherently diachronic notion, the evidence is not explanatory, but merely suggestive. The aim of this study is to generate reasonable hypotheses that are empirically testable against diachronic data.

Section 2 of this paper elaborates on the notion of auxiliation and gives the theoretical background. Section 3 discusses the evidence and proposes a classification of the elements listed in Table 1. Section 4 concludes and puts auxiliation in colloquial Sinhala into typological perspective.

2. AUXILIATION AND LEXICAL SOURCES OF AUXILIARIES. This paper treats auxiliaries as grammatical markers that develop out of lexical verbs. In accordance with a view of grammar as emergent and continually changing through usage (Hopper 1987, Barlow & Kemmer 2000, Bybee & Hopper 2001), auxiliaries are not assumed to form a uniform category. Instead, they are defined operationally as participating in the process of AUXILIATION (Benveniste 1968), which is schematized below as the development of 1a into 1b (adapted from Kuteva 2001:1):

(1)   a. verb   - argument
      b. grammatical marker - main verb

In auxiliation, argument-taking verbs undergo a semantic change from their lexical meaning towards more grammatical meaning. Along with the semantic change, the verb changes syntactically from taking arguments to taking various kinds of complements to a preference for non-finite verbal complements. At the same time, the verb may be subject to morphological and phonological reduction.

While 1b can be seen as the endpoint of auxiliation, auxiliaries tend to develop further into affixes, which motivates Heine's (1993) concept of the Verb-to-TAM chain. While grammaticization along the Verb-to-TAM chain may proceed in different ways, Heine (1993:58ff) suggests the following stages as an approximation.

Stage A - The verb has its full lexical meaning and takes an argument which typically refers to a concrete object, as in *I expect a visitor.*

Stage B - The verb has its full lexical meaning, but it takes a complement which typically refers to a dynamic situation, as in *I expect getting a tax refund.* The complement may have different forms, such as an infinitive, a gerund, a participle, or a full clause.

Stage C - At this stage the selection restrictions of the lexical meaning loosen and the verb acquires some grammatical meaning. The verb may take an etymologically identical complement, as in *I am going to go.* Stage C items typically relate to the duration, speed, or boundary characteristics of the denoted event. Even when these items take a nominal argument, these are likely to refer to events or activities. Another difference with respect to stage B is that stage C items tend to form a single semantic unit with their complements, as in *He stopped smoking.*

Stage D - This stage includes the loss of morphological variety. Items lose their ability to form imperatives, nominalizations, or the passive. Thus, stage D items show formal signs of decategorialization, they do not behave like lexical verbs anymore. Stage D items also take fewer types of complements than stage C items. For example, English *try* takes the infinitive and the gerund, English *want* only takes infinitive complements.

Stage E - At this stage syntactic indicators of decategorization emerge. Items lose their ability to be separately negated, they cannot be separated from their complements for topicalization. English auxiliaries like *can*, *may* and *must* are stage E items. Items in this stage may start to cliticize to the verbal complement and lose in phonological substance. Semantically, stage E items code only grammatical meaning.

Stage F - This stage marks the transition from a clitic to an affix. The element can still bear secondary stress.

Stage G - The affix reduces phonologically to a monosyllabic affix without stress.

Section 3 presents an analysis of the elements from Table 1 according to the criteria in Heine's stage model. All elements in Table 1 are phonological words, which means that stages F and G will not be discussed any further.

Since the process of auxiliation frequently goes along with polysemization, some items may display behaviors associated with different stages in different uses. For example, consider the English sentences *I used a toothpick* and *I used to collect toothpicks.* The second sentence shows that the lexical verb *use* has grammaticized into an auxiliary that codes habituality. However, *use* still persists as a full lexical verb, as can be seen in the first sentence. The semantic and formal differences between *use* and *use to* motivate a synchronic treatment of these as two separate items, but the development to this state of affairs has been gradual. Hence, individual items may cover a certain range on Heine's stage model.

A cross-linguistic observation is that some types of lexical verbs seem particularly amenable to development into auxiliaries. General movement verbs, posture verbs, and verbs of possession are attested as grammatical markers in many of the world's languages. It needs to be pointed out that these verbs do not only grammaticize into auxiliaries. Movement and posture verbs are also productive sources of SERIAL VERBS, which are distinguished from auxiliaries proper.

These cross-linguistically common grammaticization clines do of course not preclude more idiosyncratic developments, such as for example Korean *pelita* 'throw away' changing into a perfect marker (Bybee & Dahl 1989:58). Although the exact developments in grammaticization are not predictable, certain developments occur regularly, even across different language families. Verbs of location show a tendency to develop into aspect markers while movement verbs frequently grammaticize into tense markers. Heine (1993:47) identifies a number of common lexical sources of auxiliaries along with the grammatical functions that these typically evolve into.

Source	Grammatical functions
LOCATION	PROGRESSIVE, INGRESSIVE, CONTINUOUS
MOTION	INGRESSIVE, FUTURE, PERFECT, PAST
ACTION	PROGRESSIVE, CONTINUOUS, INGRESSIVE, COMPLETIVE, PERFECT
VOLITION	INGRESSIVE, FUTURE
CHANGE OF STATE	INGRESSIVE, FUTURE
EQUATION	RESULTATIVE, PROGRESSIVE, PERFECT, FUTURE
ACCOMPANIMENT	PROGRESSIVE
POSSESSION	RESULTATIVE, PERFECT, FUTURE
MANNER	PROGRESSIVE

TABLE 2. Lexical sources of auxiliaries with associated grammatical functions
(= Table 2.2, Heine [1993:47])

Cross-linguistic tendencies as those in Table 2 should not be taken as explanatory evidence for or against an observed change in a given language. However, they can serve as heuristics in elicitation, as the above lexical sources are good starting points to look for grammaticizing elements. Conversely, comparing auxiliation in a given language against the backdrop of cross-linguistically common tendencies may illuminate interesting grammatical peculiarities of that language.

3. AUXILIARIES AND RELATED FORMS IN COLLOQUIAL SINHALA. This section discusses the elements from Table 1 in terms of their syntactic distribution, their morphological similarity to regular main verbs, their grammatical function, and, where possible, the semantics of their lexical sources. The section is organized in accordance with the provisional classification made in Table 1 into verbal, epistemic, and stance elements.

The schema of auxiliation in 1 is not meant to specify the order of elements; auxiliaries may emerge at either side of the verbal complement, depending on word order in the respective language. Basic constituent order in Sinhala is SOV. Sinhala adheres to all of the Greenbergian word order correlates (Greenberg 1963) of SOV languages; constituents strongly tend to be right-headed. The basic constituent order in a transitive sentence is exemplified in

2a. In complex verb phrases, the fully finite element occurs at the right edge of the phrase, as in 2b. Auxiliating elements can thus be expected to be found to the right of a non-finite verb.

(2)   a.   laməya   epel gediyak   kæwwa
            child   apple CL.IND   eat-PST
            'The child ate an apple.'
      b.   mamə   kannə   yanəwa
            1SG   eat-INF   go-NPST
            'I will eat.'

3.1. VERBAL ELEMENTS. The elements discussed in this section can be used as main verbs in colloquial Sinhala, as shown in (3a-d). The sections below discuss uses of the elements that diverge in both meaning and form from these examples. In contrast to the usages shown in (3a-d), the grammaticized counterparts of the respective verbs have evolved into markers of tense, aspect, and modality.

(3)   a.   ohu   gedərə   yanəwa
            3SG   home   go-NPST
            'He goes home.'
      b.   ohu   matə   epel gediyak   dunna
            3SG   1SG-DAT   apple CL-IND   give-PST
            'He gave me an apple.'
      c.   ohu   pot   gatta
            3SG   book-PL   take-PST
            'He took the books.'
      d.   eekə   ratu   wenəwa
            it   red   become-NPST
            'It becomes red.'

The element *yannə* 'go' shows a number of signs of auxiliation. Much as with the English items *use* and *used to*, it is justified to distinguish between usage of *yannə* as a main verb and as an auxiliary. The grammatical meaning associated with the auxiliary is FUTURE TENSE. The grammaticization of a movement verb like *yannə* into a future marker is cross-linguistically very common. The construction is not mentioned in Gair & Paolillo (1997), but Garusinghe (1962:64) points out that future tense in spoken Sinhala is expressed through a periphrastic construction with *yannə*. The semantic change has loosened selection restrictions in the auxiliary. While the main verb is restricted to animate subjects, the auxiliary also occurs with inanimate subjects. The auxiliary takes only non-finite verbal complements. By the criteria outlined in section 2, *yannə* is a stage D auxiliary.

(4)   a.   * geə   yanəwa
              house go-NPST
              'The house goes.'

b.   geə   kaḍaŋ   wæṭennə   yanəwa
house   break   fall-INF   FUT-NPST
'The house will collapse.'

The auxiliary cannot form the imperative. Imperatives are generally understood to refer to some future action, so the English gloss of 4d is not grammatical either.

(4)   c.   gedərə   yannə
home   go-IMP
'Go home!'
d.   *gedərə   yannə   yannə
home   go-INF   FUT-IMP
'Will go home!'

There are compound verb constructions in Sinhala that are formed from two verbs in conjunction. The first of the verbs receives the CONVERB suffix *–la*, the second is finite. While this is a very productive process that does not necessarily alter the semantics of the individual elements, some collocates may develop a new semantics. To illustrate this, the verb *pænnə* 'jump' in conjunction with *yannə* has acquired the meaning 'escape':

(4)   e.   gemba   botələyen   pænəla   giya
frog   bottle-LOC   jump-CONV   go-PST
'The frog escaped from the bottle.'

Similar compound verb constructions can be observed with *gannə* 'take', they are discussed later in connection with that element.

The element *dennə* 'give' can also be identified as a fully grammaticized auxiliary. A distinction between usage of *dennə* as a main verb and as an auxiliary is useful, as the two elements have distinct meanings. The grammatical meaning associated with the auxiliary is PERMISSIVE, which as a grammatical function falls into the domain of deontic modality. The development a verb of giving into a permissive marker has been described by Newman (1996:236), who discusses the metaphorical motivation for the semantic extension. In an act of giving, a recipient gains control over a transferred object. In giving someone permission, the permittee gains control over an action. Permissives that derive from verbs of giving are found also in Russian, Finnish, and Mandarin (Newman 1996:189). An example is shown in 5a.

(5)   a.   ohu   maṭə   epel gediyak   kannə   dunna
3SG   1SG-DAT   apple CL-IND   eat-INF   PRM-PST
'He let me eat an apple.'
b.   ohu   maṭə   epel gediyak   kannə   idə   danəwa
3SG   1SG-DAT   apple CL-IND   eat-INF   room   give-NPST
'He lets me eat an apple.'

Example 5b shows a possible source construction for 5a. The collocation *idə dennə* 'give permission', literally 'give room', may have been reduced to just the verb, making it

structurally an auxiliary. In the absence of diachronic evidence, it is hard to determine what exactly has led to the structure that is found in Sinhala today.

Due to the permissive semantics, the auxiliary retains the ability to form the imperative and remains restricted to animate subjects. The auxiliary takes only non-finite verbal complements, which makes it a stage-D element.

The verb *gannə* 'take' differs from the two previously discussed elements, because it does not take infinitival complements by itself. It only occurs as the head of a complex predicate, which in turn may have an infinitival complement. Consider (6a-d).

(6)   a.   miniha   laməyaṭə   balaa   ganəwa
           man   child-DAT   look   take-NPST
           'The man looks after the child.'
      b.   miniha   epel gediyak   labaa   ganəwa
           man   apple CL-IND   happen   take-NPST
           'The man obtains an apple.'
      c.   miniha   horawə   allaa   gatta
           man   robber-ACC   touch   take-PST
           'The man caught the robber.'
      d.   miniha   duwannə   patan   ganəwa
           man   run-INF   start   take-NPST
           'The man starts running.'

In 6a to 6c, *gannə* heads a light verb construction which includes a verb form ending in a long –*a.* In 6a, the collocation *balaa gannə* 'look take' has acquired the meaning 'to look after someone'. Similarly in 6b, *labaa gannə* 'happen take' means 'obtain', and in 6c, *allaa gannə* 'touch take' means 'catch'. In 6d, a similar construction functions as a complex auxiliary. The collocation *patan gannə* 'start take' has fused into an auxiliary meaning 'begin'. The word *patan* never occurs outside this construction, it is unclear what part of speech it derives from, or what its own lexical meaning would be.

The 55ehaviour of *gannə* differs from the grammaticization paths that have been taken by *yannə* 'go' and *dennə* 'give'. While the latter take non-finite complements of any kind, *gannə* primarily takes specific finite complements that form collocations and develop a constructional meaning of their own. The case of *patan gannə* 'start take' is the only one of these constructions that takes a non-finite complement and thus qualifies as an auxiliary with INCEPTIVE grammatical function, which puts it into the domain of aspectual markers. As shown in 6e and 6f, *patan gannə* allows the formation of the imperative, and it also takes nominal arguments. Accordingly, it can be classified as a stage C item in Heine's taxonomy.

(6)   e.   duwannə   patan   gannə
           run-INF   start   take-IMP
           'Start running!'
      f.   miniha   randuwak   patan   ganəwa
           man   fight-IND   start   take-NPST
           'The man starts a fight.'

The form *patan gannə* illustrates that grammaticization does not only operate on single lexical items, but that entire phrases can develop into grammatical constructions (Kuteva 2001:1). Cross-linguistically, lexical verbs meaning 'take' are a particularly productive source for grammaticization. Common grammatical domains deriving from it include causatives, as well as markers of future, possession, and completion (Heine & Kuteva 2002:286).

In its non-lexical uses *wennə* 'become' functions as a tense-carrying verbal element that is comparable to an EQUATIVE copula.

Attributive sentences and predicate nominals in the present tense do not require a verbal element. However, when the attribute is meant to hold in either the future or the past, a finite form of *wennə* is required, as shown in 9b and 9c.

(9)   a.   ohu   horek
             3SG   robber-IND
             'He is a robber.'

        b.   laməya bohomə santosə una
             child   very   happy   EQ-PST
             'The child was very happy.'

        c.   ohu   horek   wey
             3SG   robber-IND   EQ-FUT
             'He will be a robber.'

A form of *wennə* is found with constructions that involve one or more of the verbal, epistemic and stance elements mentioned in Table 1. The order of these elements is regular, as the form of *wennə* occurs after stance elements like *oone* 'want/need' and *bayə* 'be afraid', but before epistemic elements like *æti* 'definitely/probably'. This syntactic distribution motivates the distinction between epistemic and stance elements that was made on semantic grounds in the introductory section. The form of *wennə* is inflected only if it occurs as the last element in the clause.

(9)   d.   maṭə   epel gediyak   kannə   oone wey
             1SG-DAT apple CL-IND   eat-INF   need   EQ-FUT
             'I will need to eat an apple.'

        e.   eyaa   gedərə   yannə bayə   wey
             2SG   home   go-INF afraid EQ-FUT
             'You will be afraid to go home.'

        f.   ohu   horek   wennə   æti
             3SG   robber-IND   EQ-INF   MUST (STRONG EPISTEMIC MODALITY)
             'He must be a robber.'

Since *wennə* does not take complements that are clearly non-finite in nature, it cannot be appropriately called an auxiliary. The complements it takes are predicative structures that are syntactically complete clauses. The reason it was included in the initial set of potential candidates was that in examples like 9c and 9d, it appears that the form of *wennə* is the only inflected element. While that is indeed the case, its complement structures are not non-finite, but simply do not require a finite element in the present tense.

The development of a verb denoting change of state into an equative copula is cross-linguistically common. Hengeveld (1992:253) discusses data from Ngalakan and Turkish. Similarly to Sinhala, the occurrence of the copula seems to be confined to specific verbal tenses in these languages.

3.2. EPISTEMIC ELEMENTS. The elements discussed in this section are used to indicate the likelihood, probability, or improbability of some event. As such, they fall into the grammatical domain of epistemic modality.

Gair and Paolillo use the label QUASI-VERBS (1997:26) to group epistemic elements and stance elements together. Their evidence for classifying these as verb-like is that they occur as predicators of clauses, and share a number of inflectional properties with lexical verbs. The rationale for grouping them together is that they share the negative characteristic of being not entirely verb-like, but distributed in very similar ways. While I am in agreement with all of these observations, I will not adopt the classification, but keep the two classes of elements apart. Section 3.3 below summarizes the syntactic, morphological, and semantic evidence motivating this decision.

Neither epistemic nor stance elements are readily accommodated in Heine's stage model of auxiliation, because these elements lack the characteristic twin role of auxiliaries, which tend to have lexical verb counterparts. For example, for the element *puluwaŋ* 'possibly' there is no corresponding lexical verb. Another hallmark of auxiliaries is polysemy. While all elements discussed in section 3.1 are polysemous to some extent, this is not the case for *næhæ* 'not' and *bæhæ* 'impossibly'.

It is a core assumption of grammaticization theory that all grammatical elements develop out of some lexical source. However, when an element has become sufficiently decategorialized as to be fully opaque, it is no longer possible to determine the lexical source in the absence of historical evidence. Accordingly, the following sections do not attempt to resolve the history of these elements, but instead discuss the synchronic evidence that would motivate a classification of these elements either as auxiliaries, or as some other category. This evidence includes syntactic distribution, morphology, and the interplay of these elements with the auxiliaries discussed in section 3.1.

For grammatical markers of negation, likely source candidates are lexical verbs meaning 'lack' or 'leave' (Heine & Kuteva 2002:333). Neither of these seems to apply in Sinhala, which leaves us with synchronic evidence. The element *næhæ* 'not' marks NEGATION in a range of different constructions, such as existential, possessive, transitive, and intransitive clauses. Existential and possessive clauses are closely related, as possessives are merely existentials with a dative possessor. Compare 11a and 11b.

(11)   a.   lamǝyek   næhæ
             child-IND   NEG
             'There is no child.'
       b.   eyaaṭǝ   lamǝyek   næhæ
             she- DAT   child-IND   NEG
             'She has no child.' (lit. There is no child to her.)

Evidence for a verb-like status of *næhæ* is that different construction types involve a set of morphologically similar negation markers that form a paradigm. Similarly, the Sinhala verb

inflects for different syntactic environments, such as co-temporal clauses, concessive clauses, and causal clauses. Besides *næhæ* there are two more forms. Predicate nominals are negated by *newey*, as shown in 11c, negation of causal sentences with *hinda* 'because' involves *næti*, as shown in 11d. While Gair and Paolillo (1997) view these forms as one paradigm, one could make the case that they are in fact three separate particles.

(11) c. | mamə | šišəyek | newey |
|---|---|---|
| 1SG | student-IND | NEG |

'I am not a student.'

d. | balənə-koṭə | miniha | laŋgə | hændunum | paṭə | tibune |
|---|---|---|---|---|---|
| look-SIM | man | close | identity | card | EX.INAN-PST.FOC |

næṭi	hinda	bayəwela	duwəla
NEG	because	afraid-CONV	run-CONV

'When we inspected the man closely, it turned out that because he had no ID on him he got scared and started running.'

Evidence for regarding *næhæ* as an auxiliary stems from the fact that it takes not fully finite complements. In 11e the complement is marked with the continuative *–genə*, while in 11f it takes the emphatic *–e* suffix.

(11) e. | eyaa | loguwak | andəgenə | næhæ |
|---|---|---|---|
| she | coat-IND | wear-CONT | NEG |

'She is not wearing a coat.'

f. | laməya | wæḍe | kəranne | næhæ |
|---|---|---|---|
| child | task | do-FOC | NEG |

'The child doesn't do the work.'

Gair and Paolillo (1997:27) state that the emphatic verb form is the default case for negation with *næhæ*. We see this point corroborated in similar sentences with auxiliaries between the lexical verb and the negating element, where the auxiliary bears the emphatic suffix.

(11) g. | minihek | hiṭiye | næhæ |
|---|---|---|
| man-IND | EX.ANIM-PST.FOC | NEG |

'There was no man.'

h. | mamə | laməyaṭə | epel gediyak | kannə | denne | næhæ |
|---|---|---|---|---|---|
| 1SG | child-DAT | apple CL-IND | eat-INF | PRM-FOC | NEG |

'I won't let the child eat an apple.'

Only in conjunction with the element *næhæ* are the above sentences finite. This, and the fact that different forms similar to *næhæ* appear in different construction types, makes it verb-like, but that also is where the similarity ends. While the term quasi-verb may thus be appropriate, there are four reasons not to view *næhæ* as an auxiliary.

First, from the data it appears that elements that can be clearly identified as auxiliaries do not occur next to each other. The tendency to avoid auxiliary stacking is cross-linguistically common (Heine 1993:23), although there are numerous counterexamples. Second, *næhæ* can be found in cliticized form, as shown in examples 11i and 11j.

(11) i.   maṭə   epel gediyak   oonæhæ
          1SG-DAT   apple CL-IND   need-NEG
          'I don't need an apple.'

      j.   maṭə   epel gediyak   oone   unnæhæ
          1SG-DAT   apple CL-IND   need   become-PST-NEG
          'I didn't need an apple.'

Third, it is problematic to view *næhæ, næti*, and *newey* as finite auxiliary forms, because there is no corresponding non-finite form. The fourth problem is that the negation markers are monosemous, lacking the characteristic polysemy of auxiliaries.

   This evidence does not preclude that *næhæ* at some point actually was an auxiliary, and it certainly does not say anything about its lexical origins, but it shows that it is by the adopted definition not an auxiliary in present-day colloquial Sinhala.

   The element *bæhæ* 'impossibly' carries meaning that is expressed in English through epistemic uses of the modal *can*, as in *That can't be right*. A difference is though that *bæhæ* does not code deontic modality, it is confined to epistemic meaning. The element is morphologically almost identical to *næhæ*, but it differs in its morphosyntactic behavior. It does not cliticize, and it is restricted to infinitive complements, which means that in predicate nominals and attributive clauses the infinitive form *wennə* is required. In this use, the equative copula *wennə* does not add to the meaning of the sentence. Occasionally it carries the implicature of futurity, though. Like *næhæ*, *bæhæ* has an alternate form. The alternative form *bæri* occurs in subordinate clauses, as shown in 12d.

(12) a.   ohu   horek   wennə   bæhæ
          3SG   robber-IND   EQ-INF   IMPOSS
          'He can't be a robber.'

      b.   ohuṭə   tibaha   wennə   bæhæ
          3SG-DAT   thirsty   EQ-INF   IMPOSS
          'He can't be thirsty.'

      c.   ohu   wæde   kərannə   bæhæ
          3SG   task   do-INF   IMPOSS
          'He can't possibly do the work.'

      d.   ohu horek   wennə bæri hinda   ohu hire   yanne   næhæ
          3SG robber-IND   EQ-INF IMPOSS because 3SG jail-LOC go-FOC   NEG
          'Since he can't be a robber, he will not go to jail.'

In summary, *bæhæ* shows some parallels with *næhæ* that warrant a classification into the same category. It is required in certain structures to yield a finite sentence, and it takes non-finite complements, but that is not enough evidence to call it an auxiliary.

   The element *æti* 'definitely / probably' is polysemous. It codes weak and strong epistemic meaning, the two of which are complementarily distributed across different construction types. The entry in Table 1 renders the meaning as 'definitely', but there are contexts where it means 'probably'. The strong epistemic meaning 'definitely' co-occurs with infinitive complements. As 13a shows, *æti* requires a copula in predicate nominals, just like *bæhæ*. Example 13b illustrates how *æti* can modify regular non-finite verb phrases. What sets *æti*

apart from *næhæ* and *bæhæ* is that it does not have an alternate form in subordinate clauses, as shown in 13c.

(13) a. ohu horek   wennə æti
     3SG robber-IND   EQ-INF MUST
     'He must be a robber.'

   b. ohu gedərə   yannə æti
     3SG home   go-INF MUST
     'He must be going home.'

   c. ohu horek   wennə æti hinda   ohu hire   yay
     3SG robber-IND   EQ-INF DEF because 3SG jail-LOC go-FUT
     'Since he must be a robber, he will go to jail.'

Where the complement of *æti* is a complete finite structure, it has the meaning 'probably'. Syntactically it is a sentence adverbial in these examples, rather than a quasi-verbal element. In 13d below, the main verb is fully finite, leaving *æti* with no grammatical function, but merely its semantic adverbial function.

(13) d. ohu wæde kərənəwa æti
     3SG task do-NPST   PROBABLY
     'He is probably doing the work.'

Gair and Paolillo (1997:36) point out another, lexical use of *æti*, which may possibly be the lexical source of the two grammatical uses discussed above. There are noun phrases such as *salli æti* 'enough money' in colloquial Sinhala, where *æti* means 'enough'. The grammaticization of an adjective meaning 'enough' into a marker of deontic and epistemic modality has been studied in Luo (Bavin 1995), who considers this grammaticization path an areal phenomenon pertaining to African languages. The case of Sinhala suggests that this cline may be more common than that. However, in order to argue for the similarity of these developments, we would need a crucial piece of evidence showing that *æti* at some point had deontic modal meaning. I do not see this evidence at present.

   Despite the fact that *æti* takes non-finite complements, a classification of it as an auxiliary cannot be sufficiently motivated. In comparison to *næhæ* and *bæhæ* it appears even less verb-like, since it does not have alternate forms, and co-occurs with fully finite examples such as 13d.

   There is evidence for viewing the element *puluwaŋ* 'possibly' as either an epistemic or a stance element, depending on what aspects of it are in focus. Gair and Paolillo (1997:26) render its meaning as English 'can', which is the deontic counterpart to its epistemic meaning 'possibly'. While the deontic meaning relates to a speaker's stance towards some state of affairs, the epistemic meaning relates to the likelihood of some event. In the data on which this study is based, the epistemic meaning dominates. It only takes infinitive complements. Like *æti*, it has the same form in subordinate clauses. There are morphological and syntactic criteria that correspond with the two meanings. Epistemic *puluwaŋ* is found to the right of auxiliaries, as shown in 14a and 14b. Stance *puluwaŋ* can occur to the left of auxiliaries, as shown in 14c.

(14) a. maṭə   epel gediyak oone   wennə   puluwaŋ
       1SG-DAT   apple CL-IND   need   EQ-INF   PROBABLY
       'I will probably need an apple.'

    b. ohuṭə   tibaha   wennə puluwaŋ hinda   mamə waturə
       3SG-DAT thirsty EQ-INF PROB   because 1SG   water
       geenəwa
       bring-NPST
    'Because he is probably thirsty, I bring some water.'

    c. maṭə   epel gediyak labaa   gannə   puluwaŋ una
       1SG-DAT   apple CL-IND   happen   take-INF   able   EQ-PST
       'I was able to obtain an apple.'

Another feature that sets *puluwaŋ* apart from the other epistemic elements is that it can occur pre-verbally, as shown in 14d. This is unusual, given that Sinhala has a strict preference for right-headed constructions. Example 14d also shows that stance *puluwaŋ* requires dative subjects. Epistemic elements, by contrast, are found with both nominative and dative subjects.

(14) d. aliyan-ṭə   jiip   ratə   puluwaŋ perəlannə
       elephant.PL-DAT jeep vehicle able   overturn-INF
       'Elephants can overturn jeeps.'

Overall, the evidence rules out a classification of *puluwaŋ* as an auxiliary, rather, it is a quasi-verb that seems to have grammaticized from an adjective. The grammaticization of epistemic markers from lexemes with the meaning 'ability' is very common (Bybee et al. 1994:187).

3.3. STANCE ELEMENTS. The elements discussed in this section are used to indicate a speaker's stance towards some event, such as for example appreciation, fear, or regret. These concepts are not generally recognized as grammatical, although stance and emotion lexemes do frequently give rise to more grammatical meanings. What warrants the discussion here are distributional similarities between stance elements and the auxiliaries discussed earlier. Grammaticization theory acknowledges that there is no strict division of 'grammar' and 'the lexicon'. To illustrate this, the English verb *try* is a less grammaticized auxiliary than for example *will*, but it shows some distributional similarities. While *will* takes only infinitive complements and regularly cliticizes, *try* takes infinitive and gerund complements. If we adopt Hopper's (1987) idea of emergent grammatical categories, we commit ourselves to the view that categories are in flux at all times. Such an open-ended view of grammar allows for degrees of auxiliarihood.

    The elements *kæməti* 'like' and *oone* 'want/need' are classified as quasi-verbs in Gair and Paolillo (1997:26), *bayə* '(be) afraid' and *kanəgatu* '(be) sorry' are discussed in Garusinghe (1962). All of these can take nominal arguments and non-finite verbal complements. However, there are also differences. In simple attributive sentences, *kæməti*, *bayə* and *kanəgatu* behave like regular adjectives, which leaves *oone* with a special status in this category. The latter is also the only polysemous element, it can mean either 'want' or 'need'. The sections below discuss each

element in detail and analyze the interplay of auxiliaries, epistemic elements, and stance elements.

The element *kæməti* 'like' can take nominal arguments, and clausal and infinitive complements. These are illustrated in 15(a-c). The subjects of sentences with *kæməti* are in the nominative. If *kæməti* is the last element in the sentence, it takes the ASSERTIVE marker -*i*, which also occurs on regular adjectives in attributive clauses.

(15) a. mamə satuntə kæmətii
     1SG   animal.PL like-ASS
     'I like animals.'

    b. ballat ekə sellam kərənə ekəṭə ṭamay laməya kæmətii
     dog   with game do- REL.PRES COMP ASS   child like-ASS
     'It is playing with the dog that the child likes.'

    c. mamə gedərə yannə kæmətii
     1SG   home go-INF like-ASS
     'I like to go home.'

It is a characteristic of stance elements that they precede auxiliaries and epistemic elements, as shown in 15d and 15e. Auxiliaries, in turn, precede epistemic elements, as shown in 15f. The correspondence of a three-fold semantic distinction to syntactic distribution is the main argument made in this paper for a distinction of auxiliaries proper, epistemic elements, and stance elements.

(15) d. mamə satuntə kæməti næhæ
     1SG   animal.PL like   NEG
     'I don't like animals.'

    e. redi   hodəpu   ekəṭə ṭamay mamə kæməti une
     clothes wash- PST.REL COMP ASS   1SG   like   EQ-PST.FOC
     'It was washing clothes that I liked.'

    f. mamə satuntə kæməti wennə yanne næhæ
     1SG   animal.PL like   EQ-INF FUT-FOC NEG
     'I will not like animals.'

Gair and Paolillo (1997:26) present evidence that *kæməti* inflects for syntactic context in the same way that lexical verbs and the epistemic elements *æti* and *næhæ* do. I have not been able to elicit these forms, but I consider it likely that the forms given in the first column of 15g are fused with a form of *wennə* that have not merged in my consultant's variety.

(15) g.

BASIC	*kæməti*	*kæməti*
CONDITIONAL	*kæməttot*	*kæməti unot*
CONCESSIVE	*kæməttat*	*kæməti unat*
FOCUS	*kæmətte*	*kæməti une*

In summary, *kæməti* does not behave very verb-like in the investigated variety of Sinhala. Its broad range of complements indicate a low degree of grammaticization; it cannot be classified as an auxiliary.

The element *kanəgatu* 'sorry' takes non-finite verb phrases as complements. It assigns dative case to its subjects. Also *kanəgatu* takes the assertive marker *-i* in sentence-final position.

(16) a.   maṭə   randuwə   gænə   kanəgatui
           1SG-DAT   fight   about   sorry-ASS
           'I'm sorry about the fight.'

      b.   maṭə   yannə   kanəgatui
           1SG-DAT   go-INF   sorry-ASS
           'I'm sorry to leave.'

Without the assertive marker the whole structure would not be finite, and hence ungrammatical. The marker is absent in examples with auxiliaries and epistemic elements, as shown in 16c and 16d. In summary, *kanəgatu* needs to be regarded as a weakly grammaticized adjective, not an auxiliary.

(16) c.   maṭə   randuwə   gænə   kanəgatu   wennə   wey
           1SG-DAT   fight   about   sorry   EQ-INF   FUT
           'I will be sorry about the fight.'

      d.   maṭə   yannə   kanəgatu   wenne   næhæ
           1SG-DAT   go-INF   sorry   EQ-FOC   NEG
           'I won't be sorry to leave.'

The element *bayə* 'afraid' takes nominal arguments, and clausal and infinitive complements. These are illustrated in 17a to 17c. Thus *bayə* behaves exactly like *kæməti* with respect to complementation; it also requires its subjects to be in the nominative case.

(17) a.   mamə   satundə   bayai
           1SG   animal.PL   afraid-ASS
           'I am afraid of animals.'

      b.   ballat   ekə   sellam   kərənə   ekəṭə   ṭamay   maṭə   bayai
           dog   with   game   do-REL.PRES   COMP   ASS   1SG-DAT   afraid-ASS
           'It is playing with the dog that I'm afraid of.'

      c.   mamə   gederə   yannə   bayai
           1SG   home   go-INF   afraid-ASS
           'I'm afraid to go home.'

A difference between the two is that *bayə* can occur with a bare subject, as in example 17d. This is due to the fact that *bayə* is also a noun 'fear', a literal gloss for 17d would be *To me there is fear*. Note that the subject is in the dative case, unlike in 17a to 17c. Like the other elements, *bayə* sheds the assertive marker when another finite element is following it, as shown in 17e.

(17) d.   maʈə   bayai
            1SG-DAT   fear-ASS
            'I am afraid.'

      e.   mamə   gederə   yannə   bayə   næhæ
            1SG   home   go-INF   afraid NEG
            'I'm not afraid to go home.'

The distributional and morphological evidence suggests that *bayə* is a weakly grammaticized adjective that has developed out of a noun. Like the previous stance elements, it is not an auxiliary.

The element *oone* can express both the concepts 'need' and 'want'. Gair and Paolillo (1997:27) identify 'must' as another sense, which is likely to be related to examples like 18c, where the meaning of 'need' shades into 'should', and maybe even 'must'. The development of weak into strong modality is a common process. *Oone* takes nominal arguments, and clausal and infinitive complements, which are illustrated in 18a to 18c. The subjects are in the dative case.

(18) a.   maʈə   epel gediyak   oone
            1SG-DAT   apple CL-IND   need
            'I need an apple.'
      b.   redi   hodənə   ekəʈə   ʈamay   maʈə   onee
            clothes   wash-REL.PRES COMP   ASS   1SG-DAT   need
            'It is washing clothes that I want.'
      c.   oyaaʈə   kərannə   onee redi   hodənə   ekay
            2SG-DAT   do-INF   need clothes   wash-REL.PRES COMP
            'What you need to do is wash clothes.'

In sentences with auxiliaries and epistemic elements, *oone* precedes the other elements.

(18) d.   maʈə   epel gediyak   oone   wey
            1SG-DAT   apple CL-IND   need   EQ-FUT
            'I will need an apple.'
      e.   maʈə   epel gediyak   oone   wennə   puluwaŋ
            1SG-DAT   apple CL-IND   need   EQ-INF   PROBABLY
            'I will probably need an apple.'

The element *oone* is different from the other stance elements in a number of respects. In attributive sentences, *kæməti*, *bayə* and *kanəgatu* take the assertive suffix *-i*, which *oone* does not. It is also the only polysemous stance element. Finally, it is the only stance element that occasionally fuses with epistemic elements, as shown in 18f.

(18) f.   maʈə   epel gediyak   oonæhæ
            1SG-DAT   apple CL-IND   need-NEG
            'I don't need an apple.'

In summary, despite a number of morphological and syntactic differences, *kæməti*, *bayə*, *kanəgatu*, and *oone* form a discernable category of their own which can be appropriately called stance elements.

4. AUXILIARIES, EPISTEMIC ELEMENTS, AND STANCE ELEMENTS. There are two basic conclusions that can be drawn from the observations in section 3. First, there is evidence for a lexical category auxiliary in colloquial Sinhala. On the basis of synchronic semantic, morphological, and syntactic evidence it can be reasonably hypothesized that a small number of lexical verbs have come to acquire grammatical functions in Sinhala, losing some of their original category characteristics in the process. These elements have grammaticized to different extents, as measured by Heine's (1993) stage model of auxiliation.

The auxiliary *yannə* 'go' is a fully grammaticized, stage D element with the function of indicating future tense. The same characterization holds for *dennə* 'give', which codes permission. The verb *gannə* 'take' differs from the two previously discussed elements, because it is no auxiliary by itself. In the complex auxiliary *patan gannə* 'start take', it codes inceptive aspect. This construction illustrates a frequent pattern in Sinhala, which is the creation of complex verbs by conventionalization of a nominal compound element. Examples of such complex verbs based on *gannə* 'take' are *balaa gannə* 'look take', which means 'to look after someone', *labaa gannə* 'happen take' which means 'obtain', and *allaa gannə* 'touch take', which means 'catch'. The element *wennə* 'become' functions as a tense-carrying verbal element that is comparable to an equative copula. It is a highly grammaticized, semantically bleached element. Despite these facts, since it does not take clearly non-finite complements, it cannot be appropriately called an auxiliary.

All grammaticization processes that can be observed in the above elements are fairly well-attested cross-linguistically. However, the idiosyncrasies and polysemies of the individual constructions also underscore the finding that grammaticization paths can be motivated in a post-hoc fashion, but never be predicted.

The second conclusion from this study is that the category of quasi-verbs, as proposed by Gair and Paolillo (1997:26), can be divided into epistemic elements and stance elements along semantic, morphological, and syntactic criteria. Syntactically, we can draw the following generalization. Stance elements are followed by auxiliaries proper, which are followed by epistemic elements. this is schematized in (19).

(19)   COMPLEMENT > STANCE > AUX > EPISTEMIC

Auxiliaries, stance elements and epistemic elements have in common that they make structures finite if they occur as the last element in a sentence.

A morphological difference between stance and epistemic elements is that the former take the assertive suffix, and the latter inflect for different syntactic contexts. Semantically, epistemic elements refer to the likelihood or factuality of some event, while stance elements code a cognizer's attitude towards some state of affairs.

As quasi-verbs are a somewhat unusual lexical category from an Indo-European point of view, it would be interesting to further analyze the lexical sources of these elements, and to investigate whether the rise of this category is a language-internal development, or if it is the result of language contact.

These two conclusions raise a theoretical point, on which I would like to end the discussion. In the works of Heine (1993), Kuteva (2001), and others it is a theoretical given that auxiliaries develop out of lexical verbs. This is a matter of definition, rather than empirical investigation, and will not be disputed here. However, the existence of quasi-verbs in Sinhala show that elements can come to function in very similar ways to auxiliaries, but have nouns (*bayə*) or adjectives (*kanəgatu*) as their lexical sources. If auxiliary-like elements can be recruited from these sources, should we rather define auxiliaries in terms of their synchronic function or in terms of their historical origins? It has been tacitly assumed that these aspects are commonly in agreement, but the case of Sinhala suggests that a revision of this assumption might be necessary. A broader definition of auxiliary would encompass all grammatical markers of tense, aspect, or modality that co-occur with non-finite verbal complements, regardless of their historical lexical source.

References

Barlow, Michael and Suzanne E. Kemmer (eds.) 2000. *Usage-based models of language.* Stanford: CSLI.

Bavin, Edith L. 1995. The obligation modality in Western Nilotic languages. *Modality in Grammar and Discourse*, ed. by Joan L. Bybee and Suzanne Fleischman, 107-34. Amsterdam: John Benjamins.

Beneviste, Emile. 1968. Mutations of linguistic categories. *Directions for historical linguistcs: A symposium*, ed. by Winfred P. Lehmann and Yakov Malkiel, 85-94. Austin: University of Texas Press.

Bolinger, Dwight. 1980. Wanna and the gradience of auxiliaries. *Wege zur Universalienforschung: Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*, ed. by Gunter Brettschneider and Christian Lehmann, 292-99. Tübingen: Narr.

Bybee, Joan L., Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, aspect and mood in the languages of the world.* Chicago: University of Chicago Press.

Bybee, Joan L. and Östen Dahl. 1989. The Creation of Tense and Aspect Systems in the Languages of the World. *Studies in Language* 13.51-103.

Bybee, Joan L. and Paul Hopper (eds.) 2001. *Frequency and the Emergence of Linguistic Structure.* Amsterdam: John Benjamins.

Gair, James W. 1998. *Studies in South Asian Linguistics: Sinhala and other South Asian Languages.* New York: Oxford University Press.

Gair, James W. and John C. Paolillo. 1997. *Sinhala.* München: Lincom.

Garusinghe, Dayaratne. 1962. *Sinhalese. The spoken idiom.* München: Hueber.

Greenberg, Joseph. 1963. *Universals of Language.* Cambridge: MIT Press

Heine, Bernd. 1993. Auxiliaries. *Cognitive Forces and Grammaticalization.* Oxford: Oxford University Press.

Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization.* Cambridge: Cambridge University Press.

Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony.* Berlin: Mouton de Gruyter.

Hopper, Paul. 1987. *Emergent grammar.* Berkeley Linguistic Society 13.139-57.

Hopper, Paul and Elizabeth Closs Traugott. 1993. *Grammaticalization.* Cambridge: Cambridge University Press.

KUHN, THOMAS. 1970. *The Structure of Scientific Revolutions.* 2nd edition. Chicago: University of Chicago Press.

KUTEVA, TANIA. 2001. *Auxiliation.* Cambridge: Cambridge University Press.

NEWMAN, JOHN. 1996. *Give. A cognitive-linguistic study.* Berlin: Mouton de Gruyter.

STEELE, SUSAN. 1978. The category AUX as a language universal. *Universals of Human Language*, Vol. III, ed. by Joseph Greenberg, 7-45. Cambridge: MIT Press.

TRAUGOTT, ELIZABETH CLOSS and BERND HEINE. 1991. *Approaches to grammaticalization. Vol. I and II.* Amsterdam: John Benjamins.

Rice University
Department of Linguistics, MS 23
6100 Main Street
Houston, TX 77005-1892
hilpert@rice.edu

THE RELATIONSHIP BETWEEN CASE MARKING AND S, A, AND O IN SPOKEN SINHALA

CARMEN JANY

University of California, Santa Barbara

1. INTRODUCTION. In this paper I examine the relationship between case marking and S, A, and O in spoken Sinhala. I will demonstrate that case roles are not assigned on the basis of grammatical relations, but rather they depend on a series of semantic and lexical principles including volitivity, animacy, semantic roles, and definiteness. This paper will furthermore provide evidence for S, A, and O in spoken Sinhala and describe how they pattern together.

Case in its most traditional sense refers to the morphological marking by which some languages indicate the grammatical relation of each argument in a clause to a predicate (DeLancey 2001). As in many other Indo-Aryan languages, case markers in spoken Sinhala do not coincide on a one-to-one basis basis with any syntactic roles (Masica 1991:367, Blake 1994). In fact, subjects and objects are not distinguished by case marking in many sentences in Indo-Aryan languages. Both can occur in the nominative case (Masica 1991). In particular the category of subject or lack thereof has been widely discussed for Indo-Aryan languages (Masica 1991), including spoken Sinhala (Gair 1976, 1990). The properties generally associated with subjects, such as agency, animacy, verb agreement, nominative case, control of reflexivity, coreferential deletion, and topicality, do not coincide in the same noun phrase in many Indo-Aryan languages (Masica 1991). Subjects functioning as experiencers rather than as agents and marked with the dative case rather than the nominative are very common in these languages (Masica 1991, DeLancey 2001). They are also found in Sinhala (Gair 1976, 1990).

Given the lack of a coherent subject category in spoken Sinhala (Gair 1976) and the multifunctionality of case markers, grammatical relations will be discussed in terms of S, A, and O following the definitions given in Payne (1997), rather than in terms of subjects and objects. According to Payne (1997), S represents the only nominal of a single-argument clause, A describes the most agent-like argument of a multi-argument clause, and O the most patient-like argument of a multi-argument clause. If there is no argument in a clause which can be identified as an agent or patient on a semantic basis, then A and O are assigned to the arguments that are treated morphosyntactically in the same manner as prototypical agents or patients respectively (Payne 1997).

Grammatical relations independent of semantic and pragmatic influences are identified by a) case marking, b) participant reference on verbs, and c) constituent order (Payne 1997). In spoken Sinhala, however, case marking is not an indicator of grammatical relations given that any argument may appear in the nominative case, and there is no verb agreement. That leaves us with constituent order as a syntactic indicator for grammatical relations. Sinhala is a verb-final language, and arguments may change position to add or change focus in a sentence. Hence, constituent order needs to be used with caution as an indicator for grammatical relations. This paper describes mainly unfocused sentences with an unmarked word order[1], which is SV or AOV in Sinhala, in order to include constituent order in the discussion of grammatical relations. However, it will only be used as a tool for the identification of S, A, and

---

[1] Unmarked refers to the basic word order whereby no constituent is put into focus.

O. In certain instances word order may function as an indicator of semantic roles, such as in transitive clauses with two inanimate arguments, both in the nominative case.

Spoken Sinhala distinguishes four cases for inanimate nouns: nominative (direct), dative, genitive, and instrumental, and six cases for animate nouns: nominative, accusative, dative, genitive, instrumental, and vocative. Animate pronouns occur in all except for the vocative, while inanimate pronouns are found in all but the vocative and the accusative case. There are different case forms for singular and plural, and definite and indefinite nouns in the singular, as well as for inanimate and animate pronouns in the singular and plural (Gair and Paolillo 1997). Except for the genitive and vocative case, all cases occur on either S, A, or O arguments. S may be marked with the nominative (unmarked), dative, instrumental, or accusative case, while A can only occur in any of the first three. O can be marked as nominative (unmarked), accusative, or dative (Gair and Paolillo 1997). These patterns are summarized below.

S:	Nominative	A:	Nominative	O:	Nominative
	Accusative		–		Accusative
	Dative		Dative		Dative
	Instrumental		Instrumental		

TABLE 1. Possible case markers for S, A, and O.

Two key concepts for the understanding of case marking in spoken Sinhala are animacy and volitivity. The distinction between animates, including humans and animals, and inanimates, including objects and plants, is pervasive in the language. Different forms for these two categories are used for case markers, pronouns, demonstratives, and numerals. The distinction between volitive and involitive verbs is for the most part encoded in the verbal derivational morphology, as the verb pairs in Table 2 illustrate. The formal distinction between the two sets surfaces in different verb tenses and aspects including past, imperfective, future, and focused forms. Semantically, the involitive verbs are associated with non-volitionality, lack of control, and lack of agency. However, the correlation is 'by no means neat or complete', as Gair and Paolillo (1997:39) assert. Some involitive verbs are also used in a volitional sense and some verbs lacking an involitive derivation are essentially involitive.

Volitive Verb		Involitive Verb	
lissannə	'to slip, to slide'	*lissennə*	'to slip'
waṭannə	'to drop'	*wæṭennə*	'to fall'
marannə	'to kill"	*mærennə*	'to die'
naṭəwannə	'to boil'	*næṭəwennə*	'to let boil'
naṭannə	'to dance'	*næṭennə*	'to dance'
gahannə	'to hit'	*gæhennə*	'to shake'
ahannə	'to listen'	*æhennə*	'to hear'
riddannə	'to hurt'	*ridennə*	'to be hurt, feel pain'
toorannə	'to explain'	*teerennə*	'to understand'
balannə	'to watch'	*bælennə*	'happen to watch'
-	-	*dænennə*	'to feel, to perceive'
dakinnə	'to see'	*peennə*	'to see'

Table 2. Volitive/involitive verb pairs.

The last pair of verbs in Table 2 demonstrates that the semantic volitive/involitive distinction is not always encoded in the verbal derivational morphology. Some verb pairs are completely distinct lexical items with different roots. In addition, not all verbs have a volitive or involitive counterpart, as *dænennə* 'to feel, to perceive' demonstrates.

This paper is divided into three main parts. First, I will describe each of the three categories: S, A, and O separately in terms of case marking, verb morphology where applicable, and constituent order. Verbless and copular clauses will be treated separately. Second, I will examine the different arguments marked with the same case in search for common semantic and syntactic properties. Third, I will combine the two analyses to discuss the relationship between syntactic roles and case marking in spoken Sinhala. I will show that argument marking is not assigned on the basis of grammatical relations, but is dependent on a series of semantic properties of the argument, such as animacy, semantic role, and definiteness, and on the semantic and lexical properties of the verb, in particular on volitivity.

2. CASE MARKING ON S, A, AND O.

2.1. S ARGUMENTS. By definition S represents the only core nominal in a single-argument clause. In spoken Sinhala there are many verbs which can take only one argument. According to Gair and Paolillo (1997), S can be in the nominative, the accusative, the dative, or the instrumental case. In the variety of spoken Sinhala examined for this paper, the instrumental case is not used for S arguments. The consultant used the nominative case in the same examples presented with instrumental case in Gair and Paolillo (1997:33). Examples (1)-(4) illustrate these different case markings.

(1) ADV   S
    [mæturuwa-iŋ passe] [ee   aliya]   [ekəpaarəṭəmə]
      chant-?   after DIST.VIS.ANIM.SG elephant.SG.DEF.**NOM**   suddenly
    OBL   P
    [wanəyaṭə]   [aayet]   [diuw-a]
    jungle-GOAL   again   run-PST
    'After he chanted, that elephant suddenly ran again into the jungle.'

(2) S   P
    [mamə]   [diuw-a]
    1S.**NOM**   run-PST
    'I ran'

(3) S   P
    [maawə]   [wæṭe-nəwa]
    1S.**ACC**   fall.INVOL-IMPF
    'I am falling (involuntarily).'

(4) S   P
    [maṭə]   [næṭun-a]
    1S.**DAT**   dance.INVOL-PST
    'I danced.' (for some external reason, such as being possessed by a spirit)

As can be seen in (1)-(4), the single argument of a clause can be in the nominative, the accusative, or the dative case. The latter two markings, however, imply lack of control and lack

of volition. This is clearly indicated by the obligatory co-occurrence of an involitive form as part of the verbal derivational morphology (Gair and Paolillo 1997). Examples (5)-(8) will illustrate the formal volitive/involitive distinction. The difference between (3) and (4) lies in the semantics of S. Accusative and dative S both occur with involitive verb stems, and the S argument has no control over the action. While the accusative marking describes an affected undergoer[2], the dative marking describes an experiencer[3], someone who receives a sensory impression. The S argument *maṭə* 'I' in (4) does not control the action, but rather 'experiences' some external force. If the S argument is visibly affected, as in (3), the accusative case is used.

A key factor in case assignment for S arguments is the volitive/involitive distinction, encoded in the verbal derivational morphology. Involitive verb forms correlate for the most part with non-volitionality, lack of control, and lack of agency. However, the correlation is not complete. Some involitive verbs are used in a volitional sense, such as *hærennə* 'to turn' (Gair 1998), and some verbs lacking an involitive derivation are essentially involitive, such as *lissannə* 'to slip'. Examples (5)-(8) illustrate imperfective and past tense inflection of volitive and involitive verb stems.

(5) S   P
     [maawə]   [næṭe-nəwa]
     1s.**ACC**   dance.**INVOL**-IMPF
     'I am dancing (involuntarily).'

(6) S   P
     [maŋ]   [naṭə-nəwa]
     1s.**NOM**   dance-IMPF
     'I am dancing.'

(7) S   P
     [maawə]   [diu**n**-a]
     1s.**ACC**   run.**INVOL**-PST
     'I ran (involuntarily).' (something made me run)

(8) S   P
     [maŋ]   [diu**w**-a]
     1s.**NOM**   run-PST
     'I ran.'

In addition to showing the derivational volitive/involitive distinction, the examples above illustrate that case marking is not determined by the lexical verb in this case, but by the semantics of the entire clause. While only nominative case can occur with verb stems lacking involitive derivation, both accusative and dative case are found in clauses with involitive verb stems. Nevertheless, dative case is only found with a few verbs, such as *næṭennə* 'dance', *diwennə* 'run', and *ridennə* 'feel pain'.

The predicates presented so far are either motion or action verbs. Verbs describing a change of state can also have arguments with different case markings, as shown in (9)-(12).

---

[2] Actor and undergoer are viewed here as semantic macro-roles generalizing across specific semantic roles. While 'actor' comprises agent, experiencer, instrument, and other specific semantic roles, 'undergoer' subsumes patient, theme, and recipient, among others.

[3] An experiencer neither controls nor is visibly affected by an action. It is someone who receives a sensory impression (Payne, 1997).

(9) S   P
     [ohu]   [mærun-a]
     3s.M.**NOM**   die.INVOL-PST
     'He died.'
(10) S   P
     [ohuwə]   [mærun-a]
     3s.M.**ACC**   die.INVOL-PST
     'He died.' (reportedly rarely used)
(11) S   P
     [malbanduna]   [kædun-a]
     vase.SG.DEF.NOM   break.INVOL-PST
     'The vase broke.'
*(12) S   P
     [malbanduna-wə]   [kædun-a]
     vase.SG.DEF-ACC   break.INVOL-PST
     'The vase broke.'

In general, only animate arguments show possible case alternations. Examples (11)-(12) demonstrate that inanimates can only be marked with the nominative case. This makes sense as only animates can control an action, act with volitivity, be visibly affected, or receive a sensory impression.

Table 3 summarizes the possible case markings found on S arguments, some properties of the arguments, and some of the verbs with which they have been elicited. While both verb stem types, volitive and involitive, can occur with nominative case, only involitive verb stems are found with accusative or dative case. The same verb stem can occur with different case markings depending on the semantics of the S argument and the entire clause. In addition, as will be shown later with O arguments, the accusative case marker can sometimes be dropped. The result is a nominative case argument with no change in meaning. This explains the involitive verb forms occurring with nominative case S.

Case marking	Verbs	Argument properties
Nominative	<u>Volitive</u>: *lisṣannə* 'slip', *duwannə* 'run', *naṭannə* 'dance', *kadannə* 'break', *naṭəwannə* 'boil' <u>Involitive</u>: *wæṭennə* 'fall', *mærennə* 'die'	animate, inanimate,
Accusative[4]	<u>Involitive</u>: *wæṭennə* 'fall', *lisṣennə* 'slip', *mærennə* 'die', *diwennə* 'run', *næṭennə* 'dance', *kædennə* 'break', *næṭəwennə* 'boil', *gæhennə* 'shake', *kæpennə* 'cut' (reflexive), *æhennə* 'hear' (passive), *issennə* 'lift' (passive)	animate, affected undergoer
Dative	<u>Involitive</u>: *næṭennə* 'dance', *diwennə* 'run', *ridennə* 'feel pain'	animate, experiencer

Table 3. Case markings on S.

---

[4] Clauses with a reflexive or a passive meaning follow the same pattern, such as for example *maavə kæpuna* 'I cut myself' where the animate affected undergoer is accusative-marked.

While Table 3 summarizes the case marking patterns found with different verbs in clauses with a single argument S, Table 4 demonstrates a hierarchy of triggers for the different case markings. If the S argument is inanimate, it can only take nominative case. If it is animate, it can occur in all three cases depending on the verb stem type and semantics. With volitive verb stems, the argument is always in the nominative case. With involitive verb stems, case marking correlates with the semantics of the argument: accusative for an affected undergoer or dative for an experiencer.

1. Animacy	Inanimate		Animate		
2. Verb Stem			Volitive	Involitive	
3. Semantics				Affected Undergoer	Experiencer
4. Case	NOM		NOM	NOM[5] ACC	DAT
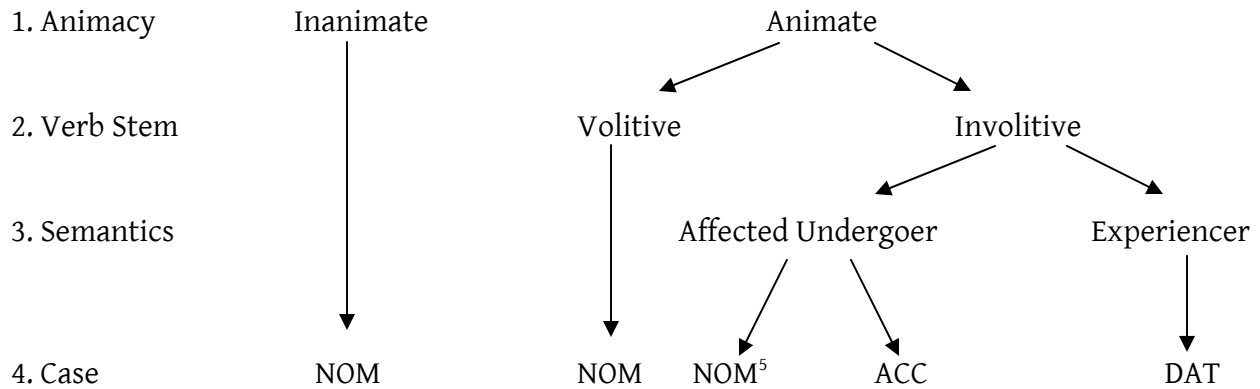
TABLE 4. Triggers for case marking on S

2.2. A ARGUMENTS. A arguments can be found in the nominative, the dative, or the instrumental case. The same as for S arguments, the instrumental case is not used in the variety of spoken Sinhala examined here. The consultant used the nominative case instead. Examples (13)-(16) illustrate the different case markings.

(13) A   O   OBL   P
      [maŋ]   [paləturə]   [pihiyə-kiŋ]   [kæpuw-a]
      1s.NOM   fruit.SG.DEF.NOM   knife.SG.IND-INST   cut-PST
      'I cut the fruit with a knife.'

(14) ADV   O
      [issrə]   [rajjuruwaŋ-ge   maalikaa-weweḍə-kərənə   minisun-ṭə]
      long.time.ago king.DEF-GEN   palace-LOC   work-do.PRES.PPL   people-DAT
      A   ADV   P
      [rajjuruwo]   [itaamat   hoňdiŋ]   [sæləkuw-a]
      king.DEF.NOM   very   well   treat-PST
      'In the old days, the king treated people working in the palace very well.'

(15) A   O   P
      [maṭə]   [sindu]   [æhe-nəwa]
      1s.DAT   music.IND.NOM   hear.INVOL-IMPF
      'I hear music.'

---

[5] This pattern only occurs when the accusative marker is dropped.

(16)   A   ADV
         [etəkoṭə   tamai] [maṭə]   [ættətəmə]
          then   EMPH   1s.**DAT**   truly
         P   O
         [teerune-e]   [meeke   bæræruŋkamə]
         understand.INVOL-FOC.PST   DEM-GEN seriousness
         'Then I truly understood the seriousness of this.'

While in (13)-(14) the A is a prototypical semantic agent, in (15)-(16) it is an experiencer receiving a sensory impression. Nevertheless, *maṭə* 'I' in (15) is more agent-like than *sindu* 'music', as it is animate and human. In addition, *maṭə* 'I' precedes *sindu* 'music' and, therefore, acts syntactically like a prototypical agent.

    Protoypical semantic agents are generally in the nominative case, as the following three examples illustrate. Their characteristics include: animacy, acting with volition, and affecting a patient.

(17)   A   O   P
         [maŋ]   [laməya-ṭə]   [gæhuw-a]
         1s.**NOM**   child.SG.DEF-DAT   hit-PST
         'I hit the child.'
(18)   A   O   P
         [maŋ]   [ohuwə]   [mara-nnə ya-nəwa]
         1s.**NOM**   3s.ACC kill-INF   go-IMPF
         'I am going to kill him.'
(19)   A   O   P
         [ohu]   [maawə]   [kæpuw-a]
         3s.M.**NOM** 1s.ACC   cut-PST
         'He cut me.'

Nevertheless, less prototypical agents in A function can also be marked with the nominative case, as in (20) and (21). While in (20) there is no affected patient and no volitional action, in (21) the A argument is not animate. Hence, in both cases the A is not a prototypical agent, but it is the more agent-like argument of the two, given its humanness in (20) and its agentive interpretation in both examples.

(20)   A   O   P
         [maŋ]   [gedərə]   [dækk-a]
         1s.**NOM**   house.SG.DEF.NOM   see-PST
         'I saw the house.'
(21)   A   O   P
         [aaňduwə]   [taxes]   [issuw-a]
         government.**NOM** taxes   change-PST
         'The government changed the taxes.'

A arguments in the dative case are never prototypical agents. Rather, they are experiencers in clauses with verbs of sensation or cognition. This is illustrated in the following examples.

(22)   A   O   P
           [maṭə]   [laməya-wə]   [pee-nəwa]
           1s.DAT   child.SG.DEF-ACC   see.INVOL-IMPF
           'I am seeing the child.'

(23)   A   O   P
           [laməya-ṭə]   [kataawə]   [teere-nəwa]
           child.SG.DEF-DAT   story.SG.DEF.NOM   understand.INVOL-IMPF
           'The child understands the story.'

(24)   A   O   P
           [maṭə]   [roṭi   pusmə]   [dænun-a]
           1s.DAT   roti.DEF.NOM smell.DEF.NOM   feel.INVOL-PST
           'I smell the roti.'

In the data examined, all of the verbs occurring in clauses with a dative-marked A argument are involitive, i.e. they belong to the e-conjugation class, called class III by Gair and Paolillo (1997). Class III verbs are 'essentially the same as those verbs that include the involitive morpheme' (Gair and Paolillo 1997:24). This shows that, the same as for S, an A in the dative case correlates with involitive verbal morphology[6] and with semantic experiencers.

    The findings for A arguments are summarized in Table 5. Dative case markings are only found on animate arguments and with involitive verb stems. There are a limited number of verbs, mostly indicating sensation or cognition, that can occur in clauses with dative-marked A arguments. Semantically, prototypical agents are marked with the nominative case, while experiencers take the dative case. Except for the accusative marked S arguments, the observed patterns are very similar for S and A arguments, as can be seen in Table 6.

Case marking	Verbs	Argument properties
Nominative	<u>Volitive</u>: *kannə* 'eat', *dakinnə* 'see', *tiintəgaannə* 'apply paint', *gahannə* 'hit', *tarəvatu kərannə* 'scold', *saləkənnə* 'treat', *andəgahannə* 'call', *kadannə* 'break', *marannə* 'kill', *kapannə* 'cut', *naṭannə* 'boil', *ahannə* 'listen', *kataa kərannə* 'talk to', *riddannə* 'hurt', *pavičči kərannə* 'use', *sænəsannə* 'console', *stutikərannə* 'thank', *udankərannə* 'help', *beerannə* 'save'	animate, inanimate, agents
Dative	<u>Involitive</u>: *æhennə* 'hear', *dænennə* 'feel', *peennə* 'see', *teerennə* 'understand' *læbennə* 'receive'	animate, experiencer

Table 5. Case markings on A.

---

[6] There is one exception to this rule. In clauses with *hambəvennə* 'meet', a verb with involitive morphology, the A appears in the nominative case. However, the verb does not retain its involitive meaning when used with a nominative A.

1. Animacy	Inanimate		Animate	
2. Verb Stem		Volitive		Involitive
3. Semantics				Cognition verbs/ experiencers
4. Case	NOM	NOM		DAT
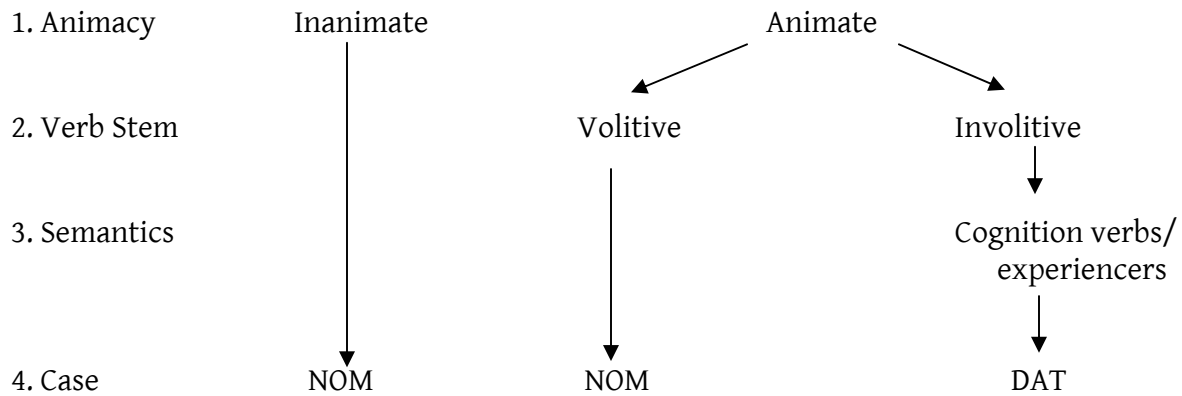
TABLE 6. Triggers for case marking on A.

2.3. O ARGUMENTS. O arguments occur in the nominative, the accusative, or the dative case. Examples with each of these case markings are given below.

(25)   A   O   P   (same as 14)
          [maṭə]   [sindu]   [æhe-nəwa]
          1s.DAT   music.IND.**NOM**   hear.INVOL–IMPF
          'I hear music.'

(26)   OBL   O   O
          [samaharə perəhærə-wələ] [makərə nætuŋ]   [siŋha nætuŋ]
          some   parade.PL-LOC dragon dance.PL.**NOM**   lion   dance.PL.**NOM**
          O
          [saha]   [paatə paatə æňduŋ   æňdə-gat   minisu
          and   color color cloth.PL   wear-PST.PPL people.NOM

          ekə ekə   vikəṭə   javənikaa   pavatvəm-iŋ
          one one   funny   performance.PL.**NOM**   perform-PPL
                                         A   P
          paarə   digee ya-nəva]   [mamə]   [dækk-a]
          road   along go-IMPF   1s.NOM   see-PST
          'In some parades I saw dragon dances, lion dances, and funny acts by people wearing colorful clothes while going along the road.'

(27)   A   O   P
          [maṭə]   [ohuwə]   [æhun-a]
          1s.DAT   3s.**ACC**   hear.INVOL-PST
          'I heard him.'

(28)   A   O   P
          [maŋ]   [meesə-ṭə]   [gæhuw-a]
          1s.NOM   table.SG.DEF-**DAT**   hit-PST
          'I hit the table.'

Examples (25)-(27) demonstrate that animate O arguments may be in the accusative case, unlike inanimate O arguments that are in the nominative case, as in (25) and (26), or in the dative, as in (28). Gair and Paolillo (1997:31) assert that the use of the accusative marker for objects (here O arguments) shows dialectal variation. Most animate arguments are marked with the accusative case, with a few exceptions. While definite O arguments always take the accusative marker, it can be omitted in some indefinite O arguments, as in (30).

(29)   A   O   P
        [ohu]   [eluva-və]   [mæruv-a]
        3s.M.NOM   goat.SG.DEF-**ACC**   kill-PST
        'He killed the goat.'

(30)   A   O   P
        [ohu]   [eluvek]   [mæruv-a]
        3s.M.NOM   goat.SG.IND.**NOM**   kill-PST
        'He killed a goat.'

The O argument in (30) is animate, the same as in (29), but it is indefinite. In general, the consultant varies in the use of the accusative marker with indefinite animate O arguments, whereas definite animate O arguments do not show such a pattern. The correlation between the accusative case marker and indefiniteness needs further investigation.

The alternation between accusative and dative case is lexically specified by the verb rather than being a property of the O argument, as a comparison of the previous examples with (31) and (32) demonstrates. Both O arguments are typical patients in that they are both animate and affected. Nevertheless, *minisu* 'people' in (31) and *laməya* 'child' in (32) are marked with the dative, while *eluva* 'goat' in (29) takes the accusative case. Following Gair and Paolillo (1997:31) 'some verbs require the dative case of an animate object'. Such verbs include: *gahannə* 'hit', *tarəvatu kərannə* 'scold', *saləkənnə* 'treat', and *andəgahannə* 'call', among others. However, in a few instances where the semantics of the verb allows it, such as with *gahannə* 'hit' in (28), this pattern extends to inanimate O arguments. In general, many of the verbs with O arguments in the dative case are verbs of speaking where the O can be interpreted as an addressee. In these cases, the dative-marking can be viewed as an extension of the ditransitive argument structure of verbs of speaking. The following examples illustrate the dative case marking.

(31)   ADV   O
        [issrə]   [rajjuruwaŋ-ge maalikaa-we weḍə-kərənə   minisun-ṭə]
        long.time.ago   king.DEF-GEN   palace-LOC   work-do.PRES.PPL people-**DAT**
        A   ADV   P
        [rajjuruwo]   [itaamat   hoňdiŋ]   [sæləkuw-a]
        king.DEF.NOM   very   well   treat-PST
        'In the old days, the king treated people working in the palace very well.'

(32)   A   O   P
        [amma]   [laməya-ṭə]   [tarəwaṭu kər-a]
        mother.SG.DEF.NOM   child.SG.DEF-**DAT** reprimand do-PST
        'The mother scolded the child.'

(33)   A   O   P
        [maŋ]   [miniha-ṭə]   [aňdəgæhuw-a]
        1s.NOM   man.SG.DEF-**DAT**   call-PST
        'I called the man.'

The same as for S arguments, only animate O arguments can take the accusative case. Inanimates in all three syntactic categories only occur in the nominative and the dative cases.

The findings for O arguments are summarized in Tables 7 and 8. As with S and A arguments, the main factor influencing case marking is animacy. While most[7] inanimates occur in the nominative case, animates are found in any of the three cases. However, only indefinite animate O arguments have been found in the nominative. The volitive/involitive verb distinction does not seem to play an essential role in the case marking of O arguments. Nevertheless, all attested dative-marked O arguments occur with volitive verbs. A clear semantic distinction between O arguments in the accusative and in the dative cannot be established at this point. While many of the dative O arguments represent beneficiaries, some accusative O arguments can equally be described as beneficiaries. The only semantic pattern observed is that many dative-marked O arguments are addresees of verbs of speaking.

Case marking	Verbs	Argument properties
Nominative	<u>Volitive</u>: *kannə* 'eat', *dakinnə* 'see', *tiintəgaannə* 'apply paint', *kadannə* 'break', *naṭannə* 'boil', (*marannə* 'kill') <u>Involitive</u>: *æhennə* 'hear', *dænennə* 'feel'	inanimate, some animate indefinite[8]
Accusative	<u>Volitive</u>: *dakinnə* 'see', *marannə* 'kill', *kapannə* 'cut', *tiintəgaannə* 'apply paint', *konitannə* 'pinch', *pavičči kərannə* 'use', *sænəsannə* 'console', *beerannə* 'save' <u>Involitive</u>: *æhennə* 'hear', *peennə* 'see', *teerennə* 'understand', *hambəvennə* 'meet'	animate
Dative	<u>Volitive</u>: *tarəvatu kərannə* 'scold', *saləkənnə* 'treat', *andəgahannə* 'call', *gahannə* 'hit', *kataa kərannə* 'talk to', *kæægahannə* 'shout', *riddannə* 'hurt', *stutikərannə* 'thank', *udankərannə* 'help', *ugananna* 'teach'	animate, inanimate, lexical specification of verb

Table 7. Case markings on O.

---

[7] Example 28 illustrates that some inanimate O arguments occur in the dative case as a function of the verb.
[8] This pattern only occurs when the accusative marker is dropped in cases with an indefinite animate O.

1. Verb (lexical property)	Most verbs				Lexical specification of verbs
2. Animacy	Inanimate	Animate			
3. Definiteness		Indefinite	Definite		
4. Case	NOM	NOM ACC	ACC		DAT
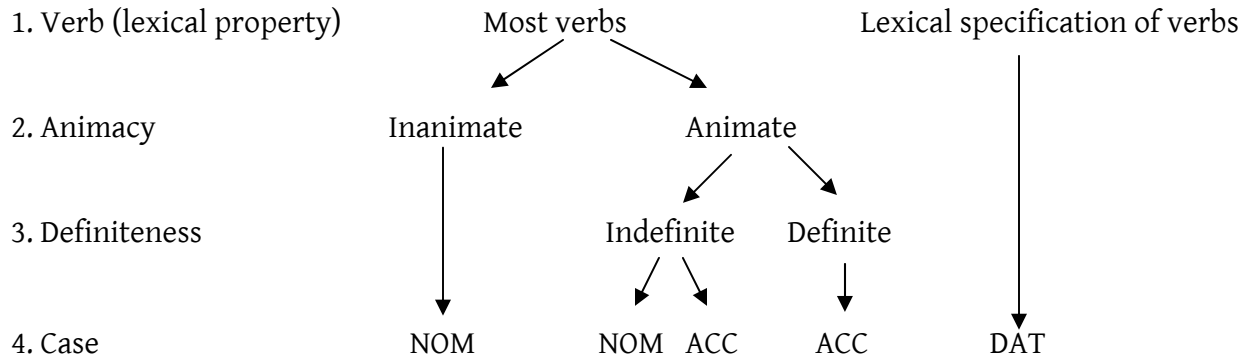
TABLE 8. Triggers for case marking on O.

3. VERBLESS AND COPULAR CLAUSES. Verbless and copular clauses behave differently from verbal clauses. It is difficult to assign S, A, and O roles is such constructions. Therefore, they are discussed separately here.

In spoken Sinhala some clauses with predicate nominals include copular verbs with no special marking on the predicate nominal, while others lack a copula but show a predicative suffix. These properties will not be discussed here. Only case markings on arguments in these clauses are described.

Equational, existential, and locational clauses show nominative case marking on their arguments. This is illustrated below. Example (34) represents an equational, (35) an existential clause, and (36) a locational construction.

(34)　NP　　　　　NP
　　　[Nimal]　　　[dostərə　　　kenek]
　　　Nimal.**NOM**　　doctor.SG.IND　person.SG.IND.**NOM**
　　　'Nimal is a doctor.'

(35)　NP　　　　　COP　　　　　NP
　　　[Nimal]　　　[i-nne]　　　　[koləm̆bə]
　　　Nimal.**NOM**　　be.ANIM–FOC.NPST　Colombo.**NOM**
　　　'Nimal is in Colombo.'

(36)　OBL　　　　　NP　　　　　　COP
　　　[gahe-e]　　　[wanḍurek]　　　[in-nəwa]
　　　tree.SG.DEF–LOC　monkey.SG.IND.**NOM**　exist.ANIM–IMPF
　　　'There is a monkey in the tree.'

Nominative case is also found in other constructions with predicate nominals. In most clauses with predicate adjectives the argument is nominative-marked, and in some clauses with predicate nominals one of the arguments takes the nominative case, as the following examples show.

(37)　NP　　　P
　　　[æyə]　　[usa-y]
　　　3S.F.**NOM**　tall-PRED
　　　'She is tall.'

(38)   NP   NP   P
        [maŋ]   [ee   laməya-ʈə]   [aasa-y]
        1s.**NOM**   DIST.SG.ANM   child.SG.DEF-DAT   like-PRED
        'I like that child.'

(39)   NP   P
        [andəre]   [siini   ka-nnə]   [harimə]   [aasa-y]
        Andare.**NOM**   sugar   eat-INF   very   like-PRED
        'Andare likes to eat sugar very much.'

The predicate expressing an emotion in (38) and (39) behaves in the same way as adjectives in spoken Sinhala in that it takes the predicative suffix *-y*. The arguments with the most resemblance to an agent, *maŋ* 'I' in (38) and *andəre* 'Andare' in (39), are marked with the nominative case, while the object of liking in (38), *ee laməya* 'that child', shows dative marking. The dative marking is also found on inanimate objects of liking, as the following example demonstrates.

(40)   NP   NP   P
        [maŋ]   [roʈiwələ-ʈə]   [aasa-y]
        1s.NOM   roti.IND-**DAT**   like-PRED
        'I like roti'

In addition to the object of liking, dative case markings occur in several predicate nominal clauses, including possessives. The dative-marked arguments all represent undergoers. This is illustrated below.

(41)   NP   NP   COP
        [maʈə]   [salli]   [tiye-nəwa]
        1s.**DAT**   money.IND.NOM   exist-IMPF
        'I have money.'

(42)   NP   NP   P
        [maʈə]   [oyaawə]   [matəka-y]
        1s.**DAT**   2s.ACC   remember-PRED
        'I remember you.'

(43)   NP   P
        [ehe innə   gamaŋ] [maʈə] [matəka-y]   [hæmə aurudæ-mə
        there be.ANIM.INF while   1s-**DAT**   remember-PRED every year-?
                                            NP
        pebərəwaari maase nætaŋ janəwaari maase] [ciina   alut aurudə]
        February   month if.not January   month Chinese   new year
        P
        [samərə-nəwa]
        celebrate-IMP
        'While I stayed there, I remember, every year in February if not in January the Chinese New Year is celebrated.'

(44) NP   P
      [maʈə]   [niḍimata-y]
      1s.**DAT**   sleepy-PRED
      'I am sleepy.'

The possessor in (41) is in the dative case, while the possessed, here an inanimate object, takes the nominative case. Animate objects of possession, such as *ballek* 'dog' have also been found in the nominative case. The animate object of memory in (42), however, shows accusative marking. The undergoers in (42) and (43), *maʈə* 'I', as well as the only argument in (44), also show dative marking. Example (44) illustrates that predicate adjectives describing a bodily function have a dative-marked argument, while adjectives defining inherent properties, as in (37), correlate with nominative-marked arguments. Other bodily functions with dative-marked arguments include *badəginii* 'to be hungry' and *siitəla* 'to be cold', as (45)-(46) illustrate.

(45) NP   P
      [maʈə]   [badəgini-yi]
      1s.**DAT**   hungry-PRED
      'I am hungry.'

(46) NP   P
      [maʈə]   [siitəla-y]
      1s.**DAT**   cold-PRED
      'I am cold.'

The findings of this section are summarized in Table 9. Contrary to the findings in clauses with verbal predicates, case marking is for the most part independent of animacy. Only animate objects of memory are accusative-marked, while inanimates are in the nominative. Arguments in the dative case can be interpreted as experiencers. To conclude, in clauses with predicate nominals, the semantic role of the argument determines its case assignment.

Case marking	Type of clause	Argument properties
Nominative	equational, existential, locational, predicate adjective (inherent property), actor of *aasa* 'like', possessed, object of memory (inanimate)	animate, inanimate
Accusative	object of memory (animate)	animate
Dative	object of *aasa* 'like', actor/possessor of *matəkə* 'remember', possessor, predicate adjective (bodily function)	animate, inanimate

Table 9. Case markings on arguments in verbless and copular clauses.

4. SEMANTIC AND SYNTACTIC GROUNDS FOR CASE MARKING.

4.1. THE NOMINATIVE CASE. The nominative case is the unmarked or direct case (Gair and Paolillo 1997). It is most often used in spoken Sinhala and appears on S, A, and O, on actors and undergoers, on animate and inanimate arguments, and with most verbs. Furthermore, it is used in equational, existential, and locational clauses, as well as with certain predicate adjectives. It is also found on inanimate objects of memory and on a possessed object or

animal. Given this great variety of arguments marked with the nominative case, no unique property can be identified. However, contrasting nominative-marked arguments with arguments marked differently, it is worth pointing out that most inanimate arguments and some animates take the nominative case. One structure where animates and inanimates behave alike is the possessed which always appears in the nominative case.

4.2. THE ACCUSATIVE CASE. The accusative case only appears on a limited number of animate arguments either in S or in O role or as an object of memory. Apart from being animate, all the arguments marked with the accusative case lack control or volitivity. They represent for the most part directly affected undergoers. Furthermore, they are rather the endpoint of an action than the starting point.

4.3. THE DATIVE CASE. The same as the nominative case, the dative case is found on arguments in S, A, and O roles. It appears mostly on animates with a few exceptions. An inanimate object of liking in clauses with *aasa* 'to like' takes dative case. Arguments marked with the dative case are never prototypical actors. The object of liking, the experiencer of memory, arguments of adjectival predicates describing a bodily function, as well as possessors, are all dative-marked and can be described as experiencers or undergoers which lack control and volitivity. They are rather passive than active participants in an event. The same as with arguments in the accusative case, dative-marked participants represent rather the endpoint of an action than the starting point.

5. CONCLUSIONS: THE RELATIONSHIP BETWEEN CASE MARKING AND S, A, AND, O. It has been shown that S, A, and O and case marking do not coincide on a one-to-basis, a fact that has been discussed for Indo-Aryan languages (Masica 1991) in general, as well as for spoken Sinhala in particular (Gair 1976, 1990, Gair and Paolillo 1997). An S argument can be marked with the nominative, the accusative, or the dative case as a function of animacy, semantic role, and volitivity. An A argument can occur in the nominative or the dative case as a function of semantic role, animacy, volitivity, and verb type. O arguments are marked in the nominative, the accusative, or the dative case depending on animacy, definiteness, volitivity, and on the lexical property of the verb. Hence, case marking is largely dependent on the animacy of the argument, its semantic role, volitivity or the lack thereof as encoded in the verb stem, and occasionally on lexical properties of the verb. It thus encodes the semantic rather than the syntactic function of an argument in a clause.

Syntactically, there is no evidence from case marking that any two syntactic roles share sufficient behavior to form grammatical relations, such as subject or absolutive. Nevertheless, a closer look reveals some subsystems. S and A pattern together in that typical agents acting with volitivity are nominative-marked, while experiencers in either S or A role take the dative case. S and O pattern together as directly affected animate undergoers are marked with the accusative. Given the different behavior of S arguments, it could be argued that Sinhala has a Split-S system, which is in part a nominative-accusative system patterning S and A together, in part an ergative system patterning S and O together. However, due to the complexity of case marking for each of the roles, it seems better to examine case marking on a semantic basis.

Given that the same case marking can be used for different semantico-syntactic roles in a clause, some inherent semantic properties of arguments, such as animacy and definiteness, as

well as word order, play an essential role in identifying the syntactic and semantic function of an argument in a clause.

The possible cooccurrences of case markings in a clause need further investigation. The examples presented in this paper demonstrate the following possible combinations of case markings in a clause: a) nominative and nominative, b) nominative and accusative, c) nominative and dative, d) dative and accusative. Given that A arguments do not take the accusative case, two accusative-marked arguments will not occur in the same clause. Furthermore, two dative-marked arguments in a clause are not possible, given that dative A arguments occur with involitive verb stems, while dative O arguments are found with volitive verbs only. However, clauses with more than two arguments have yet to be examined.

In this paper I have shown that case is not assigned on the basis of syntactic roles. Rather, it depends on a series of semantic properties of the argument, occasionally on lexical properties of the verb, and sometimes on the semantics of the entire clause.

This paper represents only a starting point in the investigation of the relationship between case marking and semantico-syntactic roles in spoken Sinhala. The correlation between definiteness of the argument and accusative case marking needs further investigation. Furthermore, clauses with more than two arguments, as well as possible combinations of case marking in a clause yet need to be examined. In addition, further evidence may be sought by analysing complex sentences with clausal arguments.

REFERENCES

BLAKE, BERRY. 1994. *Case*. Cambridge: Cambridge University Press.

DELANCELY, SCOTT. 2001. *Figure and Ground in Argument Structure*. Summer Institute in Linguistics, University of California, Santa Barbara, Lecture 3.

GAIR, JAMES W. 1976. Is Sinhala a Subject Language. *The Notion of Subject in South Asian Languages*, ed. by Manindra K. Verma. South Asian Studies, University of Wisconsin-Madison. Publication 2.

GAIR, JAMES W. 1990. Subjects, Case, and INFL in Sinhala. *Experiencer Subjects in South Asian Languages*, ed. by Manindra K. Verma and K.P. Mohanan. Stanford University: The Center for the Study of Language and Information.

GAIR, JAMES and JOHN C. PAOLILLO. 1997. *Sinhala.* München: Lincom Europa.

GAIR, JAMES and JOHN C. PAOLILLO. 1998a. Sinhala Nonverbal Sentences and Argument Structure. *Studies in South Asian Linguistics: Sihnala and other South Asian languages*, ed. by James W. Gair and Barbara Lust, 87-107. New York: Oxford University Press.

GAIR, JAMES and JOHN C. PAOLILLO. 1998b. Action Involvement Categories in Colloquial Sinhala. *Studies in South Asian Linguistics: Sihnala and other South Asian languages*, ed. by James W. Gair and Barbara Lust. 25-43. New York: Oxford University Press.

GIVON, TALMY. 1984. *Syntax: A Functional-typological introduction Vol. I.* Amsterdam: John Benjamins.

KEENAN, EDWARD. 1976. Towards a universal definition of "subject". *Subject and Topic.* ed. by Charles Li. New York: AP.

MASICA, COLIN. 1991. *The Indo-Aryan Languages.* 339-369. Cambridge: Cambridge University Press.

PAYNE, THOMAS E. 1997. *Describing morphosyntax: A guide for field linguists.* Cambridge: Cambridge University Press.

University of California, Santa Barbara
Department of Linguistics, 3607 South Hall
Santa Barbara, CA 93106
cjany@umail.ucsb.edu

THE TRANSFORMATION OF THE VERTICAL AXIS TO HORIZONTAL:
A CASE STUDY FROM SINHALA[*]

You-Jing Lin

University of California, Santa Barbara

1. INTRODUCTION. This paper presents and analyzes a case in which the Sinhala language[1] applies spatial postpositions for the vertical axis to denote horizontal relationships. Among the three axes proposed by Fillmore (1982:36-7) to be common in the linguistic treatment of spatial notions--- i.e. up/down, front/back, and left/right--- the vertical axis has usually been associated to the pull of gravity. Sinhala, however, demonstrates a case where the forms denoting the vertical axis can be used to describe spatial relationships that are in fact horizontal.

The remainder of the paper will start with an account of an experiment for spatial conception, the result of which reveals the transformed usage of the verticality terms in Sinhala (§2). The following section (§3) discusses the mental (or cognitive) manipulation of the vertical axis, which is assumed to have given rise to the phenomenon in question. The analysis is based upon a framework involving image-schema operations proposed by Ekberg (1997) (§3.1). In §3.2, I present cases in which the vertical and frontal axes are interchangeable, followed by cases in which they are not. To account for the interchangeability and restrictions, it is proposed that only by switching from the route perspective to survey perspective can the speaker activate the image-schema manipulation (Vertical → Horizontal), and the blockage of the reference point's view by one of the focal participants to the other can inhibit the perspective switching, and in turn block the transformation from the vertical to horizontal (frontal). Section 4 summarizes and concludes the paper.

2. VERTICAL TO HORIZONTAL IN SINHALA. The fact that Sinhala can use verticality terms to code horizontal meaning was discovered in the results of a mini-project designed to figure out whether a language applies egocentric/relative or absolute coding in describing spatial relationships. [2] In the project presented in this paper, two tables are set up in an L-shaped arrangement. On one of the tables five objects (i.e. a book, a tissue box, a cup, a bottle, and a walkman) are arranged in an array as shown in Fig. 1.

---

[*] I would like to thank Carol Genetti, Carlos Nash, Robert Englebretson, John Paolillo, Valerie Sultan, and Martin Hilpert for helpful discussions, input, and advice.
[1] Sinhala is an Indo-Aryan language mainly spoken in Sri Lanka. My consultant is Oshan Fernando. Aged 30 in 2005, Oshan is a native speaker of Sinhala.
[2] The project was adapted by Susanna Cumming based on the project introduced in Levinson (1996). The primary difference between the current project and Levinson's is that in the former the array is rotated 90°, while in the latter the rotation is 180°.
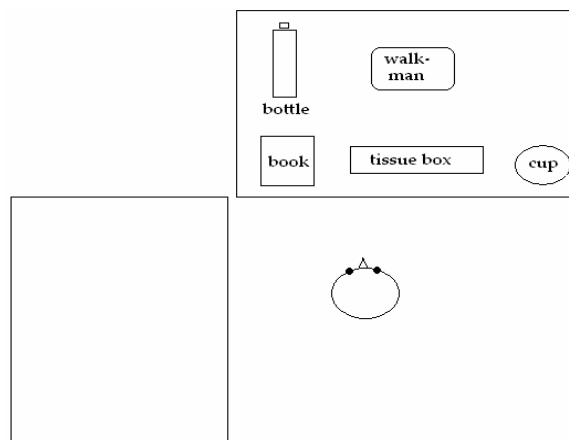
FIGURE 1. Object arrangement for the elicitation of spatial description in the mini project.

The consultant was required to describe the array so that someone else could arrange the objects in the same way according to his instruction. The consultant was actually encouraged to speak as colloquially and naturally as possible. Then the speaker was asked to move the objects to the other table and arrange them in the same way as they were arranged on the first table. The description was recorded, transcribed and glossed with the help of the consultant.

An especially appealing finding from the results of the project, which turned out to be the main theme of the present study, is the way in which Sinhala profiles two of the spatial relationships as shown in Figure 1, namely the relation between the bottle and the book, and the relative location of the walkman to the tissue box. Languages like English, Persian and Chinese tend to represent these relations with a horizontal (or 'frontal') axis (For example, 'The bottle is behind the book' and 'The tissue box is in front of the walkman'). The Sinhala consultant, however, chose to code them using vertical postpositions meaning 'above' (*uḍiŋ*) and 'below' (*yaṭiŋ*), as shown in 1-2.[3]

---

[3] Abbreviations in these examples include: ABL 'ablative', DAT 'dative, DEF 'definite', FOC 'focus', GEN 'genitive', INAN 'inanimate', IND 'indefinite', LOC 'locative', NPST 'non-past', PL 'plural'.

(1) *potə=ṭə*[4]   **uḍiŋ**[5]   *bootáləyak*   *tiye-nəwa*
     book.DEF=DAT.DEF   **above**   bottle.IND   be.INAN-NPST
     'There is a bottle above the book.'
(2) *bootəlee=ṭə*   *dakunu*   *pettə=e*   *saha*   *tišupettiyə=ṭə*
     bottle.DEF=DAT.DEF   right   side=LOC.DEF.INAN   and   tissuebox.DEF=DAT.DEF
     uḍiŋ   *vookmənekak*   *tiye-nəwa*
     above   walkman.IND   be.INAN-NPST
     'On the right side of the bottle, and above the tissue box is a walkman.'

Three axes have been proposed by Fillmore (1982:36-7) as common in the treatment of spatial notions in natural language semantics: up/down, front/back, and left/right. Among these, front/back tends to be anthropocentric, and the up/down axis refers to relations existing independently of communication act participants since it actually takes the direction of the pull of gravity as its reference. A prototypical above relation is thus one with the figure object being at the same horizontal coordinate and higher than the reference object (Hayward and Tarr 1995:78-9).

---

[4] While the object NP of simplex postpositions (i.e., postpositions that cannot be further analyzed into a relational noun and a case marker) tend to take no case marking; the object NP of composite postpositions can take not only genitive case, but also dative or even ablative case. For instance, composite postpositions based on *eliyə* 'outside' very often have their object to be in ablative case:

(1) *gedərə=iŋ*   *eliyə=e*   *lamai*   *sellaŋkərə-nəwa*
     house.DEF=ABL.DEF.INAN   outside=LOC   child.PL   play-NPST
     'The children are playing out of the house.'

The distinction made in describing the figure below suggests that different case marking on the postposition object conveys additional information between the Figure and the Ground. In this case, the key semantic determinants in the selection between dative and genitive cases are support and contact. That is to say, in describing the location of the handprint relative to the window, the Ground ('window') must take genitive case, as a whole major surface of the handprint is in contact with and occupies a part of the most salient dimension of the window. Dative case, on the other hand, highlights rather a trajector between the Figure and the Ground, thus does not serve as an apt choice here.

(2) attə   tiyenn-e   **janeele/*janeele=ṭə**   dakunu   pætte
     hand   be.INAN-FOC   **window.GEN.DEF/window.DEF.DAT**   right   side.LOC
     'The hand is on the right section/side of the window.'

English can also make such a distinction with 'The hand is **on** the window's right side' and 'The flower is **to** the window's right side'. What makes the case in Sinhala especially interesting is that such a semantic load does not fall on postpositions, but on the case marker taken by the postposition NP.
[5] An observation made by the consultant testifies to the affinity between *uḍiŋ* 'above' and *uḍə* 'on; on top of'. While analyzing the compositionality of *uḍiŋ*, he contends that it is certainly a combination of *uḍə* + *=iŋ* ('on' + ABL). The form coding the opposite orientation to *uḍiŋ* 'above', *yaṭiŋ* 'below', is also formed via the same process from *yaṭə* 'below; underneath'. The main distinction between the two groups (i.e., *uḍiŋ-yaṭiŋ* versus *uḍə-yaṭə*) lies on the factor 'attachment'.

With the arrangement of the objects shown in Figure 1, however, it is obvious that in these sentences the notion involving 'the direction of the pull of gravity' is totally absent. In fact, it looks like that the speaker adopts vertical terms to profile horizontal relations, taking objects further away to be higher, and those that are closer to be lower.

Such a usage of the vertical axis, however, does not mean Sinhala is devoid of the notions and forms for the frontal axis. The following example shows that the language does have frontal postpositions:

(3) idiripiṭə 'in front of'
    liikoṭəyak   **idiripiṭə**   perəliccə   muṭṭiyak tiye-nəwa
    stump.IND   **in.front.of** overturned pot.IND   be.INAN-NPST
    'There is an overturned pot in front of a tree stump.'

Like most Western languages, Sinhala uses what Hill (1975) calls an 'ego-opposed' strategy in coding the frontal relationship. In this strategy, the reference point (here the speaker) and the Ground (the stump) are facing each other, so the Figure (the pot), situated between the two, is in front of the stump. In other words, the example in 3 can be interpreted as:

(4) The pot is near the tree stump, on the side of the sump closest to me.

3. MENTAL MANIPULATION OF THE VERTICAL AND HORIZONTAL AXES. In this section, an account for the phenomenon in question will be proposed based on the analysis of image-schema manipulation by Ekberg (1997). A range of evidence provided in the previous literature on space and language has attested to the semantic nature of spatial relations. It has been observed that our daily perceptual interaction in the world can derive basic and simple cognitive patterns (i.e. image-schema), and spatial markers (in particular prepositions) denote the location of the Figure with respect to the Ground usually along one and rarely more than two spatial axes (Hayward and Tarr 1995:79). Image schemas, according to Johnson (1987), are schematic and retain only visual and force-dynamic properties rather than having propositional structure. They are 'constantly operating in our perception, bodily movement through space, and physical manipulation of objects' (Johnson 1987). Therefore, image-schema rotation is correlative to physical-object manipulation.

3.1. IMAGE-SCHEMATIC OPERATION: FROM VERTICAL TO HORIZONTAL. Based on the cognitive and linguistic nature of spatial conception and description, Ekberg (1997) presents four common principles for image-schema operations. The one that can be used to account for the situation in Sinhala is Principle A:

Principle A: Vertical Axis → Horizontal Axis
e.g. 'He walked up and down the corridor.' (Ekberg 1997:71)

When Principle A is applied in a deictically unspecified system, the ground level serves as the unmarked conceptual reference point. That is, in the schema, the reference point coincides with the lower end of a vertical axis that goes upward, and the reference point is away from the upper end no matter whether the vertical axis is upright or tipped, as diagrammed in Figure 2.
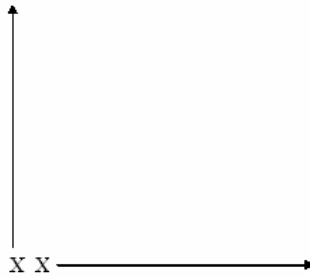
FIGURE 2. The deictically unspecified version of the (transformed) vertical axis
(Adopted from Ekberg 1997:71)   x= reference point

Such an account accords perfectly with what is observed in Sinhala as shown in 1-2. For these cases, the vertical axis ('above'/'below') can be assumed to have been 'tipped'. The reference point (the speaker) is situated at the lower pole of the axis, with the object 'above' being 'away from', and the object 'below' being 'toward' the reference point, which indeed makes a scenario that corresponds to the 'ego-opposed' frontal schema demonstrated in 3-4. A diagram illustrating the image-schema operation for 1 is provided below:

(5)
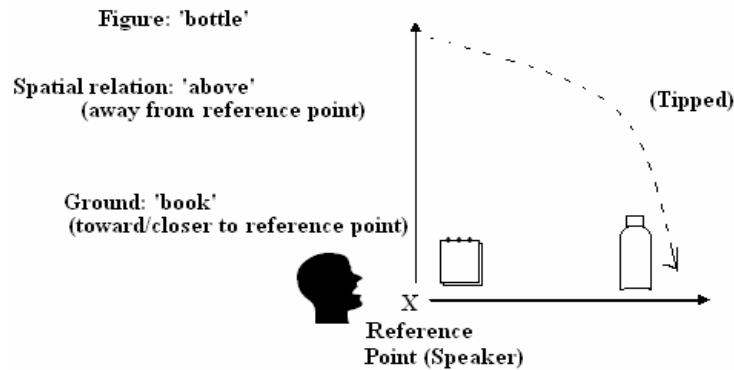
3.2. VERTICAL AND FRONTAL AXES: INTERCHANGEABILITY AND RESTRICTIONS. Further investigation with the consultant on the spatial description for Figure 1 reveals that the frontal axis is also applicable to specify the spatial relation between the book and the bottle, as well as the location of the tissue box relative to the walkman. Compare the following example with 2.

(6) bootəlee=ṭə   dakunu   pettə=e   saha tišupettiyə=ṭə
     bottle.DEF=DAT:DEF   right   side=LOC.DEF.INAN   and   tissuebox.DEF=DAT.DEF
     pitipasse   wookmənekak   tiye-nəwa
     behind   walkman.IND   be.INAN-NPST
     'On the right side of the bottle, and behind the tissue box is a walkman.'

Here the same spatial relation can be coded appropriately using either the vertical or frontal axis. This means in specific cases the two axes are interchangeable. But what are the primary principles that governs this interchangibility? Is there any limitation? These are the issues to be investigated in the remainder of this section.

There are, as one can image, situations where the vertical axis is the only apt choice. The spatial arrangement of a hat and two books in Figure 3 shows a prototypical situation for the vertical axis, in which the objects are at approximately the same horizontal coordinate, but show a significant difference on the gravitational (i.e., vertical) axis:
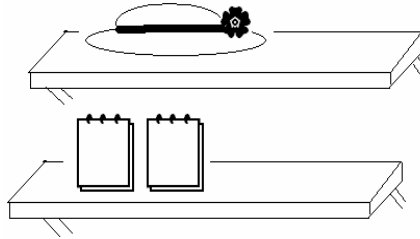
FIGURE 3.

As shown in the examples below, the spatial relation of the hat and books is characterized by the verticality postpositions *uḍiŋ* 'above' and *yaṭiŋ* 'below', which cannot be replaced by *idiripiṭə* 'in front of' and *piṭipasse* 'in back of'.

(7) toppiyə=ṭə **yaḍiŋ** pot   tiye-nəwa
     hat.DEF=DAT **below** book.PL be.INAN-NPST
     'There are books below the hat.'

(8) pot   wələṭə **uḍiŋ**   toppiy-ak tiye-nəwa
     book.PL DAT.PL **above** hat-IND   be.INAN-NPST
     'There is a hat above the book.'

By the same token, there are also situations where the frontal axis is the only apt choice. In the situation illustrated in Figure 4, the consultant is standing at the entrance of the department library, and is describing the spatial arrangement of a table and some bookshelves near the table.
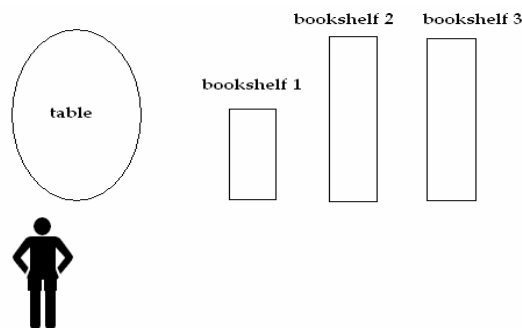
FIGURE 4.

The consultant starts with using left-right axis to locate the table and the first bookshelf:

(9) *meese=ṭə*   *dakunu pætte   poḍi   pot raakəyak   tiye-nəwa*
     table=DAT.DEF right   side-LOC small bookshelf=IND be.INAN-NPST
     'To the right side of the table, there is a small bookshelf'

But afterwards, the consultant takes the books as the Ground, and switches to the frontal axis to locate bookshelves 2 and 3. What should be noted here is that bookshelves 2 and 3 are taller than the consultant, and for this case the frontal postpositions are not interchangeable with the vertical postpositions.

(10) *ee   potraake=ṭə   **pitipasse** tawə   loku   potraakek   ui   eekə*
    that   bookshelf=DAT   **behind**   other   big   bookshelf-IND   and   that

    ***pitipasse** tawat   loku   potraakek-ut   tiye-nəwa*
    behind   another-also   big   bookshelf.IND-also   be.INAN-NPST
    'Behind the bookshelf, there is another big bookshelf also. And behind that, there is also another big bookshelf.'

It seems that the uses of the frontal axis and its interchangeability with the vertical axis are determined and restricted by two primary factors. One is the semantic nature of the frontal axis, and the other is the perspective the speaker adopts in spatial description.

Langacker (1999), in the following quote illuminates the most crucial property that distinguishes the frontal axis from other horizontal axes such as left/right:

> Consider the semantic opposition between in front of and in back of... [The diagram for each] profiles the relation involving two focal participants wherein one participant stands in the line of sight... between a viewer and the other participant... In front of takes the far participant as a landmark for purpose of locating the near participant, whereas in back of reverses those roles. (Langacker 1999:8)

Thus, for both 6 and 10, it is reasonable to say that the frontal axis is applied because both focal participants (the Figure and the Ground) and the speaker can be considered to be standing on the same line, with the 'front' one to stand in the line of sight between the speaker and the 'back' one. In fact, the spatial relation and objects described in 10 even highlights one of the frontal axis's properties that involves 'the line of sight', as bookshelf 2 are taller and bigger in size than the consultant, thus visually blocks the bookshelf (bookshelf 3) behind it from the speaker.

Perspective, on the other hand, has to do with the position from which things are viewed (Langacker 1983:123). In spatial descriptions, perspectives speakers can take are primarily categorized into two kinds. Here I would like to follow Taylor and Tversky (1996) in naming them respectively ROUTE PERSPECTIVE, in which landmarks are described with respect to a viewer moving through or situated right within the depicted space; and SURVEY PERSPECTIVE, wherein landmarks are described with respect to each other as if viewed from above. It is easy to see why the image-schema operation (Vertical → Horizontal) is only applicable when the speaker adopts survey perspective. For one thing, being in this perspective, the speaker can profile arrangement using the canonical vertical view of a map, taking objects near the reference point (the speaker) as being below, and objects further away as being above.

Nonetheless, if the size and height of a focal participant create a visual blockage between the reference point (the speaker) and the other focal participant, the survey view cannot be activated since the speaker cannot have an overall view of the object arrangement. That is to say, once a focal participant blocks the view of the reference point to the other focal

participant, the switching from route perspective to survey perspective is inhibited, and in turn the transformation from vertical to horizontal is also blocked.

An attestation to this hypothesis involving perspective, schematic transformation and the semantic nature of the frontal axis came about in the results of another mini-project. In the beginning of the project, the consultant was asked to describe the spatial setting of the buildings shown in a fraction of a campus map. Again, here we have a situation in which the vertical and frontal axes are interchangeable, as survey perspective is a very natural choice, and schematic manipulation (Vertical → Horizontal) is thus possible.
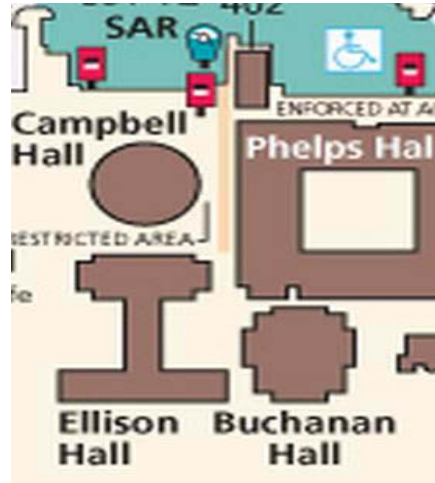
FIGURE 5.

(11) *kembelhol*   *wələṭə* **yaṭiŋ/idiripiṭi**   *elisənhol*   *tiye-nəwa*
    Campbell.Hall   DAT.PL   **below/in.front.of**   Ellison.Hall   be.INAN-NPST
    'Below Campbell Hall is Ellison Hall.'

(12) bjukənəhol   *wələṭə*   **udiŋ/piṭipasse**   *we-nnə*   felpshol   *tiye-nəwa*
    Buchanan.Hall   DAT.PL   **above/in.back.of**   be-INF   Phelps.Hall   be.INAN-NPST
    'Towards the top of Buchanan Hall is Phelps Hall.'

Then, the consultant was asked to imagine that he is standing right in front of Ellison Hall, and is asked to tell where Campbell Hall is with respect to Ellison Hall. Not surprisingly, the frontal axis is the only apt choice, for Ellison Hall can block the speaker's view to Campbell Hall. The speaker cannot have an overall view of the environment where he is situated, thus can only stick to route perspective to locate the Figure (Campbell Hall).

(13) kembəlhol   tiyenn-e   elisonhol   *wələṭə*   **piṭipasse**
    Campbell.Hall   be.INAN-FOC   Ellison.Hall   DAT.PL   **behind**
    'Campbell Hall is behind Ellison Hall.'

4. CONCLUSION. This paper accounts and analyzes a case in which Sinhala can use verticality terms to code horizontal (i.e. frontal) relation in spatial descriptions. The analysis is, on the one hand, based on the image-schema operation (Principle A: Vertical → Horizontal) proposed by Ekberg (1997). In light of such an operation, one can explain why 'in back of' can correspond to 'above', and 'in front of' can correspond to 'below', given that the reference point should be

at the ground level of a vertical axis directed upward, and being in the back is further away from the reference point than being in the front.

In explaining why human beings would prefer a vertical expression over a non-vertical one when the spatial relation in point is non-vertical, Ekberg (1999), based on evidence on language acquisition and human cognition, proposed that it is because the vertical axis is more natural, more salient, and thus easier to perceive. The facts observed in Sinhala, however, do not seem to go with this explanation. Rather, it could be the speaker's switching from route perspective to survey perspective that triggers him to adopt the canonical vertical view of a map, and in turn makes the schematic operation (Vertical Horizontal) possible. In other words, it seems in Sinhala the uses of and interchangeability between the vertical and frontal axes are determined by the semantic nature of the frontal relation, the perspective taken in depiction, as well as the applicability of the image-schema operation with respect to the adopted perspective.

<div align="center">REFERENCES</div>

BROWN, PENELOPE, and STEPHEN C. LEVINSON. 2000. Frames of spatial reference and their acquisition in Tenejapan Tzeltal. *Culture, thought, and development*, ed. by L. Nucci, G. Saxe, and E. Turiel, 167-97. Mahwah, NJ: Erlbaum.

CIENKI, ALAN. 1995. Properties of image schemas as a category/Properties of categories of image schemas. *Lexical and syntactical constructions and the construction of meaning: Proceeding of the bi-annual ICLA meeting in Albuquerque, July 1995*, ed. by Marjolijin Verspoor, Kee Dong Lee and Eve Sweetser, 69-88. Amsterdam and Philadelphia: Benjamins.

EKBERG, LENA. 1997. The mental manipulation of the vertical axis: how to go from "up" to "out", or from "above" to "behind". *Lexical and syntactical constructions and the construction of meaning: Proceeding of the bi-annual ICLA meeting in Albuquerque, July 1995*, ed. by Marjolijin Verspoor, Kee Dong Lee and Eve Sweetser, 69-88. Amsterdam, Philadelphia: Benjamins.

FILLMORE, CHARLES J. 1982. Towards a descriptive framework for spatial deixis. *Speech, place and action*, ed. by Robert J. Jarvella and Wolfgang Klein, 31-59. New York: John Wiley & Sons Ltd.

HAYWARD, WILLIAM G, and MICHAEL J. TARR. 1995. Spatial language and spatial representation. *Cognition* 55.1.39-84.

HILL, CLIFFORD A. 1975. Variation in the use of 'front' and 'back' by bilingual speakers. *Proceedings of the first annual meeting of the Berkeley Linguistic Society*, ed. by C. Cogen, H. Thompson and J. Wright. Berkeley: University of California.

JOHNSON, MARK. 1987. *The Body in the Mind: The Bodily Basis of meaning, Imagination, and Reason.* Chicago and London: The University of Chicago Press.

LAKOFF, GEORGE, and MARK JOHNSON. 1980. *Metaphors we live by.* Chicago and London: The University of Chicago press.

LANGACKER, RONALD W. 1987. *Foundations of cognitive Grammar. Vol. 1. Theoretical prerequisites.* Stanford, CA: Stanford University Press.

LANGACKER, RONALD W. 1991. *Concept, image, and symbol. The cognitive basis of grammar.* Berlin: Mouton de Gruyter.

LANGACKER, RONALD W. 1999. *Grammar and conceptualization.* Berlin: Mouton de Gruyter.

LEVINSON, STEPHEN C. 1991. Relativity in spatial conception and description. *Rethinking linguistic relativity*, ed. by John J. Gumperz and Stephen C. Levinson, 177-202. Cambridge: Cambridge University Press.

TAYLOR, HOLLY A., SUSAN J. NAYLOR, and NICHOLAS A. CHECHILE. 1999. Goal-specific influences on the representation of spatial perspective. *Memory and cognition* 27.2.309-19.

TAYLOR, HOLLY A., and BARBARA TVERSKY. 1996. Perspective in spatial descriptions. *Journal of memory and language* 35.371-91.

THORNDYKE, PERRY W., and BARBARA HAYES-ROTH. 1982. Differences in spatial knowledge acquired from maps and navigation. *Cognitive psychology* 14.560-89.

Department of Linguistics
University of California, Santa Barbara
3607 South Hall
Santa Barbara, CA 93106-3100
rgyalrong@gmail.com

DIRECT AND INDIRECT CAUSATION IN SINHALA: EXAMINING THE COMPLEXITY CONTINUUM[1]

DANIELLE MATHIEU-REEVES

Rice University

1. INTRODUCTION. The *causative*, or the causative clause, is the linguistic expression of *causation.* Causation is a somewhat abstract concept in which the occurrence of one event results in the occurrence of a separate event. It is not always evident how one can 'cause' another; it is an intangible connection between events: a series of events alone does not necessarily connote a relationship of causation. A Causative can be defined as:

> A linguistic expression that contains in semantic/logical structure a predicate of cause, one argument of which is a predicate expressing an effect (Payne 1997:176)

Because causation is not concrete, it is expected that there are different structures for causative expressions. Causative expressions come in several forms: lexical, morphological, and analytical, but each have common parts:

> -Causer - agent of predicate of cause
> -Causee - agent of caused event
> -Caused event - resultative action, i.e. 'the effect'

A lexical causative is an expression in which the caused event is part of the semantics of the verb, such as "kill" meaning "cause to die." In these expressions, the causer and the causee will be the same entity. A morphological causative is an expression with a productive change in verb form, such as an inflection that changes 'go' to 'send.'[2] This is a valence increasing operation so that an intransitive verb becomes transitive and a transitive verb becomes ditransitive. Thus an additional party, a causee or a causer, is added to the expression. An analytical causative is one in which there is a separate causative verb and which are not *de facto* valence increasing operations. Sinhala uses the latter two of these methods to express causation.

2. ICONICITY IN CAUSATIVES. Iconicity in causatives is namely concerned with the correspondence between linguistic distance and conceptual distance. Conceptual distance is directness or indirectness of a predicate of cause and a predicate of effect. Linguistic distance is the distance between two grammatical structures.

> The linguistic distance between [two structures] is least when they are fused in [the same morpheme]; greater when they are distinct but bound morphemes; and still greater when they are separate words. The linguistic distance between them is greatest of all when they are separated by one or more other words. (Haiman 1983:782)

---

[1] I offer my sincere thanks to Nissanka Wickremasinghe for all of his assistance, time and patience as well as Dr. Robert Englebretson for his guidance. Any errors in examples or analysis are mine alone.
[2] Such as example (2)

Thus the connection between linguistic and conceptual distance is:

   a. The linguistic distance between expressions corresponds to the conceptual
      distance between them
   b. The linguistic separateness of an expression corresponds to the conceptual
      independence of the object or event which it represents.
   (Haiman 1983:782-3)

Direct causation has a conceptual distance between cause and result that is lesser than indirect causation and should therefore have a causative expression with lesser linguistic distance than that of indirect causation.

Indirect causation has a conceptual distance between cause and result that is greater than direct causation and should therefore have a causative expression with greater linguistic distance than that of direct causation. Thus Haiman (1983) posits:

> If two causatives contrast within a given language, such that they correspond to structures given…, and they contrast semantically with respect to the conceptual distance between cause and result, then the conceptual distance between cause and result will correspond to the formal distance between cause and result. (783)

This is indeed the case for Sinhala. I will argue that there is a continuum of directness which matches Haiman's theory. The two types of causative expression, morphological and analytical, correspond to expressions of lesser and greater linguistic distance respectively.

   3. Morphological Causative. In Sinhala, the morphological causative is a productive change in verb form. Because the causative form is distinguished by a bound morpheme, on Haiman's continuum, it should represent more direct forms of causation.

   3.1. Unmarked morphological causative. The simple morphological causative is the most direct type of causative conceptually and is the form with the least linguistic distance. That is to say, this is the most unmarked causative clause type. With the causative verb form, valency is increased by one, making intransitive clauses transitive and transitive clauses ditransitive. The form is a simple sentence: nominatives are unmarked and animate accusatives are marked with –*wə*.

   Causative clause:
   (1)   mamə oyawə duwəwənəwa
         1SG   2SG-ACC run.NPST.CAUS
         'I make you run.'

In Sinhala, either participant in a caused event can be animate or inanimate. The only restriction is that inanimate causers must use the non-volitional verb in the past tense. There are no animacy restrictions for any of the causative clause types. This leads to clauses such as (2) where the morphological causative is used with an inanimate causer. This is a possible

exception to the tendency for the morphological causative to be used with very direct causation. Inanimate causers can provoke questions about indirectness behind the event, which would make the clause conceptionally less direct. However, because inanimate agents are non-prototypical agents, some non-prototypicality is to be expected. This does not result in a complete dismissal of Haiman's continuum in Sinhala causatives.

    (2)   gaha ohuwə ispiritaleṭə yawənəwa
           tree 3sg-acc hospital-dat go.npst.caus
           'The tree is sending him to the hospital.'[3]

Whereas English has lexical causatives, such as 'send,' the Sinhala language does not use this form of causative clause. Therefore when our consultant translated these words, a causative inflection of the verb is used.

    (3)   mamə oyawə yawənəwa
           1sg   2sg-acc go.npst.caus
           'I make you go.' or 'I send you.'

Verbs that are transitive become ditransative when the causative inflection is used.

    Ditransitive causative:
    (4)   amma   laməyaṭə bat kawənəwa
           mother child-dat rice eat.npst.caus
           'The mother feeds the child rice.'

Past tense causative forms have two different inflections, one for volitional and one for non-volitional forms. The former (5) connotes compliance or willfulness on the part of the causee. The latter (6) connotes resistance on the part of the causee, and some force, whether physical or verbal manipulation, on the part of the causer.

    (5)   mamə oyawə kadeeṭə yauwa
           1sg   2sg-acc store-dat go.pst.vol
           'I sent you to the store.'
    (6)   mamə oyawə kadeeṭə yauna
           1sg   2sg-acc store-dat go.pst.nvol
           'I sent you to the store.'

    3.2. MORPHOLOGICAL CAUSATIVES WITH POST-POSITION *lauwa* 'THROUGH'. In addition to clauses with a causative inflection of the verb, slightly less direct causation is accomplished with a causative inflection in addition to a post-position. In this type the causee is an oblique rather than the patient. This type of clause connotes that the causer uses the causee as an instrument to accomplish the caused event. The causation is direct because the causer is the agent, but less direct than the unmarked clause because some action must occur for the causee to be an instrument or channel for the caused event.

---

[3] This paper includes all combinations of causer-causee and animacy : animate-animate (1), inanimate-animate (2), animate-inanimate (14) and inanimate-inanimate (15).

(7)   mamə oya lauwa   redi   hodəwənəwa
      1SG   2SG through clothes wash.NPST.CAUS
      'I am using you to wash the clothes.'

(8)   mamə oya lauwa   sindu kiyəwənəwa
      1SG   2SG through song   say.NPST.CAUS
      'I am using you to sing.'

3.3. MORPHOLOGICAL CAUSATIVES WITH THE COMPLEMENTIZER *kiyəla* 'SAY'. The causative inflection of the verb can also be used with the *kiyəla* converb, which has grammaticized strongly into a complementizer. It is also possible to interpret this clause type as a converb with two events taking place: The causer is telling the causee something *and* a caused event happens. If this is the case, then the clause type is of greater linguistic distance as well as greater conceptual distance. If *kiyla* is to be interpreted as a converb, then the causation would be analytical rather than morphological. However, the other types of analytical causatives in Sinhala are not used with the causative inflection of the verb. Additionally, *kiyla* is used in several clause types as a complementizer in complement clauses, not simply in causative ones. Therefore, it is more likely that *kiyla* functions as a complementizer here, however both explanations are possible.

As with other clauses with converbs in Sinhala, the agent of the second action, here the caused event, can be unspecified. It is implied that the agent of the second event is the patient of the first event. In example (9) it is stated that 'you' is told something and it is implied that 'you' is made to wash the clothes. However, the agent of the caused event could be an additional unstated party, depending on context.

(9)   mamə oyaṭə   kiyəla   redi   hodəwənəwa
      1SG   2SG-DAT say.CONV clothes wash.NPST.CAUS
      'I am making you wash the clothes.' lit. 'I cause that you wash the clothes.'

To specify the causee, it can be directly stated after the converb/complementizer:

(10) nuwan maṭə   kiyəla   mawə   prəmadə kərəwənəwa
      nuwan 1SG-DAT say.CONV 1SG-ACC late   do.NPST.CAUS
      'Nuwan is making me late.' lit. 'Nuwan causes that I am late.'

4. ANALYTICAL CAUSATIVE. Analytical causatives usually involve a separate causative verb and are not *de facto* valence increasing operations. In Sinhala, as in many other languages, the caused event verb is in the infinitive.

4.1. ANALYTICAL CAUSATIVES WITH *denəwa* 'DIVE' AND *œrinəwa* 'OPEN'. These verb forms imply indirect causation: the causer lets, or does not prevent, the causee from accomplishing the caused event.

denəwa - According to the consultant, this expression is used when causer allows causee to accomplish the caused event

ærinəwa - According to the consultant, in this expression the causer retains less agency (which would make the causation less direct). This would be used when the causee accomplishes caused event through "carelessness" of the causer.

The consultant translates both forms as 'let' in English. These are more like permissive clauses than causative clauses, but it is one of the main ways in Sinhala to express indirect causation. There is greater linguistic distance here because of the compound verb form.

(11) mamə oyaṭə   sindu kiyannə denəwa
      1SG   2SG-DAT sing   say.INF   give.NPST
      'I am letting you sing.'
(12) mamə kukulawə   marannə   ærinəwa
      1SG   chicken-ACC die.INF   open.NPST
      'I am letting the chicken die.'

4.2. ANALYTICAL CAUSATIVES WITH *kriya kərənəwa* 'ACT IN SUCH A WAY THAT'. This is used for the most indirect causation in Sinhala. The linguistic distance here is greater because it is a three word verb compound, the caused event verb, and *kriya kərənəwa*. The conceptual distance for these clauses is also fairly great because it involves two caused events, one implied (action A) the other expressed (action B). Action A occurs, which leads to action B being caused. Action B is the caused event that is expressed, action A is unstated. However, there must be an unexpressed action A for this clause type to be chosen by the speaker over one of the morphological causative clause types.

(13) mamə kukulawə   marann   kriya kərənəwa
                                      ə
      1SG   chicken-ACC die.INF   act   do.NPST
      'I am causing the chicken to die.' or 'I am acting is such a way that the chicken is dying.'
(14) ohu gaha perəlennə kriya keruwa

      3SG tree topple.INF act   do.PST
      'He is causing the tree to topple' or 'He is acting in such a way that the tree is toppling.'
(15) hulaŋgə ge   wæṭenn   kriya kərənəwa
                                    ə
      wind   house fall.INF   act   do.NPST
      'The wind is causing the tree to fall.'

5. CASE STUDY WITH *kata kərənəwa* 'TALK'. This section shows the same verb, tense and participants in each causative expression, as well as a context that would have caused a particular expression to be used, rather than one of the others.

Clause Type: unmarked morphological causative
(16) mamə oyawə kata kərəwənəwa
      1SG   2SG-ACC talk do.NPST.CAUS

'I make you talk.'

Possible context: This, being the most basic causative, could be used in almost any context, as long as it wasn't indirect, or the speaker does not want to highlight the indirectness. For example, I poke you with a stick so you say 'ow'.

Clause Type: morphological causative with *lauwa* 'through'
(17) mamə oya lauwa   kata kərəwənəwa
      1SG   2SG through talk do.NPST.CAUS
      'I am using you to talk.'
Possible context: I have a mouthpiece (you) and I tell you to address the public on my behalf.

Clause Type: morphological causative with *kiyəla* 'say'
(18) mamə oyaṭə   kiyəla   kata kərəwənəwa
      1SG   2SG-DAT say.CONV talk do.NPST.CAUS
      'I cause that you talk.'
Possible context: I order you to talk and so you are doing it.

Clause Type: analytical causative with *denəwa* 'give'
(19) mamə oyaṭə   kata kərannə denəwa
      1SG   2SG-DAT talk do.INF   give.NPST
      'I am letting you talk.'
Possible context: I give you permission and because of that permission, you are talking.

Clause Type: analytical causative with *ærinəwa* 'open'
(20) mamə oyaṭə   kata kərannə ærinəwa
      1SG   2SG-DAT talk do.INF   open.NPST
      'I am letting you talk.'
Possible context: Although you would not normally talk, I am not paying attention to you, so you go ahead and talk.

Clause Type: analytical causative with *kriya kərənəwa* 'act in such a way'
(21) mamə oyaṭə   kata kərannə kriya kərənəwa
      1SG   2SG-DAT talk do.INF   act   do.NPST
      'I act in such a way that you talk.'
Possible context: I behave very rudely at a dinner so that eventually you have to say something to make me stop my poor behavior.

6. SUMMARY. In Sinhala, Haiman's iconic motivation fits with the continuum of direct-indirect causation. The most conceptually direct causation is accomplished with the least linguistically distant clause types. The least conceptually direct causation is accomplished with the most linguistically distant clause types.

Most Direct					Least Direct
-wə-	lauwa	kiyəla	denəwa	ærinəwa	kriya kərənəwa

Causative clauses in Sinhala are a confirmation of Haiman's theory of iconic motivation, and support the idea of motivated language. The distribution of causative constructions is not random or by chance. Specific conceptual events influence which form will be used over others. This leads to further research questions such as what other construction types in Sinhala are a product of iconic and economic motivation? Additionally, how many and what other languages demonstrate iconic or economic motivation in causative constructions?

REFERENCES

Gair, James W. and John C. Paolillo. 1997. *Sinhala.* München: Lincom Europa.
Haiman, John. 1983. Iconic and Economic Motivation. *Language* 59.781-819.
Payne, Thomas E. 1997. *Describing Morphosyntax.* New York: Cambridge University Press.

Rice University
Department of Linguistics, MS 23
6100 Main Street
Houston, TX 77005-1892

A PHONETIC AND PHONOLOGICAL APPROACH TO STRESS IN SINHALA VERBS

CARLOS M NASH
University of California, Santa Barbara

1. INTRODUCTION.

1.1. REVIEW OF PHONETIC STRESS. According to Cutler (2005) research on stress and stress perception has primarily focused on the acoustic characteristics of stressed versus unstressed syllables, and how listeners make use of acoustic cues to make judgements as regards the occurrence of stress. Most phoneticians agree that the three acoustic dimensions involved in the realisation of stress are duration, fundamental frequency, and intensity. These acoustic properties correspond to the perceptual phenomena of length, pitch, and loudness, respectively. Some phoneticians also include vowel quality as an additional dimension (Laver 1994, Hayward 2000). In general, stress is described as the display of prominence by the exaggeration of one or more of the phonetic parameters on certain syllables when contrasted with others (Laver 1994). Hence, a syllable displaying such prominence can be said to have possibly longer duration, higher pitch, greater acoustic intensity, and more carefully articulated phones when in contrast with unstressed syllables (Hayward 2000). However, some linguists make more specific claims as to which parameters play a larger role in the realisation of stress. Ladefoged (2003) states it is likely to be some combination of pitch, length, and loudness, with the first two playing the greatest role.

To further complicate the phonetic description of stress, it appears that its phonetic manifestation varies from language to language (Laver 1994). Based on many studies conducted on English stress, Laver claims that English exploits all four parameters. However, studies conducted by Fry (1955) claim that English relies heavily on duration and intensity. In another study conducted by Mol and Uhlenbeck in 1965 in which the intensity relationship of stressed and unstressed syllable was reversed, it shows that intensity did not affect the perception of stress (Cutler 2005). On the other hand, Fujimura and Erickson (1997) state that stress is primarily manifested in the change of the fundamental frequency. In summary, Cutler asserts that the least controversial findings as of today are (1) syllables are perceived to be stressed if they exhibit fundamental frequency excursion and (2) greater syllable duration is likewise associated with perceived stress. The more controversial finding is that intensity manipulations only weakly affect stress perception, such as the findings shown by Mol and Uhlenbeck.

As for the muscular activities involved in the production of stress, Hayward (2000) says that stressed syllables display raised subglottal pressure when compared with unstressed syllables. This increase in pressure can be accomplished by respiratory activity, by laryngeal activity, or a combination of both. Another possibility is that respiratory and laryngeal muscles are involved in the production of stress (Hayward 2000, Ladefoged 2001). Ladefoged (2001) claims that more air is pushed out of the lungs in the production of stressed syllables. Succinctly, it can be stated that a stressed syllable may have greater respiratory energy than neighbouring unstressed syllables.

Although acoustic parameters and muscular activities have been identified, it is still a difficult task to objectively measure stress. Ladefoged offers a highly impressionistic approach to identifying stressed syllables. He suggests that the best way to decide whether a syllable is

stressed is to try to tap out the beat as a word is being produced. The claim is that it is easier to produce an increased tap exactly in time with an existing increase of respiratory or laryngeal activity (Ladefoged 2001). Here Ladefoged is again correlating increased muscular activity with stressed syllables. This is apparent in his hypothesis which states that listeners perceive stress by putting together all of the cues available in order to deduce the motor activity a speaker would use to produce the same stresses. In general, stress has proven to be both the most straightforward to analyse impressionistically and the most difficult to define in purely phonetic terms (Hayward 2000).

1.2. REVIEW OF PHONOLOGICAL STRESS. Regarding stress from a phonological perspective, stress makes up part of the metrical organisation of speech. According to Kager (1999), there are conflicting forces at work in lexical stress: rhythm, quantity-sensitivity, and edge-marking. Rhythm is the pressure toward a regularly alternating distribution of weak (unstressed) and strong (stressed) syllables. Quantity-sensitivity is the pressure to match syllable weight to prominence. Edge-marking is the pressure to mark the edges of morphemes. Languages that make linguistic use of stress can be divided into two categories: fixed lexical and variable lexical stress. Laver (1994), citing Hyman's typological analysis of stress, states that the majority of fixed lexical stress languages demonstrate a preference for stress toward the initial or final syllable of a word, which points to an edge-marking, or demarcative, function. Variable lexical stress languages all demonstrate a range of different locations of lexical stress. This pattern is commonly found in languages where lexical stress makes a contrast in meaning. Hyman also reported that there were a few languages with stress that is governed by syllable weight (Laver 1994). From a phonological standpoint, stress can be regarded as a binary distinction. A two-way distinction can be drawn (i.e. stressed and unstressed) (Laver 1994). This differs from a phonetic analysis of stress, which is regarded as a gradient phenomenon. The phonetic realisation of any syllable can be said to show a greater or lesser degree of stress relative to the manifestation of some other syllable. However, some linguists make more distinctions in the levels of stress in their phonological analyses, which poses additional difficulties. As Ladefoged (2001) points out these linguists, in actuality, are conflating stress and tonic accent into 'levels of stress'.

1.3. OVERVIEW OF THIS PAPER. This paper will focus on stress in Sinhala verbs. Sinhala is an Indo-Aryan language and is spoken predominantly in the island nation of Sri Lanka. For a majority of present day Indo-Aryan languages, stress is not contrastive (Masica 1991). There are a few exceptional cases; Assamese has phonemically contrastive stress, and Siraiki has a few cases of lexically contrastive stress (Masica 1991). For a significant number of Indo-Aryan languages, stress is predicted by using a complicated set of rules, usually involving the number of syllables, whether they are open or closed, and the nature of their vowels (Masica 1991). Sinhala and Nepali, according to Masica, demonstrate a tendency for a weak word-initial stress. Letterman's 1997 dissertation on Sinhala, based on impressionistic analysis, makes five observations as regards Sinhala stress: (1) syllable weight has a role in determining stress, (2) primary stress tends to fall on the initial or peninitial syllable, whichever is heavier, (3) parsing appears to be left to right, (4) word-final stress is found when read as a list due to final glottal stop or phonetic lengthening, (5) some heavy syllables carry secondary and not primary stress. Although Letterman's findings are somewhat similar to those presented this paper, she does not attempt to systematically verify her findings using objective measurements or native

speaker intuition. The reader is left to judge her phonological account of Sinhala stress solely based on unsystematic observations.

Using instrumental analysis, this paper will attempt to derive an objective measure of stress in Sinhala verbs (Section 2). The measure and its predictions will be checked against native speaker intuition. It will be shown that higher intensity and longer duration, combined, have a high degree of correlation with stressed syllables. Once an objective measure has been developed, the verbs will be analysed for any display of predictable stress (Section 3). In this section, it will be shown that Sinhala does have a preference for stress word-initially (if the initial syllable has an onset) and on heavy syllables. This pattern will then be accounted for using an Optimality Theory approach.

2. ACOUSTIC MEASUREMENTS.

2.1. OVERVIEW OF SINHALA CONSONANTS. Sinhala presents a few challenges regarding acoustic measurements. Before measuring syllables for acoustic correlates of stress, it is important to devise a consistent method of measuring phonetic segments. In an effort to highlight the problematic segments, a brief overview of Sinhala consonants. Table 1 presents the consonant inventory of Sinhala.

	BILABIAL	ALVEOLAR	POST-ALVEOLAR	RETROFLEX	PALATAL	VELAR	GLOTTAL
STOP	p b	t d		ṭ ḍ		k g	
• BREATHY	bh	dh		ḍh		gh	
• PRENASALISED	m̆b	n̆d		n̆ḍ		ŋ̆g	
NASAL	m	n			ñ	ŋ	
TRILL		r					
FRICATIVE		s	š				h
AFFRICATE			c j				
APPROXIMANT	w			y			
LATERAL APPROXIMANT		l					

TABLE 1. Consonant inventory of Sinhala.

Of the consonants shown in Table 1, prenasalised stops and their status as a single unit need to be addressed. There is morphophonological evidence for the treatment of prenasalised stops as a single unit. In a subset of nouns, the plural form appears to be derived from the singular form by deleting the final schwa. For example, consider *li.n̆də* 'well.SG', *a.ŋ̆gə* 'horn.SG', and *hu.lə.ŋ̆gə* 'wind.SG'. The plural forms for these nouns are *liŋ* 'well.PL', *aŋ* 'horn.PL', and *hu.laŋ* 'wind.PL'. The alternation observed here between prenasalised stops and velar nasal is similar to other nasal/velar nasal alternations seen in other nouns, such as *ta.na/taŋ* 'breast/breasts', *ha.ma/haŋ* 'skin/skins', and *ka.na/ kaŋ* 'ear/ears'. Phonetic evidence is provided by the analysis of gemination. Certain verbs display gemination in the past tense, such as *a.din.nə* 'to pull' and *æd.da* 'pull.PAST'. The verb *a.n̆din.nə* 'to draw' follows a similar pattern except that when it geminates, the nasal portion of the prenasalised stop migrates to the previous syllable and becomes a coda nasal, as in *æn.da* 'draw.PAST'. Geminate affricates behave in a similar fashion, for example the noun *puc.can.nə* 'child' is phonetically [put.can.nə]. The observations above have implications as regards the distribution of prenasalised stops. It appears that prenasalised stops can only occur in onset position. Furthermore, it can be said that prenasalised stops only occur in word-medial position since there has been no evidence of word-initial prenasalised stops.

The overview of prenasalised stops brings up another challenge, that is the measurement of geminates. It is important to understand how geminates behave within the syllable. For the most part, through impressionistic and phonological analyses, it was determined that geminates are heterosyllabic. Further evidence is provided by analysing the occurrences of geminates in Sinhala. All instances of geminates so far encountered are intervocalic (e.g. *jiivatvennə* 'to live', *hiṭəgatta* 'stand.PAST', *bædde* 'fry.FOCUSED PAST'). Ladefoged and Maddieson (1996) state that geminate stops in many languages are limited to word-medial position where they usually close the preceding syllable, as well as serving as the onset of the following syllable. They also note that the vowel preceding the geminate is usually shortened to some degree. This pattern can be seen in Sinhala. For the particular consultant and tempo recorded in this study, the average length of a vowel in a non-final open syllable is 89.7 ms. The average length of a vowel preceding a geminate is about 55.6 ms. It has been well observed that coda consonants do affect the length of preceding vowels; therefore, one can conclude that a portion of the geminate is occupying the coda position of a syllable.

However, there still is the matter of proving that a portion of the geminate is the onset of the following syllable. A phonological analysis of Sinhala shows that the language does not contain complex codas. Therefore, on grounds of phonology, one is forced to split geminates across syllables. One can also demonstrate this based on phonetic evidence. Using observations from Ladefoged and Maddieson (1996), languages with a distinction of consonant length have only two distinctive lengths. Furthermore, they cite Lehiste (1966) and Eek (1984-5) whose analyses of Estonian show a third length is created by lengthening of long consonants in stressed syllables. Keeping these points in mind, the duration of stop closures in a subset of tokens were measured. It was readily noted that there were two slightly different lengths. The average length of an onset in a word-medial syllable is about 65.3 ms. However, the average length of an onset in word-initial position is approximately 102.5 ms. This difference is interesting considering that many researchers have observed some degree of stress on word-initial syllables. This 1:1.6 difference seems to be relevant to geminates as shown below in the analysis of geminate **dd**.

During the acoustic analysis of geminates, a peculiar pattern has been observed with geminate **dd**. In many of the intensity diagrams of geminate dd, there is a perturbation during the closure phase of the stop. A clear demonstration of this perturbation can be seen in Figure 1 below.
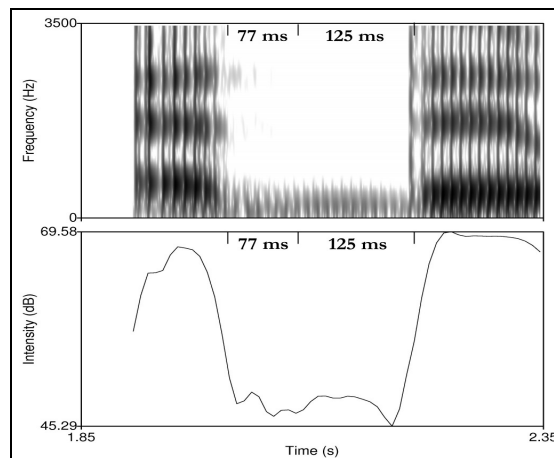
FIGURE 1. Spectrogram and intensity diagram demonstrating intensity perturbation in geminate dd.

The perturbation seen in the figure above is relatively simple compared to those seen in other tokens. In general, there is a sharp fall of intensity at the beginning of the geminate. Then there is a short rise of intensity slightly less than midway through the closure. The intensity falls again and then rises sharply during the release of the stop. At this stage of the investigation, it is uncertain as to the cause to the intensity perturbation; however, many of the intensity diagrams demonstrate this pattern. In Figure 1, the short rise of intensity begins 77 ms into the stop. The burst of the stop occurs 125 ms later. This is roughly a 1:1.6 difference, similar to the difference seen with the duration of stops in word-medial and word-initial positions.[1] If it is assumed that the syllable boundary occurs at the perturbation, then it follows that the onset is lengthened. This lengthened onset possibly corresponds to the judgements of a few observers who claim that the release of geminate stops seem to be 'stronger' or 'more prominent' than the release of single stops. As a result, based primarily on the measurements of single stops and secondarily on the observation of intensity perturbation in geminate **dd**, geminates will be split into codas and onsets with a ratio of 1:1.6.

2.2. EXPERIMENT I: MEASURING SYLLABLE INTENSITY AND PITCH. In the initial analysis of stress in Sinhala, many students in the field methods class described stressed syllables as sounding 'louder'. Therefore, the first experiment devised was to measure the intensity of each syllable and to compare the results to the judgement of the consultant and other researchers. First, a list of verbs with various syllable structures was compiled. The verbs were then recorded in frames on digital audio tape. The tapes were then converted into wave files for analysis. After converting the recording into tokens, the next step was to record native speaker judgement of stress in order to assess the degree of correlation for each experiment. In a one-hour elicitation session, the consultant was asked to listen to the sentences he had recorded one week earlier. He was then instructed to identify the verb and indicate which part of the verb sounded 'more prominent'. He was specifically instructed to identify stressed syllables in this manner so that he would not be influenced by words such as 'higher', 'louder', or 'longer' which could cause one to focus on a particular feature. He was also instructed that he may choose more than one syllable if he believes there is more than one syllable that are prominent. During the elicitation, it was noted that the consultant would bounce his index finger in the air while articulating the verbs. Ladefoged (2001) suggests tapping out the beat as the word is being produced as a useful test in determining stress. He claims that it is easier to produce an increased tap exactly in time with an existing increasing in activity (be it respiratory or laryngeal). During the elicitation session, the consultant assessed eighty verbs. The tables in this and the following sections show in bold print the syllables that he identified as being 'prominent'.

For each verb, the highest intensity for each syllable was recorded. Table 2, below, illustrates a few tokens with the highest intensity achieved for each syllable.

---

[1] A similar perturbation appears when observing other voiced continuant geminates such as /ll/ and /mm/; however, the ratio is nearer to 1: 1.4.

	Highest Intensity (in dB)					
Target	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6
ka.ḍə.nə.wa	69.07	65.68	59.96	63.93		
kæ.ḍu.wa	64.71	58.48	65.39			
ka.ḍə.wə.nə.wa	62.88	64.01	59.72	57.36	57.73	
kæ.**ḍew**.wa	64.16	61.82	56.50			
a.**di**.nə.wa	60.83	59.95	56.68	61.65		
æd.**da**	56.09	58.17				
a.**din**.ne	57.23	56.56	54.94			
æd.**de**	57.65	55.18				
a.**dii**	59.00	57.19				
ad.**də**.wə.nə.wa	54.23	62.43	57.35	55.39	54.62	
i.ňdə.**gan**.na.nə.wa	65.64	67.14	71.70	64.03	62.96	65.19
i.ňdə.**gan**.nə.wai	61.29	66.28	68.87	64.93	65.67	
i.ňdə.**gan.nan.naŋ**	60.28	64.86	66.24	60.27	62.30	
i.**ňdə**.gæ.nə	59.74	70.19	66.03	64.81		

TABLE 2. Selection of results from intensity measurements. [a]

[a] Highest values are shaded in grey. Syllables in bold print were identified by the consultant as being prominent.

In Table 2, the syllables in bold correspond to syllables judged to be stressed by the consultant, and the values in the grey cells are the highest intensities measured. In some cases, the highest intensity levels correspond to the stressed syllables (e.g. *ka.ḍə.nə.wa, æd.da, ad.də.wə.nə.wa,* and *i.ňdə.gæ.nə*). However, for the most part, intensity alone is not a very reliable indicator of stress. The fact that intensity alone is not a very good indicator of stress is noted in Ladefoged (2003). He states that stress is really not as simple as measuring intensity. It is likely to be a combination of pitch, length, and loudness, with the first two playing the greatest role. Ladefoged demonstrates his claim by showing three pitch and intensity diagrams of three identical sentences with contrastive stress appearing on a different word. From the pitch diagrams, it is clear that the words with contrastive stress have higher pitch and greater pitch movement. The intensity diagrams also show all words have nearly equal intensity. Ladefoged concludes that intensity as shown in decibels is usually not a very useful acoustic property to measure. To test this claim and its relevance to Sinhala, a follow-up experiment was conducted. Using the tokens recorded for the experiment above, the syllables were analysed for pitch. Table 3 contains the highest pitch values for each syllable in the tokens shown in Table 2.

	Highest Pitch (in Hz)				
Target	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5
ka.ḍə.nə.wa	99.84	102.62	93.56	102.49	
kæ.ḍu.wa	103.03	107.95	99.82		
ka.ḍə.wə.nə.wa	100.11	101.94	99.43	96.59	98.55
kæ.**ḍew**.wa	99.75	101.30	95.89		
a.**di**.nə.wa	104.56	113.08	97.67	106.52	
æd.**da**	95.66	100.42			
a.**din**.ne	99.71	104.02	106.98		
æd.**de**	98.34	105.29			
a.**dii**	102.31	109.99			
ad.**də**.wə.nə.wa	96.74	108.00	98.05	94.23	101.05

	Highest Pitch (in Hz)				
i.ňdə.**gan**.na.nə.wa	114.92	102.18	106.88	102.21	97.96
i.ňdə.**gan**.nə.wai	103.19	100.34	107.35	105.32	101.12
i.ňdə.**gan.nan.naŋ**	108.39	100.79	102.06	97.88	102.51
i.**ňdə**.gæ.nə	106.44	109.07	113.13	113.42	

TABLE 3. Selection of results from pitch measurements. [a]

[a] Highest values are shaded in grey. Syllables in bold print were identified by the consultant as being prominent.

Surprisingly, the highest pitch was often found on the first or second syllable of each token. However, the highest value often does not fall on the stressed syllable. Perhaps a static measurement of pitch is not an excellent method for analysing stress. Cutler (2005) suggests that syllables are perceived to be stressed if they exhibit excursion of the fundamental frequency. It follows that the syllables should be examined for the amount of pitch movement. Figure 2 shows the pitch diagram for *kaḍəwai.*
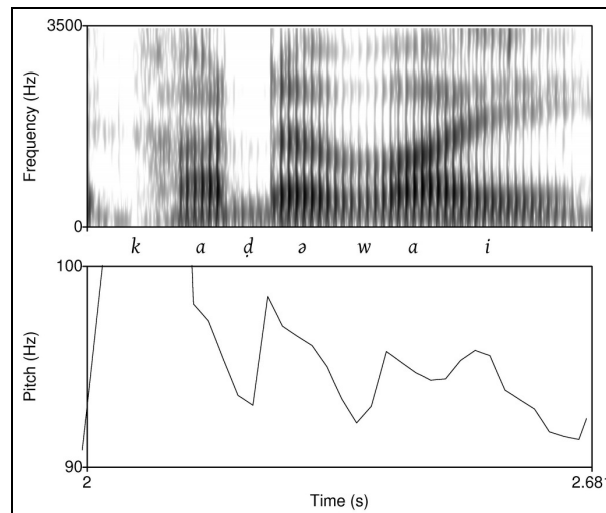
FIGURE 2. Spectrogram and pitch trace of the token *ka.ḍə.wai.*

It is important to note the high pitch values in the first syllable in the figure above. This is caused by extrapolation introduced by Praat. During the closure of the stop [k], there is no frequency information for the program to analyse. Still during the burst, there is no low frequency information (e.g. voicing). However, the burst itself contains a significant level of high frequency information. Therefore, the software extrapolates information between these two points creating a sharp rise. Once voicing begins for the vowel [a], the computer finds information corresponding to F_0. Again, the computer extrapolates information between the burst and the vowel creating a sharp fall. The same phenomenon can be seen with the stop [ḍ] (in the figure above). The only difference is that there is voicing information that is conflicting with the high energy burst, which keeps it from rising too sharply. Unfortunately, examining the pitch trace does not provide useful information in determining the stress *ka.ḍə.wai* which was found to be on the final syllable. The first syllable, if one were to ignore the extrapolation, demonstrates a movement of 5.4 Hz. The second syllable, also ignoring the extrapolation, has a movement of 4.2 Hz. The final syllable has a movement of only 4.5 Hz. Again, pitch does not

immediately provide an insight to stress. The analysis of pitch will be abandoned for the remainder of this paper. The next experiment will analyse duration in hopes of finding a better correlate for stress.

2.3. EXPERIMENT II: MEASURING SYLLABLE DURATION. The previous experiment which examined intensity as a correlate for stress had mixed results. Furthermore, pitch appeared to play only a small role in realising lexical stress. Although this is contrary to results found in many previous studies, it is not necessarily conflicting information. As stated in the introduction, all languages containing stress can use any of the four correlates (and in any combination) in the realisation of stress. This section will analyse duration as a potential correlate.

Using the tokens collected for the first experiment, each syllable was measured for duration. The measurements for selected tokens are provided in Table 4.

Target	Duration of Syllable (in seconds)					
	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6
i.ňdə.**gan**.nə.**wan**.ne	0.079	0.224	0.259	0.097	0.294	0.093
i.ňdə.**gan**.na.nə.wa	0.050	0.185	0.296	0.110	0.149	0.167
i.ňdə.**gan**.nə.**wai**	0.047	0.172	0.257	0.112	0.277	
i.ňdə.**gan**.**nan**.**naŋ**	0.065	0.172	0.289	0.281	0.260	
i.ňdə.**gan**.**naŋ**	0.043	0.224	0.311	0.239		
i.**ňdə**.gæ.nə	0.045	0.187	0.134	0.142		
i.ňdə.ga.**nii**	0.092	0.190	0.177	0.272		
i.ňdə.**gan**.nu.we	0.085	0.186	0.306	0.083	0.145	
i.ňdə.**gat**.ta	0.083	0.174	0.298	0.193		
i.ňdə.**gat**.te	0.112	0.167	0.274	0.162		
ad.**də**.wə.nə.wa	0.162	0.186	0.137	0.157	0.200	
æd.**dew**.wa	0.190	0.365	0.249			
æd.də.**wan**.ne	0.196	0.204	0.212	0.165		
æd.**dew**.we	0.218	0.353	0.215			

TABLE 4. Measurements of duration for selected tokens. [a]

[a] Highest values are shaded in grey. Syllables in bold print were identified by the consultant as being prominent.

It should be noted that the stressed syllables are mainly heavy syllables (i.e. those with codas, long vowels, or diphthongs). It follows that syllables with more segments will have longer durations. However, a few measurements seem problematic. Firstly, many of the syllables that occur in utterance-final position appear slightly longer than other syllables (e.g. *i.ňdə.gan.na.nə.wa* , *i.ňdə.gat.ta* , *æd.dew.wa,* and *æd.dew.we*). This lengthening appears to be an utterance-final phenomenon. Upon closer inspection, the duration of utterance-final CV syllables appear to be 1.2 times the length of word-medial CV syllables. If one were to take the duration of the final syllable in *ad.də.wə.nə.wa* and divide it by 1.2 to compensate for lengthening, it will have a value of 167 ms. This will then result in the second syllable being the longest and therefore corresponding to the syllable judged to be stressed.

Target	Duration of Syllable (in seconds)					
	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6
i.ňdə.**gan**.nə.**wan**.ne	0.079	0.224	0.259	0.097	0.294	0.093
i.ňdə.**gan**.na.nə.wa	0.050	0.185	0.296	0.110	0.149	0.139
i.ňdə.**gan**.nə.**wai**	0.047	0.172	0.257	0.112	0.231	
i.ňdə.**gan**.**nan**.**naŋ**	0.065	0.172	0.289	0.281	0.217	

	Duration of Syllable (in seconds)				
i.ňdə.**gan.naŋ**	0.043	0.224	0.311	0.199	
i.**ňdə**.gæ.nə	0.045	0.187	0.134	0.142	
i.ňdə.ga.**nii**	0.092	0.190	0.177	0.227	
i.ňdə.**gan**.nu.we	0.085	0.186	0.306	0.083	0.145
i.ňdə.**gat**.ta	0.083	0.174	0.298	0.161	
i.ňdə.**gat**.te	0.112	0.167	0.274	0.135	
ad.**də**.wə.nə.wa	0.162	0.186	0.137	0.157	0.167
æd.**dew**.wa	0.190	0.365	0.208		
æd.də.**wan**.ne	0.196	0.204	0.212	0.138	
æd.**dew**.we	0.218	0.353	0.215		

TABLE 5. Tokens from Table 4 with modified utterance-final syllable measurements. [a]

[a] Highest values are shaded in grey. Syllables in bold print were identified by the consultant as being prominent.

A second issue to consider is that a few heavy syllables are significantly longer than others; yet, shorter syllables are still judged to be stressed. This problem can be seen by examining the values for *iňdəgannaŋ*. The duration of the final syllable (199 ms) is shorter than the duration of the second syllable (224 ms); yet, the final syllable is judged to be stressed. The last issue to examine is the relative closeness of some values. In the token *æd.də.wan.ne*, the peninitial and penultimate syllables only differ by 8 ms, and the initial and penultimate syllables by 16 ms. Why should such a small amount of time make such a huge difference in determining stress? Based on impressionistic analysis, the penultimate syllable seems to have a higher degree of stress than the peninitial syllable, and the first syllable does not seem to be stressed. This leads to the next experiment which derives a complex calculation involving intensity and duration in order to achieve a better characterisation of stress.

2.4. EXPERIMENT III: PRELIMINARY CALCULATIONS BASED ON INTENSITY AND DURATION. So far in this paper, it has been shown that intensity alone is not a reliable correlate of stress. Also, it has been demonstrated that measuring the highest pitch per syllable or interpreting pitch movement is not profitable. The last experiment has shown that greater duration is a good indicator of a stressed syllable. However, there are still a few issues that needed resolving. This section will combine intensity and duration measurements in order to devise a better measurement for stress.

Before proceeding forward, it should be pointed out that there have been studies conducted which attempted to link duration and intensity as physical correlates of stress. One such study is that of Fry (1955). In his study, Fry chose a group of English words in which a change of lexical category is commonly associated with a shift of stress from one syllable to another (e.g. initial stress on the noun 'object' and final stress on the verb 'object'). Using spectrography, he measured the duration of all segments and the highest intensity achieved within the vowels for each target word. Since his target words were all two-syllable words, he was able to derive clean ratios which show that both duration and intensity are cues the judgement of stress. Furthermore, it was shown that vowels demonstrate the major differences in duration and intensity with a shift of stress. In the second part of his study, he manipulated vowel length and intensity levels for the target words. He then asked native English speakers to judge which syllable appeared to be stressed. The results show that duration was a more effective cue than intensity. In Section 2.2 and 2.3 of this paper, it was demonstrated that

duration of the syllable was more closely correlated with stress than intensity. Thus, the findings here correspond with those of Fry's.

However, since the correlations of the duration and intensity with stress were not precise, it is worth pursuing some method of integrating the two cues in order create a better fit. The formula derived for this experiment is based on two hypothetical situations. In the first hypothetical situation where all the syllables of a word have exactly the same intensity levels, one would expect that the syllable with the longest duration is the stressed syllable. This is simply an alternate way of stating that heavy syllables are stressed, which was shown in the previous section. In the second hypothetical situation, if all syllables of a word were exactly the same length (e.g. CV.CV.CV.CV), then one would expect the syllable with greatest intensity level to be the stressed syllable. This is based on observations from the first experiment. Using these hypothetical situations, a simple formula used in the evaluation of stress can be derived simply by multiplying the intensity of a syllable by its duration, hence $s = xi$ where s is the result in units of decibel seconds (dB×s), x the duration in seconds, and i the intensity in decibels. For example, let us assume to have a two syllable word. The first syllable has a duration of x_1 and intensity value i_1, whereas the second syllable has a duration of x_2 and intensity value i_2. We can compare the two syllables as follows: x_1i_1 ? x_2i_2. If we take the first hypothetical situation and assume the intensities to be equal, i_1 and i_2 can be replaced with a single value, i, which falls out of the equation. This means the syllable with the greater duration will have a higher overall value and is realised as the stressed syllable. The same line of reasoning applies to the second hypothetical situation. The equation is essentially taking the area of a rectangular region of an intensity diagram. Therefore, the aim of this equation is to approximate the area under the intensity curve for each syllable. A straightforward method of approximating the area of the curve for a syllable is to simply use the highest intensity value and multiply it by its duration. However, using this method will result in excessively calculating area outside of the curve. This overestimation is shown in black in Figure 3(a). An alternative to this method is to multiply the average intensity value by its duration. This method, shown in Figure 3(b), calculates less area outside the curve. It is difficult to estimate which of these methods will produce better results, since this highly depends on the shape of each curve. Some curves are flatter and will have a lesser degree of overestimation using either of these techniques, whereas sharper curves will result in a higher degree of overestimation. This stage of the investigation will use the average intensity in the calculation of the energy of the syllable.
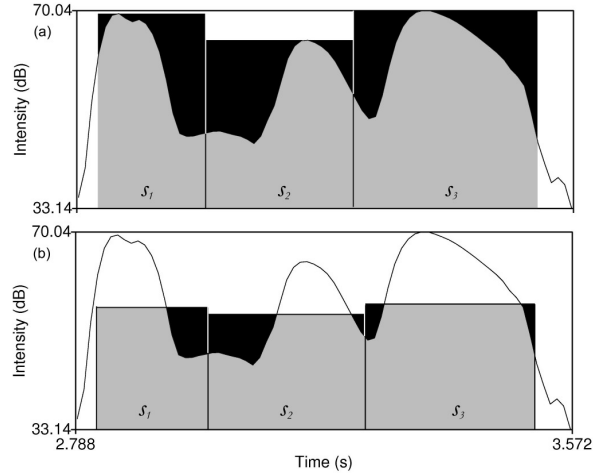
FIGURE 3. Simple methods for estimating the area under the intensity curve. [a]

[a] In (a) the highest intensity is used to define the upper boundary of each syllable. In (b) the average intensity is used to define the upper boundary. The black regions indicate the area outside of the curve that is included in the estimations.

For each syllable, the highest and lowest intensity values were recorded and then averaged. The average intensity was then multiplied by the duration of the syllable yielding a dB×sec value. This is reported below for selected conjugations of the verb *iňdəgannə* 'to sit'.

Target	Syllable (Average Intensity - Duration)					
	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6
i.ňdə.**gan**.nə.**wan**.ne	4.88	13.84	16.21	6.14	18.09	5.40
i.ňdə.**gan**.na.nə.wa	3.01	11.18	18.81	6.57	8.80	9.66
i.ňdə.**gan**.nə.**wai**	2.67	9.85	15.92	6.79	16.54	
i.ňdə.**gan.nan.naŋ**	3.72	10.22	17.45	15.79	14.08	
i.ňdə.**gan.naŋ**	2.47	13.00	18.13	13.38		
i.**ňdə**.gæ.nə	2.71	11.38	8.30	8.77		
i.ňdə.ga.**nii**	5.59	10.94	10.63	14.99		
i.ňdə.**gan**.nu.we	4.86	10.84	17.28	4.69	8.41	
i.ňdə.**gat**.ta	5.06	10.51	14.87	9.38		
i.ňdə.**gat**.te	6.95	10.30	15.24	8.31		
æd.**da**	11.88	17.14				
æd.**de**	12.98	14.35				
ad.**də**.wə.nə.wa	9.11	10.51	7.79	8.68	9.14	
æd.**dew**.wa	10.89	21.09	15.29			
æd.də.**wan**.ne	11.37	11.88	13.98	10.00		
æd.**dew**.we	12.67	20.39	13.77			
æd.də.**wan.naŋ**	10.70	13.95	17.06	17.45		
æd.də.**wai**	10.34	12.84	14.95			

TABLE 6. dB×sec values for each syllable in selected conjugations of ***iňdəgannə*** and ***adinnə.*** [a]

[a] Highest values are shaded in grey. Syllables in bold print were identified by the consultant as being prominent.

Using the measuring technique suggested above, the correlation to stressed syllables is slightly stronger. This is demonstrated by the comparing the values for *iňdəgannaŋ* in Table 6

with those in Table 5. Based on the result of this experiment, one might be able to get an even better fit by precisely measuring the area under the intensity curve for each syllable. The next experiment will do exactly that.

2.5.   EXPERIMENT IV: CALCULATIONS BASED ON INTENSITY AND DURATION. The previous experiment derived a simple equation using both intensity and duration, in order to predict stress. The equation is basically an approximation of the area bounded by the intensity curve for a syllable. Syllables having a high dB×sec value corresponded with stressed syllables. The first part of this final experiment will precisely measure the intensity×time value for each syllable. The results will be compared to native speaker's intuition.

The method of measuring the area bounded by the intensity curve presented in an intensity diagram is to sample the intensity value at small, but regular, intervals. Hence, one is taking what appears to be a continuous curve and sampling the value at regular intervals throughout the syllable. This results in a set of discrete measurements. This is the basis of discrete theory, which is used in the digital reproduction of analogue signals. The measurements can be used to recreate the original curve by extrapolating information between two measurements. Hence, smaller intervals will result in a better representation of the original curve. To perform such measurements for this experiment, an acoustic analysis program developed for Matthew Gordon's research project on the perception of stress was used. Before conducting any measurements, some information regarding the analysis program is provided.

The acoustic analysis program used for this experiment was initially designed to model the perception of acoustic signals. The program takes standard RIFF .wav files of any sampling rate and creates spectra using 11 ms windows and performing a Fast Fourier Transform on each window. This is the initial process of creating a spectrogram. Each spectrum contains the values corresponding to the intensity of 128 frequency bins (i.e. each spectrum has 128 data points). The width of each bin is one-half the sampling rate (i.e. Nyquist frequency) divided by 128 (in this case, approximately 128 Hz). The values for each bin undergo a series of transformations based on psycho-acoustic research (e.g. attenuation of frequencies based on outer and inner ear properties, refractory effects, and frequency attenuation based on the intensity of neighbouring frequencies). After these transformations, the values from each spectrum are summed up resulting in a perceptual value. However, for this experiment, all of the psycho-acoustic models have been disabled, leaving the program to simply add the acoustic intensity values from each spectrum. Figure 4 demonstrates the summation process.
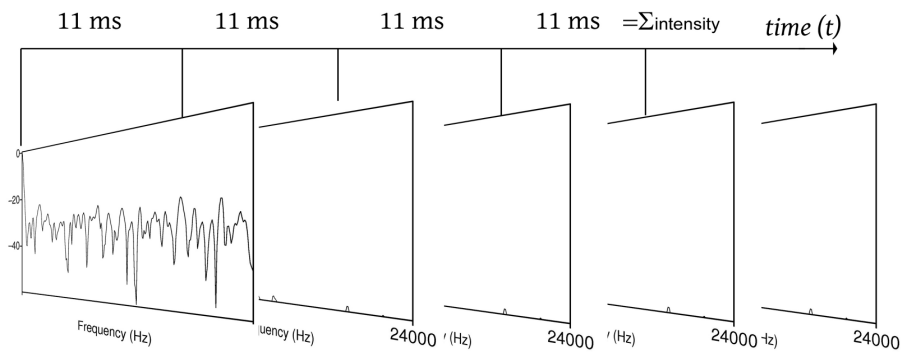
FIGURE 4. Demonstrating the summation process.

As the figure above suggests, the program moves along an audio file, in 11 ms steps, summing up the intensities. The results produced by the program are extremely large. Each spectrum provides 128 data intensity points and each syllable has a number of spectra between 4 and 30. Therefore, the results provided in this section will be linearly scaled in order to make it easier to compare values. Furthermore, since utterance-final lengthening has been regularly observed, syllables occurring utterance-finally will be scaled down to 83.3% (i.e. 1÷1.2, see Section 2.3) to compensate for lengthening.

It is important to emphasise that the values presented in this section should be interpreted relatively to each other. A syllable with an intensity-duration value of 17 is considered to have a higher degree of prominent than a syllable with a value of 10. At this stage of analysis, it is unclear how much difference is necessary in order for one syllable to be judged more stressed than the other.[2] The data provided below will be presented in three sub-sections: (1) analysis of open syllables, (2) analysis of closed syllables, and (3) analysis of syllables with long vowels. First, CV syllables will be analysed. Below are selected tokens composed of CV syllables.

	Intensity-Duration Value (Syllable)						
Target	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6	Syll. 7
i.**ĭdə**.gæ.nə	5.13	12.82	10.85	6.36			
ka.ḍə.nə.wa	12.81	8.26	11.15	8.76			
kæ.ḍu.we	12.27	11.65	8.32				
ka.**ḍə**.wə.nə.wa	13.08	8.10	8.10	9.14	9.06		
pe.nu.na	15.24	12.89	14.82				
pe.nu.ne	15.24	12.89	14.81				
a.**di**.nə.wa	11.25	12.03	9.51	15.14			

TABLE 7. Analysis of tokens with only open syllables. [a]

[a] Grey shading indicates syllables with the highest intensity value. Syllables in bold print were identified by the consultant as being prominent.

From the data presented in Table 7, it appears that prominence predominately occurs word-initially, with the exception of *iĭdəgænə* and *adinəwa*. The difference between these two tokens and the others in the table is that their initial syllable does not contain an onset. It is possible that a word-initial syllable without an onset does not receive stress. The stress shifts over to the next syllable with an onset. This analysis can be supported by examining tokens that begin with a vowel as seen below.

---

[2] Further experimentation is needed to fully answer this question. When verifying the results of this experiment with native speaker intuition, sometimes a difference of 1.6 between the intensity-duration values of two syllables was sufficient for the consultant to indicate which syllable is stressed (see *æd.da* in Table 8). Also, there appeared to be a bias of selecting heavy syllables. Considering the token i.*ĭdə.gan.nə.wan.ne* in TABLE 9, there is a 5.38 difference between the second and fourth syllables and a 6.79 difference between the second and last syllables. Since there are two other syllables with even higher values, the consultant readily identified the heavy syllables as being stressed.

	Intensity-Duration Value (Syllable)						
Target	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6	Syll. 7
æd.**da**	14.57	16.12					
æd.**de**	15.17	17.64					
ad.**də**.wə.nə.wa	13.43	12.16	8.55	8.49	7.70		
ad.**də**.**wan**.ne	12.48	14.12	16.88	12.15			
ad.**də**.**wan**.naŋ	12.92	15.68	20.70	18.03			
ad.**də**.wai	12.78	13.07	15.06				
i.**ňdə**.ga.**nii**	6.00	11.81	11.56	15.51			
i.**ňdə**.**gan**.na.nə.wa	6.24	13.10	19.57	10.99	10.49	9.62	

TABLE 8. Analysis of words beginning with a vowel. [a]

[a] The grey shading indicates the higher of the two values occurring in the first two syllables. Syllables in bold print were identified by the consultant as being prominent.

When comparing only the first two syllables of the tokens in Table 8, the second syllables show slightly higher values than the first ones, with the exception of *addəwənəwa*. Furthermore, the first CV syllable, granted if only a V or VC syllable precedes it, typically has a higher intensity value when compared to other CV syllables within the same word (e.g. *addəwanne*). It can be generalised that the first syllable containing an onset will have some degree of stress.

The next two syllable types, closed syllables and those containing long vowels, are relatively straight-forward. Table 9 and Table 10 present tokens of these types along with their intensity values.

	Intensity-Duration Value (Syllable)						
Target	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6	Syll. 7
i.**ňdə**.**gan**.nə.**wan**.ne	1.78	14.42	18.39	9.04	18.20	7.63	
i.**ňdə**.**gan**.na.nə.wa	6.24	13.10	19.57	10.99	10.49	9.62	
i.**ňdə**.**gan**.**nan**.naŋ	8.98	10.87	17.71	15.16	13.68		
i.**ňdə**.**gan**.nu.we	4.35	12.64	18.46	6.61	10.13		
i.**ňdə**.**gat**.ta	4.55	11.94	18.84	11.50			
i.**ňdə**.**gat**.te	6.42	12.48	18.09	9.18			
kæ.**ḍew**.wa	10.28	19.66	12.57				
kæ.**ḍew**.we	11.16	19.55	10.24				
ka.**ḍə**.wan.naŋ	11.20	7.04	18.23	15.00			
sud.də.kə.rə.nə.wa	17.83	10.34	14.29	10.09	12.82	12.63	
sud.də.ke.ruw.wa	14.95	10.11	15.66	19.08	11.36		
sud.də.kə.**ran**.ne	16.92	8.83	12.72	18.98	13.70		
sud.də.kə.**ran**.naŋ	16.38	11.10	11.74	16.93	15.26		
sud.də.kə.rə.wə.nə.wa	17.82	9.36	12.19	10.81	10.10	9.99	13.97
pen.**nan**.naŋ	18.85	22.22	18.30				
a.**din**.naŋ	7.39	17.36	15.68				
ad.**də**.**wan**.naŋ	12.92	15.68	20.70	18.03			

TABLE 9. Analysis of closed syllables. [a]

[a] Grey shading indicates closed syllables. Syllables in bold print were identified by the consultant as being prominent.

Target	Intensity-Duration Value (Syllable)						
	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6	Syll. 7
i.ňdə.**gan**.nə.**wai**	6.11	14.22	16.51	9.53	17.43		
i.**ňdə**.ga.**nii**	6.00	11.81	11.56	15.51			
ka.**ɖai**	11.08	19.98					
ka.də.wai	10.91	10.24	20.06				
sud.də.kə.**rai**	14.78	11.29	12.04	18.72			
sud.də.kə.rə.**wai**	14.30	8.52	10.49	10.65	16.58		
pee.nə.wa	26.83	10.95	11.12				
peen.ne	35.31	13.43					
pe.**nei**	14.09	22.96					
a.**dii**	9.17	19.25					
ad.**də**.wai	12.78	13.07	15.06				

Table 10. Analysis of syllables with long vowels. [a]

[a] Grey shading indicates syllables with long vowels or diphthongs. Syllables in bold print were identified by the consultant as being prominent.

In Table 9, closed syllables consistently had the highest intensity-duration values, while in Table 10, syllables with long vowels or diphthongs had the highest values. Table 11 compares the values gathered in this experiment with native speaker judgement. The number of grey cells in each row corresponds to the number of prominent syllables indicated by the consultant.

Target	Intensity Value (Syllable)						
	Syll. 1	Syll. 2	Syll. 3	Syll. 4	Syll. 5	Syll. 6	Syll. 7
i.ňdə.**gan**.nə.**wan**.ne	1.78	14.42	18.39	9.04	18.20	7.63	
i.**ňdə**.**gan**.na.nə.wa	6.24	13.10	19.57	10.99	10.49	9.62	
i.ňdə.**gan**.nə.**wai**	6.11	14.22	16.51	9.53	17.43		
i.**ňdə**.**gan**.nan.naŋ	8.98	10.87	17.71	15.16	13.68		
i.**ňdə**.**gan.naŋ**	3.97	14.32	19.27	12.20			
i.**ňdə**.gæ.nə	5.13	12.82	10.85	6.36			
i.**ňdə**.ga.**nii**	6.00	11.81	11.56	15.51			
i.ňdə.**gan**.nu.we	4.35	12.64	18.46	6.61	10.13		
i.**ňdə**.**gat**.ta	4.55	11.94	18.84	11.50			
i.**ňdə**.**gat**.te	6.42	12.48	18.09	9.18			
ka.də.nə.wa	12.81	8.26	11.15	8.76			
kæ.ɖuw.wa	14.49	14.07	10.31				
ka.ɖan.ne	14.19	19.07	10.90				
kæ.ɖu.we	12.27	11.65	8.32				
ka.ɖan.**naŋ**	10.26	16.12	13.28				
ka.**ɖai**	11.08	19.98					
ka.**ɖə**.wə.nə.wa	13.08	8.10	8.10	9.14	9.06		
kæ.**ɖew**.wa	10.28	19.66	12.57				
kæ.**ɖew**.we	11.16	19.55	10.24				
ka.**ɖə**.wan.naŋ	11.20	7.04	18.23	15.00			
ka.ɖə.wai	10.91	10.24	20.06				
pee.nə.wa	26.83	10.95	11.12				
peen.ne	35.31	13.43					
pe.nu.ne	15.24	12.89	17.78				
pe.**nei**	14.09	22.96					

Intensity Value (Syllable)						
pen.nan.ne	22.43	22.95	12.67			
pen.nu.we	20.92	15.08	12.67			
pen.nan.naŋ	18.85	22.22	18.30			
pen.**nai**	20.26	20.67				
a.**di**.nə.wa	11.25	12.03	9.51	15.14		
æd.**da**	14.57	16.12				
a.**din**.ne	7.48	18.78	13.28			
æd.**de**	15.17	17.64				
a.**din**.naŋ	7.39	17.36	15.68			
a.**dii**	9.17	19.25				
ad.**də**.wə.nə.wa	13.43	12.16	8.55	8.49	7.70	
æd.**dew**.wa	11.44	26.12	15.28			
ad.**də**.**wan**.ne	12.48	14.12	16.88	12.15		
æd.**dew**.we	14.28	28.82	13.40			
ad.**də**.**wan**.naŋ	12.92	15.68	20.70	18.03		
ad.**də**.**wai**	12.78	13.07	15.06			

TABLE 11. Sample tokens and measurements. [a]

[a] Syllables in bold print are judged to be stressed by the consultant. Values in grey are the higher values. The number of cells shaded for each word corresponds to the number of syllables selected by the consultant.

Of the 26 words which the consultant indicated as having one stressed syllable, 18 (69.2%) of the syllables corresponded the highest intensity value measured in the word. Of the 11 words where he chose two syllables (hence, 22 syllables), 19 (86.4%) of the syllables corresponded to the measurements. Of the 4 words where he indicated three stressed syllables (that is 12 syllables), all 12 (100%) corresponded to the measurements. A total of 60 syllables were identified by the consultant. Forty-nine (81.7%) of the syllables he identified were characterised by high intensity-duration measurements.

In summary, intensity and duration appear to be the key correlates in the realisation of stress in Sinhala. Native speaker intuition often correlates with the results yielded by the acoustic analysis conducted in this experiment. Based on this data, a generalisation can be made as regards stress sites. Firstly, heavy syllables are stressed. This is perhaps due to the number of segments, or moras in phonological perspective, within the syllable which in turn translates to longer duration. Secondly, the first syllable containing an onset is also stressed. At times, this initial stress appears to be weaker than the stress found in heavy syllables. If the initial syllable is CV, it will not be perceived as prominent as a CVV or CVC stressed syllable. However, if the initial syllable is CVV or CVC, then the syllable will be perceived as equally prominent as other heavy syllables. Using these generalisations, one is now able to account for stress sites in a phonological framework.

3. A PHONOLOGICAL ACCOUNT OF STRESS. From the experimental data presented in Section 2, one is able to derive a general pattern for locating stress in Sinhala verbs. This pattern can be formalised within a phonological framework. Choosing the appropriate framework is crucial. This section will commit to an Optimality Theory (OT) approach and show how OT accounts for the stress pattern observed.

Kager (1999) states that word stress patterns are typically governed by conflicting forces. The interaction of conflicting metrical constraints has been observed and reported in many pre-OT studies. The forces in conflict are rhythm, quantity-sensitivity, and edge marking. Rhythm is the pressure towards a regularly alternating distribution of weak and strong

syllables. Quantity-sensitivity is the pressure to match syllable weight to prominence. Edge-marking is the pressure to mark the edges of morphemes. One can describe the pattern seen in the Sinhala verbs as a conflict of quantity-sensitivity with edge-marking (i.e. the realisation of stress on heavy syllables and the realisation of a weak stress word-initially). The central idea of Optimality Theory is that surface forms of a language reflect solutions to conflicts between competing demands. This makes OT the ideal framework to describe Sinhala stress. In order to derive the necessary constraints and rankings, one needs to compare the surface form and the suboptimal forms. This will be accomplished below.

Before deciding which constraints are needed to describe Sinhala verb stress, one first must recognise the fact that a foot-based approach is not appropriate. Firstly, there is a requirement that the head of every foot be the locus of stress. Furthermore, the foot is usually analysed as disyllabic or bimoraic. This assumption is used to describe the regular alternating distribution of weak and strong syllables. However, there is no evidence of this alternation in Sinhala. For example, the verb *ka.ɖə.wə.nə.wa* has been reported by the consultant as only having initial stress, where as *i.ňdə.gan.nan.naŋ* has a weak stress on the second syllable and heavy stress on the following three syllables. Therefore, one must not appeal to foot-based constraints. There are constraints that rely on the syllable. This is more appropriate for this task.

Having eliminated foot-based constraints, one can now begin with the analysis. The most straight-forward conflicting force to describe is syllable weight. As shown in the preceding section, heavy syllables are stressed. In this case, suboptimal candidates are defined as those with unstressed heavy syllables.

(1) i.ňdə.**gan.nan.naŋ** > i.ňdə.**gan.nan.**naŋ, i.ňdə.gan.**nan.naŋ** > i.ňdə.gan.**nan.**naŋ

The suboptimaility of the three rightmost forms in (1) is due to a constraint enforcing quantity-sensitivity. There is a close relation between syllable weight and the degree of prominence. Therefore, these forms are in violation of Weight-to-Stress-Principle (Kager 1999).

(2) **Weight-to-Stress-Principle (WSP)**
     Heavy syllables are stressed.

WSP is violated anytime a heavy syllable is not stressed. It is important to note that this constraint could potentially be cumulative, in other words accrue multiple violations. For example, in (1) the right-most form would accrue two violations for having two unstressed heavy syllables.

The second conflicting force is the pressure for word-initial stress. However, there are two parts to this pattern: (1) a force pulling stress towards the left-edge of the word, and (2) the demand for an onset in the initial stressed syllable. Before defining the constraints behind these forces, we turn to current research on weight-sensitive stress for some insight as regards onset.

Gordon (forthcoming) notes that the vast majority of weight-sensitive stress systems ignore onsets in the calculation of syllable weight. Yet, it has been shown that some language do demonstrate clear cases of onset-sensitive stress. Gordon cites the work of Everett and Everett (1984) on Pirahã, an indigenous language spoken in Brazil. Pirahã has a five-way weight hierarchy in which both the onset and the rime play a role in syllable weight. The five-way hierarchy is as follows: KVV > GVV > VV > KV > GV where K stands for a voiceless

consonant and G for a voiced consonant. He also cites Iowa-Oto which has stress on the first syllable unless it is onsetless, in which case stress is then on the second syllable. This is the case with Sinhala. Other languages containing this same onset-sensitive system are Lamalamic, Umbuykam, Parimankutinma, Banawá, and Arrernte. Gordon's article is an attempt to explain the basis for onset-sensitive stress and its rarity relative to rime-sensitive stress. He presents compelling evidence based on studies of the auditory system. For example, the auditory system is most sensitive to a stimulus at its onset before auditory sensitivity declines. This is typically referred to as *adaptation.* Once there is silence or a decrease of stimulation, the auditory system recovers.

In his phonological description of these systems, Gordon proposes a few important constraints, two of which will be used to describe Sinhala stress. First, notice in Section 2.5 that all but one of the syllables with an intensity-duration value has an onset. It can be stated that there is a demand for stress bearing syllables to have an onset. In the conjugations of the verb *iňdəgannə* the initial syllable is never stressed.

(3) i.**ňdə**.**gat**.ta > i.ňdə.**gat**.ta, i.ňdə.gat.**ta**, i.ňdə.gat.ta,

This pattern is addressed by employing a prominence constraint that bans the occurrence of stress on onsetless syllables (Gordon forthcoming).

(4) *PROM[∅[X]$_R$]$_σ$

No syllable lacking an onset can carry prominence.

The second part of this conflicting force is the demand to align with the left edge of the word. Kager (1999) and Gordon (forthcoming) offer a set of alignment constraints that can be used to address this demand. In general, the alignment constraint is defined as in (5).

(5) **Generalised Alignment**

ALIGN (Cat$_1$, Edge$_1$, Cat$_2$, Edge$_2$) =

∀ Cat$_1$ ∃ Cat$_2$, such that Edge$_1$ of Cat$_1$ and Edge$_2$ of Cat$_2$ coincide.

In order to use this constraint, the categories and edges must be defined. The claim is that a stressed syllable occurs word-initially. Therefore, the two categories are stressed syllable (σ́) and prosodic word (PrWd). The edges involved are the left edge of the stressed syllable and the left edge of the prosodic word. The definition of the specific align constraint is given in (6).

(6) ALIGN (σ́, L, PrWd, L)

For every stressed syllable there exists a prosodic word that the left edge of the stressed syllable matches the left edge of the prosodic word.

It is apparent from the data seen in Section 2.5 that not all stressed syllables occur at the left-edge of the prosodic word. Given the manner this constraint is defined, there is now a need to crucially rank the constraints in (4) and (6). Consider the ordering of WSP and the ALIGN constraint in the following tableau.

(7)

Input: /sud.də.ke.ruw.wa/	WSP	ALIGN (σ́, L, PrWd, L)
a. ☞ ˈ**sud**.də.ke.ˈ**ruw**.wa		***
b. ˈ**sud**.də.ke.ruw.wa	*!	

In (7), the winning candidate is the attested form is. This is due to the ranking of WSP above ALIGN. If the order is reversed, the winning candidate will have only one stressed syllable which must occur at the left edge. This is demonstrated in (8).

(8)

Input: /sud.də.ke.ruw.wa/	ALIGN (ó, L, PrWd, L)	WSP
a. ˈsud.də.ke.ˈruw.wa	*!**	
b. ☞ ˈsud.də.ke.ruw.wa		*

Notice in (8a) the fourth syllable, which is behaving in accordance to the Weight-to-Stress Principle, accrues three violations of the ALIGN constraint since it is three syllables removed from the left edge. The winning candidate in (8b) avoids this by violating the WSP constraint. Therefore, it has been established that WSP must be ranked above ALIGN.

To determine the overall hierarchy, it must be shown how *PROM is ordered relative to the other two constraints. First, the ordering between *PROM and ALIGN will be considered. The potential ordering for these two constraints can be shown by examining a verb in which the initial syllable is not stressed, as in (9).

(9)

Input: /ad.də.wan.ne/	*PROM[∅[X]ᵣ]σ	ALIGN (ó, L, PrWd, L)
a. ☞ ad.**də.wan**.ne		*, **
b. **ad.də.wan**.ne	*!	**
c. **ad**.də.wan.ne	*!	

The tableau in (9) shows the winning candidate is the one that does not violate *PROM. Therefore *PROM must be ranked above ALIGN. Now one must determine if there is a ranking between *PROM and WSP. This is demonstrated in (10).

(10)

Input: /ad.də.wan.ne/	*PROM[∅[X]ᵣ]σ	WSP
a. ☞ ad.**də.wan**.ne		*
b. **ad.də.wan**.ne	*!	

When *PROM is ranked above WSP, candidate (10a) surfaces as the winning candidate. This is attested. Therefore, the overall ranking is *PROM[∅[X]ᵣ]σ >> WSP >> ALIGN (ó, L, PrWd, L). However, as it turns out there is one more constraint needed. Consider the following example of *iňdəganii* in (11).

(11)

Input: /i.ňdə.ga.nii/	*PROM[∅[X]ᵣ]σ	WSP	ALIGN (ó, L, PrWd, L)
a. i.**ňdə**.ga.**nii**			*, ***!
b. ☞ i.**ňdə**.ga.**nii**			***

In (11b) the winning candidate violates the ALIGN constraint one time fewer than the attested candidate. This undesired outcome is prevented by allowing at most one unstressed syllable to

separate the leftmost stress from the left edge of the prosodic word (Gordon forthcoming). This constraint is called LAPSE LEFT.

> (12) **LAPSE LEFT** (Gordon 2002)
> A maximum of one unstressed syllable separates the leftmost stress from the left edge of a stress domain

LAPSE LEFT must be ranked above the ALIGN constraint in order prevent the outcome seen in (11). This is demonstrated in (13).

(13)

Input: /i.ňdə.ga.**nii**/	LAPSE LEFT	ALIGN (ó, L, PrWd, L)
a. ☞ i.**ňdə**.ga.**nii**		* ***,
b. i.ňdə.ga.**nii**	*!	***

With LAPSE LEFT, the attested candidate is the winning one. Although it has been stated that the purpose of LAPSE LEFT is to prevent an excessive amount of unstressed syllables from appearing at the left edge of the prosodic word, it can also be thought of as a force that pulls a stressed syllable to the left. Initially, the ALIGN constraint was used to pull all stressed syllables towards the left edge. So, does it appear that the ALIGN constraint is redundant? The answer is no. It not only pulls all stressed syllables to the left, but it also restricts the number of stressed syllables appearing within the word. The following tableau demonstrates this phenomenon.

(14)

Input: /i.ňdə.ga.**nii**/	LAPSE LEFT	ALIGN (ó, L, PrWd, L)
a. ☞ i.**ňdə**.ga.**nii**		* ***,
b. i.**ňdə.ga**.**nii**		*, ** **!*

So far there has been no restriction on the possibility of an open syllable being a stressed. In (14b) the open syllable is stressed; therefore, it accrues two more violations than the candidate in (14a).

The constraint hierarchy established prior to (11) was *PROM[∅[X]ᵣ]σ >> WSP >> ALIGN (ó, L, PrWd, L). In (14) it has been shown that LAPSE LEFT outranks ALIGN. This leads to the question of how does LAPSE LEFT relate to *PROM[∅[X]ᵣ]σ and WSP. To determine its proper position within the hierarchy, LAPSE LEFT will be compared first to WSP. The two candidates in (15) differ only in respects to obeying LAPSE LEFT and WSP.

(15)

Input: /ad.də.wai/	LAPSE LEFT	WSP
a. ☞ ad.**də.wai**		*
b. ad.də.**wai**	*!	

As demonstrated above, in order to have the attested candidate in (15a) be the winning candidate, LAPSE LEFT must be ranked higher than WSP; otherwise the suboptimal candidate (15b) will surface. Now LAPSE LEFT will be compared to *PROM[∅[X]R]σ .

(16)

Input: /ad.də.wai/	*Prom[∅[X]ᵣ]σ	Lapse Left
a. ☞ ad.**də.wai**		
b. **ad.**də.**wai**	*!	
c. ad.də.**wai**		*!

At first glance it appears that *Prom[∅[X]ᵣ]σ outranks in (18). However, notice that the attested candidate in (18a) does not violate either constraint, whereas candidates (18b) and (18c) violate one or the other. If one were to swap the ordering of these two constraints, the attested candidate would still be the winning candidate. This is demonstrated in (17).

(17)

Input: / ad.də.wai /	Lapse Left	*Prom[∅[X]ᵣ]σ
a. ☞ ad.**də.wai**		
b. **ad.**də.**wai**		*!
c. ad.də.**wai**	*!	

Therefore, there is no crucial ranking between Lapse Left and *Prom[∅[X]ᵣ]σ. The finalised constraint hierarchy is as follows: *Prom[∅[X]ᵣ]σ , Lapse Left >> WSP >> Align (σ́, L, PrWd, L). To conclude the analysis of stress in Sinhala verbs, two tableaux with the full constraint rankings are presented in (18) and (19).

(18)

Input: /sud.də.kə.ran.naŋ/	*Prom[∅[X]ᵣ]σ	Lapse Left	WSP	Align (σ́, L, PrWd, L)
a. ☞ sud.də.kə.**ran.naŋ**				*** ****,
b. sud.**də.**kə.**ran.naŋ**			*!	* *** ****, ,
c. sud.də.kə.**ran.naŋ**			*!	****
d. sud.**də.**kə.**ran.naŋ**			*!	** ****,
e. sud.də.kə.**ran.naŋ**		*!	*	*** ****,

(19)

Input: /ad.də.wan.ne/	*Prom[∅[X]ᵣ]σ	Lapse Left	WSP	Align (σ́, L, PrWd, L)
a. ☞ ad.**də.wan**.ne			*	* **,
b. **ad.**də.**wan**.ne	*!			**
c. ad.də.**wan**.ne		*!	*	**
d. **ad.**də.wan.ne	*!		*	
e. ad.**də.wan.ne**			*	* ** *!**, ,

4. Conclusion. Section 2 examined the possibility of finding an objective procedure in the identification of phonetically stressed syllables. It was shown that intensity and duration appear to be the key parameters in the realisation of stress in Sinhala. Based on an equation which integrated the two parameters, a generalisation can be made regarding the stress pattern: (1) heavy syllables are stressed and (2) the first syllable containing an onset is also stressed. At times, the initial stress appears to be weaker than the stress found in heavy syllables. Section 3 focused on the development of a phonological account within an OT

framework. Four conflict pressures were identified: (1) a ban on stress on onsetless syllables, (2) an allowance of an unstressed syllable word-initially, (3) matching prominence with syllable weights, and (4) a pull of stressed syllables to the left, which also governs the number of stressed syllables within the word.

REFERENCES

CUTLER, ANNE. 2005. Lexical Stress. In David B. Pisoni and Robert E. Remez (eds.), *The Handbook of Speech Perception.* Oxford: Blackwell Publishers. 264–289.

FRY, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27. 765–768.

FUJIMURA, OSAMU AND DONNA ERICKSON. 1997. Acoustic phonetics. In William J. Hardcastle and John Laver (eds.), *The Handbook of Phonetic Sciences.* Oxford: Blackwell Publishers. 65–115.

GAIR, JAMES W. AND JOHN C. PAOLILLO. 1997. *Sinhala.* Munich: Lincom Europa.

GORDON, MATTHEW. 2002. A factorial typology of quantity insensitive stress. *Natural Language and Linguistic Theory* 20, 491–552.

GORDON, MATTHEW. Furthcoming. A perceptually-driven account of onset-sensitive stress. To appear in *Natural Language and Linguistic Theory.*

HAYWARD, KATRINA. 2000. *Experimental Phonetics.* Harlow, England: Longman.

KAGER, RENÉ. 1999. *Optimality Theory.* Cambridge: Cambridge University Press.

LADEFOGED, PETER AND IAN MADDIESON. 1996.*The sounds of the world's languages.* Oxford: Blackwell Publishers.

LADEFOGED, PETER. 2001. *A course in phonetics, 4th edition.* Ft Worth: Harcourt College Publishers.

LADEFOGED, PETER. 2003. *Phonetic data analysis: an introduction to fieldwork and instrumental techniques.* Oxford: Blackwell Publishing.

LAVER, JOHN. 1994. *Principles of Phonetics.* Cambridge: Cambridge University Press.

LETTERMAN, REBECCA S. 1997. The effects of word-internal prosody in Sinhala: a constraint-based analysis. PhD dissertation. Cornell University.

MASICA, COLIN. 1991. *The Indo-Aryan Languages.* Cambridge: Cambridge University Press.

MOL, H. G. AND UHLENBECK, G. M. 1965. The linguistic relevance of intensity in stress. *Lingua* 5. 205–213.

INFORMATION PACKAGING IN SINHALA: A PRELIMINARY STUDY OF ADVERBIAL CLAUSES IN FOCUS CONSTRUCTIONS

VALERIE SULTAN

University of California, Santa Barbara

1. INTRODUCTION. A single utterance in discourse carries within it three different components that contribute to the conveyance of a particular message—syntactic, semantic and pragmatic. To illustrate this, consider the English sentences provided below (boldface indicates stress).

(1)   a. Boo-Boo loves bones.
        b. Bones Boo-Boo loves.
        c. Bones Boo-Boo hates.
        d. Boo-Boo loves **bones**.

Each of these sentences can be compared with one or more of the others to demonstrate the presence of either the syntactic, semantic or pragmatic component. Sentences 1a and 1b both provide the same statement about the world and are thus semantically equivalent, but they differ in terms of their syntax (seen with the change in word order) and pragmatics (seen with the fronting, hence focus, of 'bones'). On the other hand, sentences 1b and 1c are syntactically and pragmatically equivalent but differ with respect to their semantic meaning ('hates' versus 'loves'). Finally, sentences 1a and 1d have the same semantic and syntactic components, but the stress-focus on 'bones' causes these sentences to differ pragmatically.

For the purposes of this paper, the most important comparisons just discussed are those concerning 1a, 1b, and 1d. This is because it is when looking at these examples that we see evidence of what Vallduví and Engdahl (1996) refer to as 'information packaging'. As can be seen in the three sentences expressed in 1a, 1b, and 1d, utterances may express the same propositional content despite changes in sentence structure or intonation or both, but they are not, as Vallduví and Engdahl (1996:459) point out, 'interpretively equivalent in absolute terms'. In fact, these sentences differ because of the extrapropositional, or pragmatic, contribution to meaning, and therefore cannot be used interchangeably in the same context (Szendröi 2004; Vallduví and Engdahl 1996). In other words, it is not the message that is different, but the way in which the message is packaged that is different. In order to address this type of difference, Vallduví and Engdahl (1996:460) refer to this behavior as 'information packaging' which they define as 'a structuring of sentences by syntactic, prosodic, or morphological means that arises from the need to meet the communicative demands of a particular context or discourse'. This is to say that speakers design their talk, both at the discourse and sentence levels, according to their beliefs about what hearers can be assumed to know or have in mind in a given context. As an example of this, it is possible to refer back to the sentences provided above. In these examples, both 1b and 1d are utterances constructed with the expectation that the hearer is aware that there is something that Boo-Boo loves, whereas the object of the love—bones—is presumed to be information that is either new to the hearer or contrasts with the hearer's previously held beliefs. The same, however, cannot be said for 1a.

One of the primary means by which information packaging is represented in a sentence is through the use of focus-ground partitions. Such partitions divide a sentence into the ground—the part which is presumed to be known (Andrews 1990; Szendrői 2004) or predictable (Givón 1990) by the hearer and thus anchors the sentence to the previous discourse or the hearer's 'mental world' (Vallduví and Engdahl 1996)—and the focus—a new, informative (Andrews 1990; Szendrői 2004) or less predictable (Givón 1990) part that contributes to the discourse or the hearer's 'mental world' (Vallduví and Engdahl 1996). The definition of information packaging presented earlier notes that speakers can use morphology, syntax, and prosody to meet different communicative demands, so it is to be expected that the focus-ground division is often represented in the morphosyntax of a language, e.g. with special focus constructions. Such a representation is seen clearly in Sinhala.

Sinhala has an extensive focus construction, as has been widely discussed in the literature (Gair 1970, 1998 [1983], 1998 [1985], 1998 [1989], Gair and Paolillo 1997, Gair and Sumangala 1991, Herring and Paolillo 1995, Kariyakarawana 1998). While it appears that the Sinhala focus construction is likely to have derived from contact with Dravidian languages (Gair 1998 [1985]), it has undergone a great deal of internal development and diversification since that presumed historical influence, and as a result there is an increased role and wider range of discourse uses of focus in Sinhala syntax (ibid). Almost all of these discourse uses, to be discussed in more detail later, revolve around the idea of information packaging by either pointing to information that is expected to be unknown to the hearer or contradicting what is assumed to be known or believed by the hearer. Moreover, these focus constructions tend to involve focusing one of the constituents of the clause, hence bring attention to that constituent's new or contradictory information status.

However, it is possible for speakers of Sinhala to capitalize on the interplay of focus constructions and information status to mark the information status of interclausal relations as well as the information status of constituents. This presents a challenge to traditional notions of focus and information flow, as both have been treated as relevant only with respect to referents in a noun phrase, whereas in Sinhala, both are used to refer to referents of predications, i.e. events and states. Furthermore, the pragmatic factors motivating the use of the focus structure in Sinhala is the same for the referents of both noun phrases and predications. In the preliminary study presented in this paper, I will show that this appears to be the case for a set of data in which the focus form of verbs are used in matrix clauses when there exists a set of particular characteristics with respect to their modifying adverbial clauses. Specifically, it will be shown that focus comes into play when an adverbial clause expresses a new event which provides an explanation for the given or inferred event expressed in the matrix clause.

In order to accomplish this goal, the current paper will begin with a general discussion of adverbial clauses, paying special attention to their different interpropositional functions and discourse roles. Following this will be a brief overview of the structure and use of Sinhala focus, which will lead into a section devoted to the examination of Sinhala adverbial clauses in constructions with focused matrix verbs. Finally, the conclusion will address what these findings mean for the interplay between adverbial clauses, focus, and information packaging.

2. TYPES OF ADVERBIAL CLAUSES. As is noted by Thompson and Longacre (1985), adverbial clauses are those that modify a verb phrase or a sentence. Cross-linguistically, three of the devices used to mark subordinate clauses are also seen to mark adverbial clauses. These are: 1)

adverbial particles (either with or without lexical content), 2) special verb forms (i.e. those not used in independent assertions), and 3) word order. Sinhala utilizes the first two mechanisms. The table below provides a list of the adverbial particles and the verb morphology that is allowable with these morphemes (note that in Sinhala the verb precedes the adverbial particle, which will be shown in the examples in the following section). Also included in this table are verb forms that do not co-occur with an adverbial particle, but rather express the adverbial relationship via bound morphology.

Verb Form	Gloss	Adverbial Particle	Gloss	Interpropositional Relationship
-at	CONC.PRES	----	----	Concessive
-a	PST	*wunat*	'even though'	Concessive
-ot	COND.PRES	----	----	Conditional
-naŋ	COND.PST	----	----	Conditional
-iŋ		*passe*	'after'	Time
-mə		----	----	Simultaneous
		sandaha	'in.order.to'	Purpose
-məʈə		*amətərəwə*	'in.addition.to'	Additive
-gat	PPL.REFL	*gamaŋ*	'while'	Simultaneous
		hinda	'because'	Reason
		nisaa	'because'	Reason
nætti	NEG.ADJ	*hinda*	'because'	Reason (negation)
-nə	ADJ	----	----	Reason
		atərədi	'while'	Simultaneous
		gamaŋ[1]	'while'	Simultaneous
		koʈə	'while'	Simultaneous
		hinda	'because'	Reason
		nisaa	'because'	Reason
		pinisə	'in.order.to'	Purpose
-nnə	INF	----	----	Purpose
		issella	'before'	Time
		kaliŋ	'before'	Time
-pu	PST.ADJ	----	----	Reason, Simultaneous
		gamaŋ	'while'	Simultaneous
		hinda	'because'	Reason
		nisaa	'because'	Reason

TABLE 1. Sinhala adverbial clause adverbial particle and verb forms

Note that there are only three adverbial particles that occur with more than one verb form—*gamaŋ* 'while', *hinda* 'because', and *nisaa* 'because' may be used with any of the three participle forms (*-nə, -gat,* and *-pu*)[2]. In all of these cases, the determining factor governing the use of one verb form over another is the timing of the event in the adverbial clause with respect to the event expressed in the matrix clause. The aspectual relationship between the

---

[1] There is one form with *-nnə* that is followed by both *gamaŋ* and *koʈə*, but it is the form *innə* which does not have a corresponding *-nə* form, so it is doubtful that this form actually represents the infinitive in these cases.
[2] This excludes the co-occurrence of *hinda* with the negative existence morpheme *nætti*

other adverbial particles and their respective verb forms becomes evident when one takes into consideration the semantics of the relationship between the adverbial clause and the matrix clause. With the exception of *gamaŋ,* all adverbial particles expressing simultaneity occur with the present adjectival participle *-nə*[3]. In addition, purpose adverbial clauses, which indicate that the act in the adverbial clause is unrealized at the time of the event in the matrix clause, must be expressed with either the present adjectival participle *-nə* or the infinitive *-nnə*. The infinitive is also the only form allowable with morphemes meaning 'before'.

A brief glance at the right-hand column of Table 1 shows that adverbial morphology can express a number of interpropositional relationships between the matrix and adverbial clauses. The next section provides further exploration and illustration of these types.

2.1. INTERPROPOSITIONAL RELATIONSHIPS. Thompson and Longacre (1985) provide a thorough description of the different interpropositional relationships that adverbial clauses can have with the modified matrix clause. They divide the adverbial clauses of the attested languages of the world into twelve basic types, further categorizing them into two groups. The classification they provide appears below.

Time	Simultaneous
Location	Conditional
Manner	Concessive
Purpose	Substitutive
Reason	Additive
Circumstantial	Absolute

TABLE 2. Thompson and Longacre's Classification of Adverbial Clauses (1985:177)

As Thompson and Longacre note, some of these interpropositional relationships are expressed through other grammatical means, i.e. relative clauses, and this is the case for Sinhala. As a result, only those relationships that utilize an adverbial clause are discussed here. Explanations and Sinhala examples (where possible) are provided for each of these types in the following subsections[4].

TIME. Time adverbial clauses concern the sequencing relationship between clauses, typically marked either by verbal affixes or by independent morphemes along the lines of the English 'when', 'before', 'after', and so on. In the Sinhala example below, we see the use of the independent adverbial particle *issella* 'before' with a special verb form to mark a time adverbial (the adverbial clause is highlighted).

(2) Turtle Hatchery, Sentence 15 (Santa Barbara)
      hari   welaawətə   matai   magee   yaaluwatai   mee
      so   time.SG.DEF.DAT.   1SG.DAT.and   1SG.GEN   friend.SG.DEF.DAT.and   1PROX
          just.in.time

---

[3] It is also possible for a verb with the suffix *-mə* to imply simultaneity (like *-nə*) or purpose (like *- nnə*) (the latter with a special adverbial particle), but this suffix is likely related to the emphatic *-mə* and does not inflect for tense or aspect

[4] All explanations are taken from Thompson and Longacre's description

baaldiyə   koociyə   e-nnə   issella   kooci   paareŋ
bucket.SG.DEF.   train.SG.DEF   come-INF   before   train   road.SG.DEF.ABL
                                                            train.track

ehaaṭə   ga- nnə   puluwaŋ   wun-a
other.side.DAT   take-INF   can   do-PST
 'Just in time, my friend and I were able to take the bucket to the other side of the
train track before the train came.'

Note that in this case the adverbial is a full clause with a predicate and its arguments. It is also marked with a special time morpheme (i.e. 'before'). In addition, the verb is in the infinitive form, which cannot be used in independent assertions except for imperatives. All of these features make the highlighted clause an adverbial time clauses.

MANNER. In many languages, a manner clause can be introduced with a subordinator such as 'like'. One of the means of expressing manner in Sinhala is through the use of the adverbial particle *widiyəṭə* 'as/like'. An example of this appears below.

(3)   Elicited
Adverbial: kooṭə   gahee   trikoonəyak   hede-nə   widiyəṭə
            stick.SG.DEF   tree. SG.DEF.LOC   triangle.SG.IND   make-ADJ   as/like

            heettukəra-nnə
            lean-IMP

            'Place the stick on the tree so that it forms a triangle' (lit. 'Lean the stick on
            the tree like making a triangle')

PURPOSE. Thompson and Longacre point out that the interpropositional relationships purpose and reason are often expressed with the same morphology as both provide explanations for the event expressed in the matrix clause. The difference, they note, is that purpose clauses describe an event that is unrealized at the moment of the main event, which can be indicated by a language's grammar. In Sinhala, purpose can be expressed with the use of the infinitive verb form without any other subordinating particle[5]. The example below shows this. Note that the adverbial clause does not have an expressed A argument and therefore represents a general A.

(4)   Chinese New Year, Sentence 4 (Santa Barbara)
    ciina   alut   aurudə   samэra-nnə
    Chinese   new   year.SG.DEF   celebrate.INF

    wiwidə   wiwidə   saŋdaršэnэ   saha   perэhærэ   pawat-wэnэwa
    various   various   show.PL   and   parade.PL   hold-CAUS
    'To celebrate the Chinese New Year various shows and parades are held.'

---

[5] Sinhala also has subordinating particles that are equivalent to the English 'in order to', *pinisə* and *sandaha,* which are used with other verb forms, but these do not appear in examples as they did not appear in the data collected.

REASON. As was noted earlier, purpose and reason are often expressed with the same morphology, but some languages use an adverbial particle that explicitly expresses a causal relationship. Sinhala is an example of the latter type of language, as can be seen with the example below.

(5) <u>Frog Story, Lines 15-16 (Rice)</u>

ballaṭə		dæn	kohoməhari	oluwə	eliyəṭə	ga-nnə	wiḍiyak
dog.SG.DEF.DAT		now	somehow	head.SG.DEF	out.DAT	take-INF	way.SG.IND

næti	hinda	balla	daŋgələ-la	daŋgələ-la	janeelen
NEG.ADJ	because	dog.SG.DEF	fidget-PPL	fidget-CONV	window.SG.DEF.ABL

eliyəṭə	pænn-a
out.DAT	jump-PST

'Now, because the dog had no way of taking his head out, the dog kept fidgeting about and jumped out the window.'

CIRCUMSTANTIAL. Circumstantial adverbial clauses provide information about how the event expressed by the matrix clause came to be. In English, this is usually expressed with the adverbial particles 'by' or 'without'. The one example of a circumstantial adverbial clause in the Sinhala data uses a time adverbial particle, but as Thomspson and Longacre point out, often time clauses and cause clauses are conflated. This example appears below.

(6) <u>Tsunami, Sentence 1 (Santa Barbara)</u>

mamə	sunaamiə	gænnə	šrilaŋkawe	saha
1SG	tsunami.SG.DEF	about	Sri.Lanka.LOC	and

aasiyawe	sunaamiə	gænnə	dænə-gatee
Asia.LOC	tsunami.SG.DEF	about	to.know-REFL.FOC.PST

əntarjaaləyæ	pwuətpatak	kiyəwə-nnə	gamaŋ
on.the.internet	newspaper.SG.DEF	read-INF	while

'I got to know about the tsunami in Sri Lanka and Asia while reading a newspaper on the Internet.'

SIMULTANEOUS. According to Thompson and Longacre, when two events co-occur at the same time, then languages provide a mechanism by which speakers can express that one is the backgrounded event that provides the context for the main event. This can be done one of two ways—either with a marker explicitly indicating simultaneity or with an aspect marker. As can be seen in the example below, in Sinhala, both are used together—a free adverbial particle indicating simultaneity is used along with the present adjectival form of the verb.

(7) <u>Frog Story, Lines 24-25 (Rice)</u>

laməya	gembawə	bimə	hoyə-nə	koṭə
child.SG.DEF	frog.SG.DEF.ACC	ground	search-ADJ	while

miiyek	lamǝyage	nahayǝ	hæpuw-a
mouse.SG.IND	child.SG.DEF.GEN	nose.SG.DEF	bite-PST

'While the child was searching for the frog on the ground, a mouse bit the child's nose.'

CONDITIONAL.   Most languages mark two kinds of conditional—reality conditional and unreality conditionals.   The former refers to real, habitual or past situations, whereas the latter refers to those events that we imagine or predict (Thompson and Longacre 1985). Syntactically, these events are usually represented with an adverbial particle such as 'if', as with English.   In Sinhala, however, conditional statements are marked only by verbal suffixes— -*ot* in the non-past tense and –*naŋ* in the past tense.   An example of each of these appears below.

(8)   Elicited

Present:

ohu	væṭun-ot	maṭǝ	ohuwǝ	alla-nnǝ	puluwaŋ
3SG.M	fall-COND.PRES	3SG.DAT	3SG.M.ACC	catch-INF	can

'If he falls, I can catch him'

Past:

oyaa	hoňdǝṭǝ	vædǝ keraa-naŋ	hoňdǝ	lakunu	ga-nnǝ	tibun-a
2SG	good.DAT?	work.do-COND.PST	good	grade.PL	get-INF	keep-PST
						could.have

'If you had worked hard, you would have gotten good grades.'

CONCESSIVE.   Concessive adverbial clauses mark a concession against which the matrix clause is contrasted.   According to Thompson and Longacre, there are two general subcategories within the broader category of concessive—definite and indefinite.   Definite concessive clauses are usually marked by an adverbial particle like 'although' and can be identified by the fact that they can be paraphrased by the statement 'in spite of the fact that …' (note the definite noun phrase).   Indefinite concessive clauses are those which indicate the sense of 'no matter what' or 'whatever'.

As with the conditional clauses in Sinhala, the concessive adverbial clauses are marked by verbal morphology alone, but this is true only in the past tense.   If the verb in the adverbial clause takes the non-past marker then it must be followed by the adverbial particle *wunat* 'even though'.   This can be seen with the data below.

(9)   Elicited

Present:

balla	mas	ka-nǝwa	wunat	apee	kukulaṭǝ
dog.SG.DEF	meat	eat-IMPF	even.though	1PL.GEN	chicken.SG.DEF.DAT

haani	kera-nne	næǽ
harm	do-INF	NEG

'Although our dog eats meat, she won't take our chicken'

Past:	balla	mas	kææw-at	apee	kukulaṭə	haani
	dog.SG.DEF	meat	eat-CONC.PST	1PL.GEN	chicken.SG.DEF.DAT	harm

	kera-nne	nææd
	do-INF	NEG

'Although our dog ate meat, she wouldn't take our chicken'

SUBSTITUTIVE. Substitutive adverbial clauses express a relationship in which the matrix clause event replaces the adverbial clause event, the former being the unexpected event and the latter the expected one. This is expressed in English with 'instead of' and 'rather than'. In Sinhala, substitutives are constructed by using the morpheme *nætuwə* 'without'. Interestingly, sentences constructed in this way can mean either that the event in the adverbial clause was replaced by that in the matrix clause or that both events were supposed to occur, but the one in the adverbial clause did not occur.

(10) Elicited

æwidi-nnə	ya-nne	nætuwə	ohu	tiwi	bæluw-a
walk-INF	go-FOC.PRES	without	3SG.M	TV	watch-PST

'He watched TV instead of going for a walk' or 'He watched TV without going for a walk'

ADDITIVE. Some languages have morphology that indicates a relationship in which one event occurs in addition to another. In English, phrases such as 'in addition to' and words like 'besides' are used to express this relationship. Despite the fact that Sinhala has converbal affixes on verbs, it is possible to construct an additive adverbial clause using the subordinating morpheme *amətərəwə* and a special verb form with the suffix *-məṭə*. A Sinhala example is provided below.

(11) Elicited

keek	picci-məṭə	amətərəwə	ohu	kukis	hadə-nəwa
cake	bake-?	in.addition.to	3SG	cookies	make

'In addition to baking a cake, he is making cookies'

ABSOLUTIVE. The interpropositional category absolute is a broad category, which must meet the following conditions (Thompson and Longacre 1985:200-201):

1. The clause is marked in some way as being subordinate
2. There is no explicit signal of the relationship between the main and subordinate clause
3. The interpretation of the relationship is inferred from the pragmatic and linguistic context.

These clauses are used when there is no need to explicitly specify how the main and adverbial clauses are related. They can be identified by special marking on the verb (often nominalization) and a general adverbial particle. In Sinhala, this can be accomplished with converbs, as is seen in the example below (cf. Taylor current volume).

(12)   Andare Sugar Story, Sentence 17 (Santa Barbara)
      itiŋ   andəre   ehemə   putaaʈə   kiyə-la
      therefore   Andare   in.that.way   son.SG.DEF.DAT   say-PPL

      aayet   maaligaawəʈə   giy-a
      again   palace.SG.DEF.DAT   go-PST

      'Therefore, having said that to his son, Andare went to the palace.'

2.2. DISCOURSE ROLES. Thompson and Longacre's (1985) discussion of the discourse roles of adverbial clauses points to two main functions. When an adverbial clause is predicated (through lexical overlap) with another clause in the story, its function is to aid in the progression of the narrative to its goal. When it is not predicated, its function is usually to contribute information that is only relevant to the matrix clause that it modifies.

Ramsay's (1985) findings support and extend Thompson and Longacre's. In her work on preposed versus postposed adverbial clauses in English, she finds a relationship between position and function. Those adverbial clauses that appear before the matrix clause act as a cohesive device, advancing the narrative. On the other hand, those that appear after the matrix clause are only locally significant, completing the information provided in the matrix clause.

Related to the concept of local relevance is work on the role between foreground and background information and independent versus dependent clauses. While this work has been fairly controversial, as the definitions and determinations of foreground versus background information are fuzzy, there does seem to be a relationship between clause type and narrative role. Tomlin (1985) tries to operationalize foreground and background, and he shows that adverbial clauses do tend to contain background information. When looking at the findings discussed earlier, this result is not surprising, as clauses that are only locally relevant and do not advance the narrative are likely to be background as well.

3. BRIEF OVERVIEW OF FOCUS IN SINHALA.

3.1. THE SYNTACTIC STRUCTURE OF SINHALA FOCUS. The primary means by which Sinhala brings a certain element into focus is the use of tense-based focus morphology on the verb (-*nne* if the verb is in non-past and –*e* if the verb is in the past). The focused element then usually appears postverbally, causing a shift in the more typical SOV constituent order (and hence the characterization of Sinhala as having variable constituent order). The example below shows a sentence with different constituents focused.

(13)   Basic: nimal   dælak   ekkə   maalu   allə-nəwa
              Nimal   net.SG.IND   with   fish.PL   catch
              'Nimal is catching fish with a net' (as a general statement of fact)
   a.   Focus: nimal   dælak   ekkə   alla-nne   maalu
              Nimal   net.SG.IND   with   catch.FOC.PRES   fish.PL
              'It is fish that Nimal is catching with a net'
   b.   Focus: nimal   maalu   alla-nne   dælak   ekkə
              Nimal   fish.PL   catch.FOC.PRES   net.SG.IND   with
              'It is with a net that Nimal is catching fish'

    c. Focus: dælak   ekkə maalu <mark>alla-nne</mark>   nimal
               net.SG.IND with fish.PL catch.FOC.PRES Nimal
               'It is Nimal that is catching fish with a net'

In each of the sentences above, the verb is specially marked for focus, and the focused element—'the fish', 'with a net', and 'Nimal' respectively—follows. Oftentimes, though, if the focused element appears before the verb or *in situ* then it is marked with a focus morpheme such as *tamai* 'indeed' (see the example below). This is not, however, always the case, and it is even possible for *tamai* to appear after a post-verbal focused element. The reasons behind the use of *tamai* or lack thereof appears to be discourse-based and needs to be investigated further.

(14) <u>Tsunami, Sentence 13 (Santa Barbara)</u>
      itiŋ   ehemə   tamai   maŋ sunaamiyə   gænə   muliŋmə
      therefore in.that.way indeed 1SG tsunami.SG.DEF about first

       dænəgatte
       know-REFL.FOC.PAST

    'Therefore, that was how I first got to know about the tsunami.'

The focus constructions discussed in this paper refer to those cases in which the focus morphology appears on the verb, regardless of whether or not the focused element appears with *tamai*.

3.2. THE VARIOUS ROLES OF SINHALA FOCUS. Gair (1998 [1985]) points out that although Sinhala focus may have derived from contact with neighboring Dravidian languages, the use of focus has diversified and become a more central part of Sinhala grammar since that historical contact. This section of the paper discusses some of the main areas in which focus forms can be found[6].

PRESENTATIONAL AND CONTRASTIVE. The two most common typological functions of focus constructions are presentational and contrastive. In both cases, the focused element is something the hearer is assumed not to know, either because it is new (presentational) or because it contradicts what the hearer presupposes (contrastive). In Sinhala, both of these structures appear alike syntactically and are distinguished only by context. An example of each is provided below.

     Presentational
(15) <u>Chinese New Year, Sentence 2 (Santa Barbara)</u>
     mamə wædəkar-ee waarta karuwek hæṭiyəṭə
     1SG   work-FOC.PST report do-NOM as

    'I worked as a reporter' (new information)

---

[6] Note that this paper only includes examples with focus morphology on the verb. There are other ways of expressing focus, but that is not covered here.

Contrastive
(16) <u>Elicited</u>
ṭikeṭekak   gatt-e   nimal
ticket.SG.IND   take-FOC.PST   Nimal
'Nimal bought a ticket.' (as opposed to another person)

OTHER.   In addition to the discourse-based use of focus in declarative clauses, focus structures have developed as obligatory elements in questions and negations and have also grammaticized in some common collocations (i.e. epistemic stance using *maŋ hitanne* 'I think', equationals using focus forms of *kiyanna* 'to say', and locationals using focus forms of *tiyenna*). They also appear with certain adverbial clauses. While the literature on Sinhala focus has addressed many of the functions of Sinhala focus, the role of focus with adverbial clauses has yet to be discussed (Gair 1970, 1998 [1983], 1998 [1985]; Gair and Paolillo 1997; Gair and Lelwala 1991; Herring and Paolillo 1995; Kariyakarawana 1998; Paolillo 1994). Providing a preliminary analysis to discover the motivation governing the use of focus forms with adverbial clauses is the goal of the current paper. For this reason, the following section covers adverbial clause focus constructions in greater detail.

4. FOCUSED ADVERBIAL CLAUSES AND SINHALA DISCOURSE.
4.1. DATA. The data used in the current paper come from a collection of eleven stories of lengths varying from 1½ to 4 minutes. These stories were elicited from two different speakers in two separate field methods classes, one at Rice University and the other at the University of California, Santa Barbara. A total of 50 sentences with adverbial clauses appear in these eleven stories. In the analysis phase, all of the adverbial clauses and their respective matrix clauses were analyzed and classified according to the following features:

1. Relative order of matrix and adverbial clauses
2. Semantics of interpropositional relationship (cf. Thompson and Longacre (1985))
3. Presence of focus morphology on adverbial clause verbs
4. Information status of event/state of adverbial clause
5. Predication of the event/state of adverbial clause in preceding or following sentences[7]
6. Presence of focus morphology on matrix clause verbs
7. Information status of event/state of matrix clause
8. Predication of the event/state of matrix clause in preceding or following sentences

With respect to information status, all of the adverbial clauses were coded according to whether they were New, Given, or Inferred following Chafe's (1976) definitions in which New refers to information which the speaker assumes the addressee is not expected to know at that point, Given information is that which the speaker assumes to be in the addressee's consciousness, and information that is Inferred may not be directly in the speaker's consciousness but can be easily accessed from given context (i.e. that someone was tired can

---

[7] Predication was measured by whether or not the event was mentioned in an earlier or later sentence. The reason that event and sentence was chosen is that both of these represent complete ideas, and it was necessary to see if the complete idea referenced in the adverbial clause was referenced elsewhere.

be inferred if it is known that s/he rested). To illustrate the information status distinction, as well as each of the other categorizations, an example from the stories appears below with the respective analysis.

(17) <u>Chinese New Year, Sentence 1 (Santa Barbara)</u>

mamə	aurudu	tunəkaṭə	issella	santabarbarawaṭə	e-nnə	kaliŋ
1SG	year.PL	three.IND.DAT	before	Santa.Barabra.?.DAT	come-INF	before

haŋkaŋ	welə	aurudu	dekak	wæḍəkər-a
Hong.Kong	LOC.PL	year.PL	two.IND	work-PST

'Before I came to Santa Barbara three years ago, I worked in Hong Kong for two years.'

In this example, the sentence is the very first in the narrative. The interpropositional relationship between the matrix and the adverb is along a time dimension (i.e. 'before'). The intrasentential ordering of the clauses is adverbial then matrix. Neither the verb in the adverbial clause nor the one in the matrix clause have focus morphology. The event in the adverbial clause (coming to Santa Barbara) can be inferred from the context, as the speaker is telling the story in Santa Barbara but is originally from Sri Lanka. This event, however, is not mentioned anywhere else in the narrative and is thus not predicated by another sentence. On the other hand, while the event described in the matrix clause is new information, it is referred to in the very next sentence of the discourse. This information is summarized in the table below.

1. Intra-sentential order: Adverbial, Matrix
2. Adverbial type: Time
3. Adverbial verb focus: No
4. Adverbial info. status: Inferred
5. Adverbial predication: None
6. Matrix verb focus: No
7. Matrix info. status: New
8. Matrix predication: First Next

TABLE 3. Summary of analysis of sentence in example 28

As was noted, the example just presented does not have a verb with focus morphology in either the matrix or the adverbial clause, but it was mentioned in the section on focus constructions in Sinhala that adverbial clauses are one of the instances in which focus can be seen; and in the data collected for this investigation, there were a total of 7 sentences with adverbial clauses in which a verb carried the focus morpheme (one of which was eliminated from the analysis because it involved negation, which necessitates the use of the focus form of the verb). Therefore, it is worthwhile to determine what, if any, features of sentences with an adverbial clause call for a focus construction. The analysis prepared for this paper indicated some general patterns for those particular sentences, and these patterns will be addressed in the next section.

4.2. GENERAL PATTERNS OF ADVERBIALS IN FOCUS CONSTRUCTIONS.   When comparing the characteristics of the sentences with both adverbial clauses and focus constructions, certain patterns emerged. The table below provides the information concerning the characteristics of the six relevant sentences found in the data.

From the data in the table above, there are patterns that become evident. In the row concerning Adverbial Clause Type, we see that the interpropositional relationship in these cases is predominantly reason. There are only two cases which differ, one which is purpose-based and the other which is circumstantial-based; however, the semantics of purpose and circumstantial are related to reason (note that reason and purpose are often represented with the same morphology because they 'can be seen as providing EXPLANATIONS for the occurrence of a given state or action' (Thompson and Longacre 1985:185, emphasis in original)). As a result, all of these can be subsumed under the category 'explanatory'.

With respect to Clause Order, all of the examples provided appeared in the order matrix clause followed by adverbial clause. This is not surprising if we consider the most typical word order in focus constructions noted in section 3.1.2., i.e. focused element following the focus form of the verb. In these cases, therefore, the adverbial clause is the focused element, taking its expected post-verbal position. It is also for this reason that all of the matrix verbs are focused, whereas those in the adverbial clause are not (indicated in the table by 'Yes' in the row 'Matrix Verb Focus' and 'No' in the row 'Adverbial Verb Focus').

The remaining columns concern the information status of the clauses and whether the situations (events/states) of the clauses are predicated elsewhere in the discourse. All of the matrix clauses in these examples refer to an event that is expected to be known by the hearer, either because it was given in the previous discourse or because it can be inferred from the circumstances in which the story was told, and all except for one is predicated by the sentence just preceding them. In addition, none of these clauses are mentioned again in the rest of the story. On the other hand, when looking at adverbial clauses, all of the entries except one are both new and not predicated by any other sentence. The one exception, entry 6, involves an adverbial clause whose event is given in the preceding sentence, and it will be discussed in the section concerning exceptions.

The correlates presented in the table above provide an impetus for determining the functional motivations for focusing these particular adverbial clauses. These motivations are the focus of the following section.

4.3. IMPLICATIONS OF THE PATTERNS. The data just presented indicated that adverbial clauses in focus constructions tend to share the following characteristics: 1) An explanation-based interpropositional relationship, 2) An intrasentential order of matrix followed by adverbial clause, 3) A matrix clause that is expected to be known by the hearer because it was mentioned in the just preceding sentence or because it can be inferred from the circumstances, and 4) An adverbial clause that is both new and unique in the discourse. In the following two subsections, the relationship among these characteristics is examined and the exception to these patterns mentioned earlier is explained in light of this relationship. The final subsection presents evidence for the uniqueness of the characteristics of focused adverbial clauses by comparing them with the other adverbial clauses found in the data.

	1	2	3	4	5	6
Story	*Andare*	*Andare*	*Mahadænəmutta*	*Tsunami*	*Vesak*	*Yaalə*
Clause Order	Mat, Adv	Mat, Adv	Mat, Adv	Mat, Adv	Mat, Adv	Mat, Adv
Adverbial Clause Type	Purpose	Reason	Reason	Circumstantial	Reason	Reason
Adverbial Verb Focus	No	No	No	No	No	No
Adverbial Information Status	New	New	New	New	New	Given
Adverbial Predication	None	None	None	First Next	None	First Prev.
Matrix Verb Focus	Yes	Yes	Yes	Yes	Yes	Yes
Matrix Information Status	Given	Given	Inferred	Inferred	Given	Given
Matrix Predication	First Prev.	First Prev.	First Prev.	None[8]	First Prev.	First Prev.

TABLE 4. Summary of the features of sentences with adverbial clauses and focus constructions[9]

[8] Despite not being mentioned earlier (as it is the first sentence in the story), this matrix clause is considered inferred because the prompt was 'Tell me how you learned about the tsunami', hence making the hearing about the tsunami (expressed in the matrix clause) given.

[9] As was noted earlier, the sentence number in the discourse was also examined, but does not appear to play a role, as focused adverbial clauses can occur anywhere in a discourse.

THE CORRELATIONS. One of the key related features of sentences with both adverbial clauses and focus constructions concerns the clause order and the verbal focus forms. A common placement for focused constituents in Sinhala focus constructions is postverbal. The fact that we find adverbial clauses following matrix verbs with focus forms indicates that the same behavior occurs with elements larger than a phrase, namely with clauses that bear a relationship to the matrix verb. Therefore, Sinhala speakers have the ability to focus an even broader range of elements.

The question then arises what would motivate a Sinhala speaker to use a focus construction with an adverbial clause, and it is here that we see how information status plays a role. The first thing to note is that there is a difference in the information status of the two clauses with the adverbial clause expressing new information, a characteristic not found in those sentences with a non-focus-marked matrix verb (a point that is addressed in §4.3.3). More specifically, the event in the matrix clause is given whereas the event in the adverbial clause is new. As was discussed in both the introduction and the section on the types of Sinhala focus structure, a primary function of focus in languages is to point to new participants. Extending what was found earlier concerning the extrapolation of post-verbal focused participants to post-verbal focused clauses, we can argue that a similar extrapolation is occurring here. Namely, it is not just new participants of a discourse that are focused, but new information as a whole, including new states and events. Furthermore, the interpropositional relationship between the clauses serves to explain the event in the matrix clause. This is expected because in these cases, we have focused new information that modifies only a particular given event as opposed to modifying the narrative at large, and such information is likely to provide an explanation—hence its use with purpose, reason, and circumstantial interpropositional relationships.

In order to illustrate how the correlations work, it is worthwhile to look at some of the examples from the data. In example 19, we see the very beginning of the story of Mahadaenemutta. The sentence of interest is the second, but the first and third have been provided for context.

(18)   Mahadaenemutta (Santa Barbara)

Sentence 1: ekomatekədawəsəkə   laŋkaawe   dakunu   prədeešəye   kæægallə
             once.upon.a.time   (Sri)Lanka.LOC   south   province.SG.DEF   Kaegalla

             kiyən-a   nagərəye   mahadænəmutta   kiye-la
             know.as-PST   town.SG. DEF.LOC   Mahadaenemutta   know.as-PPL

             siiyakenek   hiṭiy-a
             grandfather.person.SG.DEF   exist-PST
                     old.man

             'Once upon a time in the town known as Kaegalla in Sri Lanka's Southern province there was an old man knows as Mahadaenemutta.'

Sentence 2: mahadænəmuttaṭə   ehemə   mamə   aaw-e   eya   itaamat
             Mahadaenemutta.DAT   that.way   1SG   come-FOC.PST   3SG   very

             ugat   pudgələyek   hæṭiyəṭə   gamee   minisu
             wise   person.SG.IND   as.DAT   village.SG.DEF.LOC   man.PL

> | saləkə-pu | nisaa |
> |---|---|
> | consider-PST.ADJ | because |
>
> 'The name came in that way to Mahadaenemutta because the people of the village considered him to be a very wise person.'

Sentence 3:

mahadǽnəmuttaṭə	gooləyo	pasdenekut	hiṭiy-a
Mahadaenemutta.DAT	follower.PL	five.people.and	exist-PST

'Mahadaenemutta also had five followers.'

In this example, the character of Mahadaenemutta is introduced in the first sentence. Because the hearer can be expected to know from the previous sentence that the man had been given the name Mahadaenemutta, the new information in the clause is the circumstances or reasons that led to the giving of his name. This contrast in information status is represented syntactically by the focusing of the adverbial clause, represented with a focus morpheme on the verb and the immediately postverbal position of the adverbial clause. Notably, the state described by the adverbial clause is not mentioned in the following sentence, and in fact is not mentioned at any other place in the story.

In example 20, the story of Andare and his eating of the sugar in front of the palace has just begun. The relevant sentence to the current study is the fifth sentence of the story. The first three sentences establish respectively the existence of the jester Andare, that he usually worked at the king's palace, and that workers in the king's palace were treated very well by the king. The fourth, fifth, and sixth sentences appear below (as with the earlier example, the surrounding sentences provide context for the sentence under study).

(19)   <u>Andare Sugar Story (Santa Barbara)</u>

Sentence 4:

dawəsak	da	andəree	rajjəmaaligaawəṭə	weḍəṭə	udee
day.SG.IND	?	Andare	palace.SG.DEF.DAT	work.DAT	morning

ya-nə	koṭə	maaligaawə	issərəha	siini	goḍak	elə-la
go-ADJ	while	palace.SG.DEF	in.front	sugar	pile.SG.IND	spread-PPL

tiye-nəwa	andəree	dǽkk-a
keep-IMPF	Andare	see-PST

'One day, while Andare was going to the palace to work in the morning, he saw a pile of sugar spread out in front of the palace.'

Sentence 5:

ee	siini	ehemə	elə-la	tibbe	weele-nnə
DIST	sugar	that.way	spread-PPL	keep-FOC.PST	dry-INF

'The sugar was spread in that way to be dried.'

Sentence 6:

mokədə	dawas	kiipəyəkəṭə	issəlla	huŋgak	wǽssə	ewi-la
because	day.PL	few.SG.IND.DAT	before	lot.SG.IND	rain	come-PPL

siini	malu	huŋgak	temi-la
sugar	sack.PL	lot.SG.IND	wet-PPL

'Because a few days ago a lot of rain came and many sacks of sugar got wet.'

The first sentence in example 30 establishes the foundation upon which the rest of the story will be built by introducing the sugar that Andare will soon eat. The next sentence provides an explanation for the unexpected spreading of the sugar on the ground. In this sentence, the adverbial clause expressing purpose is focused and immediately follows the focused verb. As with the preceding example, this focusing is done with the focus form of the verb in the matrix clause and the post-verbal position of the adverbial clause[10]. The reason for the focusing of the adverbial clause in this example is again a difference in information status and predication. The information in the matrix clause of sentence 5 is introduced in the immediately preceding clause, whereas the purpose explanation provided by the adverbial clause is not mentioned elsewhere in the story. In addition, this explanation is only relevant to the sentence to which it belongs.

As is seen with the above examples, the relationships among the features of the adverbial clauses in focus constructions also support the findings of both Ramsay (1985) and Thompson and Longacre (1985) concerning the discourse roles of adverbial clauses. As was discussed earlier, Ramsay's study found that in English, a difference in position of the adverbial clause reflected a difference in discourse function, with one position indicating a more limited focus, elaborating the matrix clause, and another position acting as a means to advance the narrative. This point is made again by Thompson and Longacre, as they note that adverbial clauses that share an intraparagraph relation involve paraphrasing another element of the paragraph, whereas those that do not only contribute local background to the surrounding sentence. The adverbials in this study all appear after the matrix clause. In addition, they all provide new information that modifies a matrix clause containing an event that is already known, thus limiting the scope of the matrix. As a result, these adverbial clauses are not re-predicated, as they do not constitute a significant event that advances the plot. Therefore the results here support both of these studies.

THE EXCEPTION. The correlates just discussed were consistent among all of the examples except for the one example from the *Yaale* story, in which the event in the adverbial clause is given in the preceding sentence, resulting in a lack of difference in information status between the matrix and the adverbial clauses. This particular sentence is the last sentence of the actual narrative. It appears in the example below along with its surrounding sentences.

(20)   Yaale (Santa Barbara)
Sentence 11:   mæturuwa-iŋ   passe   ee   aliya   ekəpaarəṭəmə
               chant-?   after   DIST elephant.SG.DEF one.?.EMPH
                                                             suddenly

               wanəyaṭə   aayet   diuw-a
               jungle.SG.DEF.DAT again   run-PST

               'After he chanted, that elephant suddenly ran again into the jungle.'

---

[10] It needs to be noted that the reason clause following the infinitive verb *weelennə*, beginning with *mokədə* 'because', is actually not an adverbial clause but a separate sentence that has dropped the inflected verb *tibba* 'place-PST' from the end.

Sentence 12:   itiŋ   meekə   apiṭə   itaamat pudumə   awastaawak
                 therefore   1PROX.SG.INAN 1PL.DAT very   incredible occasion.SG.IND

'Therefore this was to us a very incredible occasion'

Sentence 13:   mukədə ekə atəkiŋ   apiṭə   pee-nne   eekə
                 because one hand.SG.IND.ABL 1PL.DAT see-FOC.PRES DIST.SG.INAN

ee   aliyə   aayet wanəyaṭə   diuw-e
DIST elephant.SG.DEF again jungle.SG.DEF.DAT run-FOC.PST

ṭrækə	mahattəya	maturə-pu	hinda-i	kiyə-la
tracker	gentleman.SG.DEF	chant.PST.ADJ.	because-COMP	say-PPL
				COMP

anit   pættəṭə   apiṭə   mætir-iimə gænə kisimə
other side.SG.DEF.DAT 1PL.DAT chant-NOM   about any

wišwaasəyəkut   næhæ
belief.SG.IND.and NEG.have

'Because on the one hand we saw the elephant run into the jungle again because the ranger chanted; on the other hand, we had no belief about chanting'

Sentence 14:   itiŋ   eekə   tamai   magee   keṭi   kataawə
                 therefore DIST.SG.INAN indeed 1SG.GEN short story.SG.DEF

'Therefore this is my short story.'

In order to determine why this particular sentence differs from the others in the collection, it is necessary to establish the motivations behind the use of the focus form here. The storyteller begins this sentence stating that the event just mentioned was *itaamat pudumə awastaawak* 'a very incredible occasion'. This is a key statement in determining what is happening with the sentence under study, as the simple running of the elephant into the jungle would not be remarkable on its own. Rather, it is the fact IT WAS BECAUSE THE RANGER CHANTED that the elephant ran into the jungle that is noteworthy in this narrative. Therefore, what we see here is another function of focus forms, namely highlighting an unexpected, thus noteworthy, interpropositional relationship. The unexpected reason relationship between the two events is what is important. The focus is on the entirely unexpected causal interpropositional relationship between the two events. This provides further evidence that not only can participants be highlighted, but events and their interrelationships may as well.

CHARACTERISTICS OF NONFOCUSED ADVERBIAL CLAUSES. As has been shown, all sentences with focused adverbial clauses share particular features. The question that must now be addressed is whether or not these features are unique to sentences with focused adverbial clauses in the data collected. In order to establish that this is indeed the case, it is necessary to examine the characteristics of non-focused adverbial clauses and compare them with focused adverbial clauses, specifically looking at: 1) intrasentential order of matrix and adverbial clause, 2) information statuses of matrix and adverbial clauses, and 3) interpropositional relationship.

Because of the nature of focus structures, namely the typical post-verbal position of the focused element, it is no surprise that all of the focused adverbial clauses follow their matrix clause. It is also to be expected that in sentences with non-focused adverbial clauses, the order will likely be an adverbial clause followed by the matrix, and this is what is demonstrated in the data. In all but three sentences with non-focused adverbial clauses, the adverbial clause appears first.

Due to the fact that the positioning of focused adverbial clauses can so easily be related back to the syntax of focus as a whole, the remaining two characteristics—information status and interpropositional relationship—are more central to determining whether or not the focused adverbial clauses have a special discourse purpose. In the first case, it is important to determine whether or not there is a distinction with respect to the information status of the matrix and adverbial clauses. The table below provides data for the four possible permutations[11] of information status for both nonfocused and focused adverbial clauses.

Information Status Matrix/Adverbial	Non-focused Adverbial Clauses	Focused Adverbial Clauses
New/New	10	0
New/Given	26	0
Given/New	2	5
Given/Given	5	1
TOTAL	43	6

TABLE 5. Information Status of Matrix and Adverbial Clauses

As is evidenced in the table above, sentences with non-focused adverbial clauses predominantly have new information in the matrix clause and very frequently given information in the adverbial, whereas focused adverbial clauses always have a given event in the matrix clause with the adverbial clause primarily containing new information. Note that in most cases, regardless of focus, the information status of the matrix clause and adverbial clause are opposite of one another. The fact that sentences with focused adverbial clauses have an inverse information status relationship to those with nonfocused adverbial clauses is not surprising when considering the discourse role of these clauses. As was noted earlier, the focused adverbial clauses tend to have a limited scope, only modifying their respective matrix clause (and are hence often new with given matrix clauses). On the other hand, given adverbial clauses tend to act as narrative ties, linking previous events with a new event in the matrix clause (cf. Ramsay's (1985) findings on postposed and preposed adverbial clauses and Tomlin's (1985) work on adverbial clauses and foreground and background). It is also worth noting that both cases in which a nonfocused adverbial clause was new while the matrix clause was given were both expressing an interpropositional relationship of simultaneity, which points to the significance of interpropositional relationship.

Although all of the focused adverbial clauses had an explanatory relationship with their matrix clause (either purpose, reason, or circumstantial), it is not the case that only focused adverbial clauses have such a relationship, as the data indicate that nonfocused adverbial clauses may also share an explanatory relationship with their matrix clause. However, nonfocused adverbial clauses are far less restricted than focused adverbial clauses, as they are

---

[11] For ease of reference, Inferred information status is collapsed with Given in this table.

able to express a variety of interpropsositional relationships in addition to cause. This is seen in the table below.

Interpropositional Relationship	Non-focused Adverbial Clauses	Focused Adverbial Clauses
Explanatory:		
Reason[12]	6	4
Purpose	4	1
Circumstantial	1	1
Simultaneous	21	0
Time	8	0
Absolutive	3	0
TOTAL	43	6

TABLE 6. Interpropositional Relationships of Adverbial Clauses

Table 6 indicates that while focused adverbial clauses are limited to explanatory interpropositional relationships, non-focused adverbial clauses have more varied functions. Indeed, nonfocused adverbial clauses appear to predominantly express temporal relationships between the events in the matrix and adverbial clauses. However, there is overlap with respect to explanatory interpropositional relationships, so it is worthwhile to determine what, if anything, sets focused adverbial clauses apart from their non-focused counterparts when considering only explanatory interpropositional relationships.

	Non-focused Explanatory Adverbial Clauses (Total = 11)	Focused Explanatory Adverbial Clauses (Total = 6)
Clause Order		
Mat/Adv	2	6
Adv/Mat	9	0
Information Status Matrix/Adverbial		
New/New	3	0
New/Given	6	0
Given/New	0	5
Given/Given	2	1

TABLE 7. Explanatory Adverbial Clauses

From this table, it is clear to see that focused adverbial clauses must have a matrix clause in which a given event is expressed followed by an adverbial clause that is most often new

---

[12] There is an additional adverbial clause with a reason interpropositional relationship, but as it also is a negative sentence, thence requiring focus verbal morphology, it is not possible to determine if the adverbial clause is focused or not, so it is not included in the count.

information (depending on the function of the adverbial clause). On the other hand, there are only two nonfocused explanatory adverbial clauses that follow their matrix clauses. Interestingly, these two are also the two in which both the matrix and adverbial clauses express given events. While this points to another intriguing area of study, for the current purposes, it is important to note that only focused explanatory adverbial clauses have a matrix concerning a given event followed by an adverbial concerning a new event.

The data presented in this section show that for each independent feature of focused adverbial clauses, that feature is predominantly expressed by the focused adverbial clauses and almost absent in nonfocused adverbial clauses. More importantly, however, when considering the intersection of all three features, we find that only the focused adverbial clauses simultaneously have a clause order of matrix followed by adverbial, an explanatory interpropositional relationship, and a given adverbial clause describing a new matrix clause. However, it is important to recognize that due to the limited data set, the results discovered in this study are preliminary and further investigation is required to ensure that these results concur with larger sets of data.

5. CONCLUSION. Discourse is based on the interaction of two or more people, and in order for this discourse to flow smoothly, participants keep track of one another's state of knowledge so as to provide just the right amount of information. One way in which this is indicated in the grammar is through the use of information packaging mechanisms such as focus. Most of the literature concerning focus attends to the fact that focus constructions are used to introduce sentence participants that are either new to the hearer or contradictory to his/her presuppositions. However, one mechanism that languages can use to introduce new information that will be only locally relevant is through adverbial clauses, so it should be possible for these elements of a sentence to be in focus as well. In the current paper, it is shown that Sinhala speakers do just this. The evidence provided indicates that in the cases examined here, adverbial clauses become the focused element of a sentence when they provide new information about a matrix clause that contains given information. Moreover, this only occurs when the adverbial clause is only relevant to the immediate sentence as opposed to the surrounding narrative.

REFERENCES

ANDREWS, AVERY. 1990. The major functions of the noun phrase. *Language Typology and Syntactic Description, vol. 1*, ed. by Timothy Shopen, 62-154. New York: Cambridge University Press.

CHAFE, WALLACE L. 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. *Subject and Topic*, ed. by Charles N. Li, 25-55. New York: Academic Press.

FERNANDO, WIROSHANA. 2005. Personal Communication (University of California, Santa Barbara Field Methods Course).

GAIR, JAMES W. 1970. *Colloquial Sinhalese Clause Structures.* Paris: Mouton.

GAIR, JAMES W. 1998 [1983]. Nonconfigurationality, Movement and Sinhala Focus. *Studies in South Asian Linguistics: Sinhala and Other South Asian Languages*, ed. by Barbara C. Lust, 65-563. New York: Oxford University Press.

GAIR, JAMES W. 1998 [1985]. Sinhala Focused Sentences: Naturalization of a Calque. *Studies in South Asian Linguistics: Sinhala and Other South Asian Languages*, ed. by Barbara C. Lust, 155-169. New York: Oxford University Press.

GAIR, JAMES W. 1998 [1989]. Sinhala Nonverbal Sentences and Argument Structure. *Studies in South Asian Linguistics: Sinhala and Other South Asian Languages*, ed. by Barbara C. Lust, 87-107. New York: Oxford University Press.

GAIR, JAMES W.; JOHN C. PAOLILLO. 1997. *Sinhala.* München: Lincom Europa.

GAIR, JAMES W.; LELWALA. SUMANGAIA. 1991. What to focus in Sinhala. In the proceedings of the *Eastern States Conference on Linguistics*, 93-108.

GIVÓN, TALMY. 1990. Contrastive Focus Constructions. *Syntax: A Functional-Typological Introduction, vol. II.*, 699-737. Amsterdam: John Benjamins Publishing Company.

HERRING, SUSAN C.; JOHN C. PAOLILLO. 1995. Focus position in SOV languages. *Word Order in Discourse*, ed. by Pamela Downing and Michael Noonan, 163-198. Amsterdam: John Benjamins Publishing Company.

KARIYAKARAWANA, S. M. 1998. *The Syntax of Focus and Wh-Questions in Sinhala.* Colombo, Sri Lanka: Karunaratne & Sons Ltd.

LUST, BARBARA C., ed. 1998. *Studies in South Asian Linguistics: Sinhala and Other South Asian Languages.* New York: Oxford University Press.

PAOLILLO, JOHN C. 1994. The Co-Development of Finiteness and Focus in Sinhala. *Perspectives on Grammaticalization*, ed. by William Pagliuca, 151-170. Amsterdam/Philadelphia: Benjamins.

RAMSAY, VIOLETA. 1985. The functional distribution of preposed and postposed 'if' and 'when' clauses in written discourse. *Coherence and Grounding in Discourse, Typological Studies in Language, vol. 11*, ed. by Russell Tomlin, 383-408. Amsterdam: John Benjamins.

SZENDRÖI, KRISZTA. 2004. Focus and the interaction between syntax and pragmatics. *Lingua* 114.229-254.

TAYLOR, CHRIS. 2005. Perfect, Sequence, Recapitulation: Common Construal and the Sinhala Conjunctive Participle. Santa Barbara Working Papers: Sinhala. 2005.

THOMPSON, SANDRA A.; ROBERT E. LONGACRE. 1985. Adverbial clauses. *Language Typology and Syntactic Description, vol. 2*, ed. by Timothy Shopen, 171-234. New York: Cambridge University Press.

TOMLIN, RUSSELL S. 1985. Foreground-background information and the syntax of subordination. *Text* 5.85-122.

VALLDUVÍ, ENRIC; ELISABET ENGDAHL. 1996. The linguistic realization of information packaging. *Linguistics* 34.459-519.

WICKREMASINGHE, NISSANKA SIRIMEVAN. 2005. Rice University Field Methods Course.

University of California, Santa Barbara
Department of Linguistics
3607 South Hall
Santa Barbara, CA 93106
vsultan@umail.ucsb.edu

PERFECT, SEQUENCE, RECAPITULATION:
COMMON CONSTRUAL AND THE SINHALA CONJUNCTIVE PARTICIPLE

Chris Taylor

Rice University

1. Introduction. In recent years, the term 'converb' has increasingly been used to describe constructions with cross-linguistically comparable forms and functions which include nonfinite verbal affixation, dependency on a finite verb, clause linking, and the sequencing of events (Bickel 1998, Genetti 2005, Hasplemath and König 1995, Masica 1991). Among such constructions, two broad areal categories have been suggested (Bickel 1998), European and Asian converbs, differing primarily with respect to the potential for what Bickel refers to as 'narrative chaining': Asian converbs perform a clause chaining function in addition to various types of modification, whereas the European type 'does not include chaining functions but rather stands in a binary relation to the main verb'. In the sparse typological converb literature, one form potentially instantiative of the former category comes from the Indo-Aryan language Sinhala, namely the conjunctive participle.[1] Examples include the following.

(1) booṭəle wætila kæḍuna
    bottle   fall-PPL break-PST
    'The bottle fell and broke.'

(2) galəkə   hæpila lamǝya-i   persgeḍi okomə bimə   wætuna
    stone-IND hit-PPL child-CONJ pears   all   ground fall-PST
    'After hitting the stone, the boy and the pears all fell to the ground.'

(3) siri wattə-ṭə gihilla pol   kaḍəla   wæṭak   bæňḍəla gedərə
    Siri estate-DAT go-PPL coconuts break-PPL fence-IND tie-PPL home
          giyaa
          go- PST
    'Siri went to the estate, picked coconuts, built a fence and went home.' (Gair and Paolillo 1997:49)

As examples (1)-(3) illustrate, the Sinhala conjunctive participle (which is morphologically marked by the suffix -*la*) performs several of the abovementioned functions characteristic of converbs. For instance, in each example the conjunctive participle expresses temporal sequence, and in (3) we observe narrative chaining, claimed to be characteristic of Asian converbs. Moreover, the verbal form exemplified here does not indicate time reference *per se*, and as such is less finite than the past tense form which occurs clause-finally.

Despite these similarities between the Sinhala conjunctive participle and Asian converbs, the former may also occur as a NONDEPENDENT PREDICATE when expressing perfect aspect, which distributionally appears to violate the converbal criteria of nonfiniteness and dependency (Genetti 2005, Hasplemath 1995; cf. Nedjalkov 1995 concerning the former). This function,

---

[1] Hereafter, the terms 'conjunctive participle' and 'participle form' will be used synonymously.

which Gair (2003) appropriately describes as 'an unusual if not unique feature among South Asian languages', is illustrated by (4) and (5) below.

(4)   mahattea   gihilla
        gentleman   go-PPL
        'The gentleman has gone.' (Gair 1970:153)

(5)   mamə   Renu-wə   dækka   habei   dæn   æyə   gihilla
        I   R-ACC   see-PST   but   now   3F.SG   go-PPL
        'I saw Renu but now she has gone.'

While in many respects the Sinhala conjunctive participle functions as a converb (per the definition put forward by Genetti 2005), utterances such as (1)-(5) illustrate a type of multifunctionality absent in similar South Asian verbal forms, namely, the functions of both nonfinite clause linking and nondependent predication. Faced with this duality of function, we must decide how best to characterize the relationship between the different uses.

One possibility is to analyze the two functions of the conjunctive participle as homonymous. On this view, the observed variation in use is taken to be indicative of two formally-identical morphemes with semantically-unrelated functions, i.e. the converbal functions illustrated by examples (1)-(3), and the expression of perfect aspect in main clause predication. Such an account is flawed, however, in that it fails to capture fundamental similarities in scene construal among the conjunctive participle's different uses, thus resulting in a missed generalization of descriptive significance.

In contrast, I will argue for a polysemy analysis of the Sinhala conjunctive participle. On this view, certain qualities of the construal traditionally accorded to the expression of perfect aspect are shown to crosscut the interpretations of the two aforementioned grammatical functions, nondependent predication and clause linking. Regarding the latter, I will discuss two specific functions—event sequencing and recapitulation—that provide evidence for analyzing the conjunctive participle as one form with related senses. Specifically, the analysis will demonstrate a parallel between a state's continued relevance to the speech act and the conceptual interrelatedness of certain event sequences.

The paper is structured as follows. After describing the data and methodology used for the study in Section 2, I provide a brief overview of the Sinhala conjunctive participle's form and functions in Section 3, each of which are subsequently discussed in Section 4. A summary of the findings common to each function follows in Section 5.

2. DATA AND METHODOLOGY. The data used in this study come from three sources, including published literature on Sinhala, structured elicitation, and a small corpus of ten texts. From the latter, which comprises nine narratives and one recipe, only those instances in which the conjunctive participle functions as a clause linker or nondependent predicate were included in the analysis. Such criteria were necessary in order to exclude other uses of the verb form in question which do not fall within the scope of the present study, such as its use in what Genetti (2005) refers to as 'conventionalized collocations', as well as the participial form of the verb *kiyannə* 'tell'—*kiyəla*—which functions as a complementizer. In total, 66 target instances of the conjunctive participle were collected from the corpus and coded for the following parameters: (1) the type of interpropositional relation (where relevant), (2) the number of conjunctive

participles in the turn, (3) the presence of a same-turn finite verb form, (4) whether the occurrence constitutes an instance of recapitulation, and (5) the position of the conjunctive participle relative to the subject. These variables were chosen in the interest of identifying the most common functions of the participle in our corpus, which are discussed in Section 4.

3.1. FORM. There are three base forms from which inflected Sinhala verbs are 'built' (Gair 1976, Gair 2003, Gair and Paolillo 1997), which include two tensed bases—nonpast and past—and the participial base. Examples of each base form of the verb *balannə* 'look' are provided in Table 1 below.

Base	Form
non-past	*balə-*
past	*bælu-*
participial	*balə-*

TABLE 1. Base Forms of *balannə* 'to look'

As the three forms above illustrate, the non-past and participial base forms are in some cases identical. To form the conjunctive participle, the morpheme *-la* is suffixed to the participial base, yielding *baləla*.

3.2. FUNCTIONS. As was illustrated by examples (1)-(3) above, the conjunctive participle expresses the temporal SEQUENCE OF EVENTS, typically in cases of same-subject reference across clauses. This referential quality of utterances containing the participle is only a tendency, though, in contradistinction with many Indo-Aryan languages (Masica 1991). The utterance in (6) illustrates this point: here, the subject of the first clause, *kocciə* 'train', differs from that of the second, *api* 'we'.

    (6)   kocciə æwilla   api jannə   giya
            train   come-PPL   1PL go-INF go-PST
            'The train came and we left.'

The verb form in question is also often employed successively within a turn to express a sequence of more than two events and/or states (as in example (3)). Such utterances invariably end with a tense-marked verb. In this way, the Sinhala conjunctive participle may be described as a CLAUSE CHAINING strategy (Longacre 1985, Crain 1992, Genetti 2005), similar in function to participle or converbal constructions in other languages (e.g. Genetti (2005) for Dolakha Newar; Terrill (2003) for Lavukaleve, and Tikkanen (1995) for Burushaski).

In a similar yet syntactically-distinct function, the conjunctive participle serves to repeat information expressed by an immediately preceding clause. Genetti (2005:49) terms this use of the participle construction in Dolakhae RECAPITULATION, describing it as 'a process common in South Asian narrative, where one begins a syntactic sentence by repeating, often in abbreviated form, the substance of the preceding finite clause or sentence.' As (7a)-(7c) below demonstrate, this function of the conjunctive participle is similar to those discussed above, in that the recapitulated event or state is sequenced with a following event or state.

(7) a.   itiŋ   andərə   ṭikak   hitəla   daval-ṭə   gedərə   giya
             then   A.   a.little   think-PPL   afternoon-DAT   home   go-PST
             'Then Andare pondered for awhile and went home for the afternoon.'

      b.   gedərə   gihilla   eya-ge   putaa-ṭə   anḍegahalla   andərə   kiwwa
             home   go-PPL   3M.SG-GEN   son-DAT   call-PPL   A.   say-PST
             'After going home and calling his son, Andare said...'

      c.   mamə   rajəmaaligaavə-ṭə   yanəwa
             I   palace-DAT   go-PRES
             'I am going to the palace...'

In (7b), the initial phrase *gedərə gihilla* '(after) going home' repeats information expressed by the last clause in (7a), namely, that Andare 'went home for the afternoon', *davalṭə gedərə giya*. Moreover, the repeated event is temporally sequenced with two subsequent events: 'calling his son' and a speech event, as we see in (7b). In this way, the conjunctive participle's recapitulative use performs the same sequencing function we see in examples such as (1)-(3).

The last function of the Sinhala participle construction to be examined here is illustrated by examples (4)-(5) above and (8) below. In these and similar utterances, the form in question cannot be characterized as dependent, as it occurs either as a monoclausal predicate or as the final verb in a complement clause. The conjunctive participle's use as a nondependent form imposes an aspectual construal of the situation describable in terms of perfect aspect, as indicated by the English translation in (8). Here, the state of having gone, expressed by means of the participle construction, relevantly persists until and bears on the arrival of the speaker's interlocutor in Texas.

(8)   oya   Texas   wələṭə   enə-koṭə   mamə   Indiawə-ṭə   gihilla
      2SG   T.   PL.INAN.DAT   come-PRES-REL-when   1SG   India-DAT   go-PPL
      'When you came to Texas I had gone to India.'

With the preceding overview in mind, each of these functions is now considered in turn, beginning with event-sequencing and clause chaining.

4.1. EVENT SEQUENCING. As Gair and Paolillo (1997) point out, the conjunctive participle is the most common way of expressing a sequence of actions or events in Sinhala. The corpus data used for the present study indicate that, in the least, such event sequences favor same-subject reference;[2] that is, when one conjunctive participle co-occurs in a turn with a finite verb, the two predicates share a subject. Consider (9)-(11).

(9)   æyə   saŋgi-tə   ahala   æṭuwa
      3F.SG   music-DAT   hear-PPL   dance-PST
      'She heard music and danced.'

---

[2] Although in the corpus used here no cases of event sequencing by means of the conjunctive participle exhibited a change in subject, such examples are attested in the literature (Gair 2003, Gair and Paolillo 1997, Masica 1991) and my elicitation data.

(10) waṇḍura palleha-ṭə æwilla   ṭoppi   goḍə issuwa
      monkey   down-DAT   come-PPL   hat-PL heap steal-PST
      'The monkey came down and stole the hats.'

(11) itiŋ   andare ṭikak   hitəla   daval-ṭə   gedərə giya
      then   A.   a.little think-PPL afternoon-DAT home   go-PST
      'Then Andare pondered for awhile and went home for the afternoon.'

In each of these utterances, we observe that one noun phrase serves as subject for both the conjunctive participle and finite verb. For instance, in (9), 'she' both hears music and dances. Similarly, in (10), 'the monkey' serves as subject for the two clauses, performing both actions depicted by the utterance. This affinity for depicting two consecutive events involving only one non-patient participant functioning as the grammatical subject of both clauses stands in contrast with the subject reference behavior of other strategies for expressing similar instances of event sequencing. To this end, at least two other forms are also available, namely, the PRIOR TEMPORAL form and the use of an instrumentalized verb immediately followed by the lexeme *passe* 'after'. The existence of these potential alternatives to the use of the conjunctive participle makes necessary an explanation of one's use over another in a particular context and syntactic environment. Although a comprehensive explanation of this sort is beyond the scope of the present analysis, I present a few preliminary observations below.

First, as was noted above, the expression of an event sequence involving same-subject reference across clauses favors the use of the conjunctive participle. By contrast, all of the utterances taken from the corpus which contain a combination of an instrumentalized verb and *passe* (6/6 total), as well as half of the utterances containing the prior temporal form (2/4 total), express a sequence of events involving a change in subject. Consider (12) below.

(12) ṭikkə   welawak giya-in passe ṭoppi welenda   nægiṭṭa
      a.little time-IND go-INST after hat-PL merchant awaken-PST
      'After a little time went by, the hat seller woke up.'

The content of the first clause in (12) proves indicative of this form's use in the corpus and elicitation data. Here, the phrase *ṭikkə welawak giyain passe* 'after a little time went by' renders the temporal relation expressed by the instrumentalized verb-*passe* combination more transparent. In such cases, the use of this construction as a clause linking strategy entails both nonsimultaneity of the events (or states) and an intervening, nonpunctual temporal interval. For example, in (12) a short, nonpunctual duration of time passes before the hat seller awakes. This quality of events depicted by the verbal construction in (12) is suggested not only by the translation, 'After X, Y...' but also by elicited minimal pairs varying only in the use of either the conjunctive participle or the combination of an instrumentalized verb and *passe*, illustrated in (13) and (14) below.

(13) itin ookə dækka-in passe mage   yaaluwek kiwwa
      so   that   see-INST   after 1SG-GEN friend-IND say-PST
      'So after seeing that, a friend of mine said...'

(14) itin ookə dækəla passe mage yaaluwek kiwwa
so that see-PPL after 1SG-GEN friend-IND say-PST
'So seeing (having seen) that, a friend of mine said...'

According to the language consultant, the utterance in (13) depicts a situation in which the 'seeing' event concludes a short time before the speaker's friend begins to talk; in other words, the first and second event do not overlap. By contrast, the utterance in (14) can be interpreted as involving temporal overlap, the first event preceding the second inceptively, or, alternatively, the two events may be interpreted as noncoextensive. Thus, with respect only to temporal sequencing, the instrumentalized verb strategy appears to specify a more fine-grained circumstantial relation between the linked clauses, whereas we observe a coarser depiction of the event-sequence temporally in the case of the conjunctive participle.

In this respect, then, the two forms differ in regard to the level of circumstantial specificity afforded by each's use. Moreover, as the corpus data show, the two forms exhibit a degree of complementary specialization with respect to subject reference; the conjunctive participle being used in same-subject sequences, and the instrumentalized form elsewhere. In counterpoint to this complementary distribution, the prior temporal form—which occurs four times in the corpus—sequences events involving two non-patient participants as well as those involving one. Consider (15a)-(15c) and (16).

(15) a. balla daŋgələla-daŋgələla janee-len eliyə-ṭə pænna
dog fidget-REDUP window-from out-DAT jump-PST
'The dog kept fidgeting about, and jumped out of the window.'
b. eliyə-ṭə pænəla
out-DAT jump-PPL
'(He) jumped out.'
c. wæṭunaamə botəle biňdila laməya balla-wə beerəgattə
fall-PRTMP bottle break-PPL child dog-ACC rescue-PST
'As (he) fell, the bottle broke and the child rescued the dog.'
(16) andəree-ṭə meekə æhunamə andəree kiwwa rajjuruwan-ṭə
A.-DAT this hear-PRTMP A. say-PST king-DAT
'When Andare heard this, he said to the king...'

As the English translations suggest, the events in (15c) and (16) either overlap (as in case of the former) or nearly overlap (as in the latter). For instance, in (15c), the termination of the falling event and the bottle's breaking coincide. In (16), a period of near punctual duration separates the two events depicted by the utterance. Thus, the prior temporal form contrasts with the instrumentalized verb-*passe* combination, in that they express different temporal relations. The two forms are similar, however, as they both specify a circumstantial relation, in contradistinction with the conjunctive participle, whose use expresses only the gross temporal relation of sequence. In this way, the participle form contrasts with both the instrumentalized verb and prior temporal form, which both express a more specific temporal relation.

The latter verb forms also differ distributionally from the conjunctive participle in that they do not form chains. At least two factors bear on this disparity, the first of which being the

explicit marking of interpropositional relations. As Genetti (2005:43) demonstrates in her discussion of participial and adverbial clauses in Dolokhae, such marking limits the latter's 'freedom of occurrence, and makes them suitable for their discourse function of expressing rhetorical relations.' Genetti continues by arguing that because of this discourse function, adverbial clauses 'do not easily combine into long chains.' Such an analysis accords well with the data and observations of the present study.

For example, the first alternative to the conjunctive participle considered above—namely, the instrumentalized verb form—co-occurs with the independent lexeme *passe* 'after', signaling that the event has come to an end and that another event follows. Similarly, as Gair (2003:811) points out, the prior temporal suffix "is historically derived from a lexical form *hamə*", which means 'when'. This degree of temporal specificity, which is absent in the conjunctive participle, prohibitively reduces the ease with which these adverbial forms could combine into chains.

The second factor potentially contributing to this inability to form chains bears directly on the remainder of the analysis. In addition to the differences between the conjunctive participle and the two temporal alternatives discussed above, the data indicate that the former and latter contrast conceptually in the construal imposed by their use on the relation between the two sequenced events. To illustrate this dissimilarity, first consider (17a)-(17c) below.

(17) a. wandura kehelgediya kææwa-in passe mæruna
     monkey banana eat-INST after die-PST
     'After the monkey ate a banana he died.'

   b. wandura kehelgediya kææwaamə mæruna
     monkey banana eat-PRTMP die-PST
     'When the monkey ate a banana he died.'

   c. wandura kehelgediya kaala mæruna
     monkey banana eat-PPL die-PST
     'The monkey ate a banana and died.'

In (17a), the combination of instrumentalized verb and *passe* profiles the nonoverlapping temporal relation between the two events, namely, the monkey eating a banana and the event of its death. I use the term PROFILE here in the sense of Langacker (1987, 1991), in which a form-meaning pair, such as the verbal construction in (17a), brings into focus 'a particular substructure' of the conceptual content evoked by the construction's use (Langacker 1987:183). This substructure, which may be either a thing or relationship, constitutes one element of the form-meaning pair's 'scope of predication' or 'base.' Together, the base and profiled element evoked by a construction form a relationship which imposes a particular construal on a situation, such as the consecution of two events, as in (17a).

In this example, the base involves two events sequenced temporally, one preceding the other, with a nonpunctual duration of time interposed. It is this nonoverlapping temporal relation that is profiled by the instrumentalized verb-*passe* combination. Similarly, the utterance in (17b) profiles a specific temporal relation, namely one of near-simultaneity, with the first event minimally-preceding the second. Moreover, in (17c), the use of the conjunctive participle also profiles a temporal relation between the events expressed by each clause, but in this case, the relation is less fleshed-out, indicating only consecution. In this way, the three forms appear to be reasonably similar in function, differing only minimally with respect to the

nature and degree of the temporal relation specified by each. Furthermore, a correlational interpretation of the event sequence in each utterance above is possible; that is, one may infer a relation between the two events beyond that of temporality. The possibility of such an interpretation of each utterance raises a descriptive question, namely, whether such a construal of the events results from the conventional profile imposed by each verb form or is arrived at primarily through an interaction of world knowledge and discourse context. One source of possible explanation comes from elicited utterances, such as (18a)-(18b).

(18) a. kurula sindukiwwa-in passe mage   amma   aawa
       bird   sing-PST-INST   after 1SG-GEN mother come-PST
       'After the bird sang, my mother arrived.'
    b. kurula sindukiwwaamə mage   amma   aawa
       bird   sing-PRTMP   1SG-GEN mother come-PST
       'When the bird sang, my mother arrived.'

Each of these examples depicts a situation involving two events occurring in succession. However, given our knowledge of bird songs and the coming and going of people, the two events are not interpreted as standing in a correlational relation, only one of temporal sequence. Thus, the two verb forms in (18a) and (18b) do not appear to encode a correlation between events beyond that of temporality. With this in mind, we would expect that if the conjunctive participle encoded only temporal consecution, it could felicitously substitute for either verb form in the examples above. This, however, is not the case, as the language consultant rejected the participle's replacement of either the instrumentalized verb or prior temporal form in this and similar utterances, as is illustrated in (18c) below.

    c. kurula sindukiyəla passe mage   amma   aawa
       bird   sing-PPL   after 1SG-GEN mother come-PST
       *'The bird sang and my mother arrived.'

This disparity in usage provides evidence that the Sinhala conjunctive participle does conventionally profile a correlational relation between the two events in addition to a relation of temporal sequence.

One potential counterargument to such a proposal relies on distributional evidence, namely, the fact that the conjunctive participle overwhelmingly favors event sequences involving same-subject reference. However, as was discussed in Section 3.2, the participle construction can depict event sequences with distinct subjects, as illustrated by example (6), repeated here as (19).

(19) kocciə æwilla   api jannə giya
     train   come-PPL 1PL go-INF go-PST
     'The train came and we left.'

The significance of such utterances lies in the nature of the relation between the two events. In (19), they are not only sequenced, but also exhibit a correlation beyond that of succession. This relationship can be expressed by the English translation Masica (1991:400)

refers to as 'the most literal' rendering of the conjunctive participle, namely, 'Having done Y, X...' Thus, applying this translation, (19) would read 'The train having come, we left.' The use of the PERFECT in this translation captures the continued relevance of the train's arrival to the event of departure expressed by the second clause. Such a sequentially-interrelated relevance of events is not evidenced by utterances such as (18a) and (18b) above. Instead, in these and similar utterances, the events are construed as standing only in a temporal relation.

4.2. CLAUSE CHAINING. In a related function, the conjunctive participle can occur several times in one utterance expressing a sequence of events. This capacity to form CLAUSE CHAINS (Crane 1992, Genetti 2005, Longacre 1985, Myhill and Hibiya 1988, Terrill 2003) is illustrated by (20) and (21) below.

(20) æyə   nægitəla   koopi   hadəla   pattəre   kiyəwəla   giya
      3F.SG   awaken-PPL   coffee   boil-PPL   paper   read-PPL   go-PST
      'She woke up, made coffee, read the paper and left.'

(21) miniha   gallak   ussəla   wandura-ṭə   gahalla   duwəla   heŋguna
      man   rock-IND   lift-PPL   monkey-DAT   throw-PPL   run-PPL   hide-PST
      'The man picked up a rock, threw it at the monkey, ran away, and hid.'

In both of these utterances, we observe a series of events, temporally-sequenced, involving one subject shared by each clause. Moreover, the sequence of events in each example exhibits a type of correlational coherence absent in utterances such as (18a)-(18b) above. For instance, the events in (20) taken together constitute a larger 'macro-event,' namely, what may be termed a prework morning ritual. Each clause thus describes one subevent, the completion of which brings the utterance's subject one step closer to the culmination of the event chain: departure for work. In this way, the completion of each act—waking up, making coffee, and reading the paper—bears relevantly on the subsequent event in the chain.

Similarly, in (21), the use of the conjunctive participle to express the sequence of actions carried out reflects a 'correlational curve' with an inception (picking up a rock) and completion (hiding). As in the preceding example, what may be described here as a monkey attack comprises several subevents, culminating in the event depicted by the tense-marked verb *heŋguna* 'hide'.

Thus, the two preceding functions of the conjunctive participle—(simple) event sequencing and clause chaining—correspond conceptually in their construal of event sequences. Specifically, as demonstrated by the discussion of examples illustrative of both functions, the state resulting from an anterior action, such as making the coffee or picking up a rock, persists relevantly until the inception of a subsequent event. In this way, each use of the conjunctive participle profiles both a correlational relation between events and a coarse temporal relation.

4.3. RECAPITULATION. In addition to the preceding functions, the conjunctive participle is also used in cases of recapitulation, as described in Section 3.2. In this capacity, the form in question not only performs the discourse function of repetition, but also serves to sequence two events; one expressed by the repeated information and another predicated by a following clause. Furthermore, the observed correlational relation between successive events is also in evidence, as illustrated by (22a)-(22b).

(22) a. baisikəle galəkə   hæpuna
          bicycle   stone-IND   hit-PST
          'The bicycle hit a stone.'
      b. galəkə   hæpila laməya-i persgeḍi okomə bimə wæṭuna
          stone-IND hit-PPL child-CONJ pears   all   ground fall-PST
          'After hitting the stone, the boy and the pears all fell to the ground.'

In this case, as a result of the first event, the boy riding the bicycle loses his balance and consequently falls to the ground. Thus, the state of imbalance relevantly bears on the boy's fall. Again, the correlation between these two events can be captured in English by translating the second utterance as 'Having hit the stone, the boy and his pears all fell to the ground'.

4.4. NON-DEPENDENT PREDICATION. As was noted in Section 3.2, the Sinhala conjunctive participle is unique among Indo-Aryan languages in its capacity to function as a non-dependent predicate, either monoclausally or as the final verb in the clause. Moreover, in such cases, the participle expresses perfect aspect, as illustrated by the following example.

(23) oya   heṭə   enə-koṭə   mamə California wələṭə   gihilla
      2SG   tomorrow come-PRES-REL-when   1SG C.   PL.INAN.DAT go-PPL
      'When you come tomorrow, I will have left for California.'

Following Comrie (1976:52), I take perfect aspect to indicate 'the continuing... relevance of a past situation'. We observe this sense precisely in (23) above, in which the continued relevance of 'having gone' persists until the interlocutor's expected arrival. Similarly, in (24a)-(24b), we see that the resulting state of the theft relevantly bears on the man's observation that his food has been taken, depicted by the participial form of the verb *kərəla* 'do' in the expression *horəkam kərəla* .

(24) a. ohuge   baharyawə hoyənə-gaman
          3M.SG-GEN   wife   look-PRES-REL-when
          'While looking for his wife...'
      b. horek   tamange   kææmə horəkam kərəla kiyəla ohu   dææka
          robber self-GEN food   theft   do-PPL COMP   3M.SG   see-PST
          'he saw that a robber had stolen his food.'

As examples (23) and (24a)-(24b) show, the Sinhala conjunctive participle can function as a nondependent predicate, occurring as either the main clause verb or embedded in a complement clause. In such cases, we observe a meaning consistent with the interpretation traditionally attributed to the expression of perfect aspect, as the English glosses suggest.

5. COMMON CONSTRUAL. To summarize the findings common to each function considered above (event sequencing, clause chaining, recapitulation, and nondependent predication), we observed first, in cases of two-event sequences, that the conjunctive participle profiles a

correlational relation between the events in addition to a rough temporal relation, as we see in example (25) below.

(25) maŋ gihilla ee kaaryaləyin æhua mage bææg ekə kohedə kiyəla
    I go-PPL that office ask-PST I-GEN bag one where COMP
    'I went to that office and asked, "Where is my bag?"'

Here, the conjunctive participle not only sequences the events of going and asking, but also profiles the correlational coherence between the two events. This correlation, which was shown to be in evidence for the related functions of clause chaining and recapitulation, involves the continued relevance of a resultant state bearing on the event expressed by the following clause. In example (25) above, the resultant state of the speaker going to 'that office' relevantly bears on the inquiry made once there. This relationship among events has been observed in Dravidian and Indo-Aryan languages by Lindholm (1975) and Masica (1991:400), respectively, the latter noting that, in regard to the conjunctive participle's use as a clause linkage strategy, 'not just any two clauses may be so linked: they must have what [Lindholm] calls "natural relevance"—an elusive concept when one tries to define it, but independently cited by other investigators.'

With respect to the construal imposed by the conjunctive participle's use a nondependent predicate, we observe a similar relationship in the expression of perfect aspect. In such cases, the participle profiles the continued relevance of a resultant state to the speech act, as well as the moment of a past or future event, as illustrated in (26) below.

(26) gəhæni kukula-wə marəla dæn hæmotəmə kaanə puluwan
    woman chicken-ACC kill-PPL now everyone eat-INF can
    'The woman has killed the chicken and now everyone can eat.'

In this example, the resultant state of the first event, namely that of killing a chicken, relevantly bears on the speaker's immediate situation at the time of the utterance. Specifically, the state expressed by the second clause follows as a consequence of killing the chicken. Thus, the construal evoked by the participle's use in utterances such as (26) parallels the construal imposed by its use as a clause linkage strategy, exemplified in (25). In both cases, a correlational coherence obtains between two situations which involves the resultant state of a prior event relevantly persisting until and directly bearing on a succeeding event. In this way, the scene construal characteristic of perfect aspect conceptually unites the syntactically-disparate functions.

6. Conclusion. The significance of the findings presented here are twofold. First, I have presented evidence in favor of a polysemy analysis of the Sinhala conjunctive participle. Specifically, I have argued that a 'common construal' is in evidence for each of the participle's distinct syntactic functions. This construal, which involves a correlational relationship between a prior event and a subsequent situation, crosscuts each of the conjunctive participle's functions discussed above.

Second, I have shown that the form in question performs a number of the functions typical of converbal constructions, despite its capability to serve as a nondependent predicate. These observations contribute to the ongoing typological dialogue interested in establishing a

crosslinguistic prototype of such forms. Moreover, the data discussed above underscore the disadvantages of emphasizing definitional criteria, such as nonfiniteness and nondependency, at the expense of a prototype model. By narrowing their scope to a neatly delimitable set of forms, such approaches potentially exclude candidates for analysis which would deepen our understanding of how converbal functions are formally-instantiated crosslinguistically.

REFERENCES

BICKEL, BALTHASAR. 1998. Review of Converbs in Cross-Linguistic Perspective, by Martin Haspelmath and Ekkehard König, eds. *Linguistic Typology* 2.381-397.

CRANE, LAURA DIANE. 1992. Clause Chaining in Nepali Discourse. Los Angeles, CA: UCLA dissertation.

GAIR, JAMES W. 1970. *Colloquial Sinhalese Clause Structures.* The Hague: Mouton.

GAIR, JAMES W. 1976. The verb in Sinhala, with some preliminary remarks on Dravidianization. *International Journal of Dravidian Linguistics* 5.2.259-273.

GAIR, JAMES W. 2003. Sinhala. *The Indo-Aryan Languages*, ed. by George Cardona and Dhanesh Jain, 766-817. London: Routledge.

GAIR, JAMES W. and John Paolillo. 1997. *Sinhala.* München: Lincom Europa.

GENETTI, CAROL. 2005. The Participle Construction of Dolokha Newar: Syntactic
  Implications of an Asian Converb. *Studies in Language* 29.1.35-87.

HASPELMATH, MARTIN. 1995. The Converb as a Cross-Linguistically Valid Category. *Converbs in Cross-Linguistic Perspective*, ed. by Martin Haspelmath and Ekkehard König, 1-55. Berlin:
  Mouton de Gruyter.

HASPELMATH, MARTIN. and EKKEHARD KÖNIG. 1995. *Converbs in Cross-Linguistic Perspective.* Berlin: Mouton de Gruyter.

LANGACKER, RONALD. 1987. *Foundations of Cognitive Grammar, Vol. 1.* Stanford: Stanford University Press.

LANGACKER, RONALD. 1991. *Foundations of Cognitive Grammar, Vol. 2.* Stanford: Stanford University Press.

LINDHOLM, JAMES. 1975. *The Conceptual Basis of the Tamil Adverbial Participle.* Chicago, IL: University of Chicago dissertation.

LONGACRE, ROBERT. 1985. Sentences as combinations of clauses. *Language Typology and Syntactic Description 2: Complex Constructions*, ed. by Timothy Shopen, 235-286. Cambridge: Cambridge University Press.

MASICA, COLIN. 1976. *Defining a Linguistic Area.* Chicago: University of Chicago Press.

MASICA, COLIN. 1991. *The Indo-Aryan Languages.* Cambridge: Cambridge University Press.

MYHILL, JOHN and JUNKO HIBIYA. 1988. The Discourse Function of Clause Chaining. *Clause Combining in Grammar and Discourse*, ed. by John Haiman and Sandra A. Thompson, 361-398. Amsterdam: John Benjamins.

NEDJALKOV, IGOR´. 1995. Converbs in Evenki. *Converbs in Cross-Linguistic Perspective*, ed. by Martin Haspelmath and Ekkehard König, 441-463. Berlin: Mouton de Gruyter.

TERRILL, ANGELA. 2003. *A Grammar of Lavukaleve.* Berlin: Mouton de Gruyter.

TIKKANEN, BERTIL. 1995. Burushaski Converbs in their South and Central Asian Context.
  Converbs in Cross-Linguistic Perspective, ed. by Martin Haspelmath and Ekkehard König, 487-528. Berlin: Mouton de Gruyter.

Rice University
Department of Linguistics, MS 23
6100 Main Street
Houston, TX 77005-1892
sastre@rice.edu

RELATIVE CLAUSES IN SINHALA

Benjamin R. Walker

Rice University

1. Introduction. Relative clauses are clauses which modify a noun. These clauses add information about the modified noun, called the head noun, which cannot be conveyed with a single adjective. Instead, relative clauses use an entire clause to capture the quality to be imparted to the head noun. Relative clauses contain a relativized noun phrase, NPrel, which is coreferential with the head noun. Also relative clauses are marked by some sort of relativizer, whether a verb suffix, particle, or pronoun (Payne 1997:326). Sinhala constructs its relative clauses using the 'gap' strategy and non-finite verb forms in clauses placed before the head noun.

2. Data. During the academic year of 2004-2005, Nissanka Sirimevan Wickremasinghe, a native speaker of Sinhala from Sri Lanka, provided elicited examples and seven texts in Sinhala. From this database 100 relative clauses were recovered, 9 from the texts and the rest from elicitation sessions.

3. Construction of Relative Clauses in Sinhala. In the collected data, Sinhala demonstrated a predominantly SOV word order. In accordance with Greenberg's word-order correlates, we find relative clauses preceding the head noun they modify.

(1) laməya [ohu ændəpu]   redi   heduwa
     child   3SG   wear-PST-REL   clothes   wash-PST
     'The child washed the clothes that he wore.'[1]

In example (1) above, the basic structure of relative clauses in Sinhala can clearly be seen. The basic clause *laməya redi heduwa*, 'the boy washed the clothes', demonstrates the predominate word order of declarative clauses in Sinhala. The verbal element *heduwa,* 'wash,' comes at the end of the clause. The subject, *laməya*, 'child,' and then the object, *redi*, 'clothes,' precede the verb. In accordance with the word order correlates, the relative clause *ohu ændəpu*, 'which he washed,' precedes the noun that it modifies, *redi.* It should also be noted that the declarative word order is maintained within the relative clause. In (1), *ohu*, 'he,' the subject of the relative clause precedes the verb element, *ændəpu*, 'wash,' and the object, Nprel, coreferential with the clothes in this case, is omitted.

Sinhala relative clauses are formed with a 'gap,' covered later in the paper, and a non-finite verb form. The verb forms used in relative clauses are labeled 'nonfinite,' because they do not have the same inflection as main verbs in independent, declarative clauses and cannot stand alone as the main verb of such a clause. The verbs found in relative clauses have one non-past form and two past forms.

---

[1] Relative clauses will be bracketed for easier recognition throughout this paper.

PRES	NONPAST-REL	PAST-REL 1	PAST-REL 2
kərənəwa 'do'	*kərənə*	*kərəpu*	*keruwə*
hodənəwa 'wash'	*hodənə*	*hodəpu*	*heduwə*
dakinəwa 'see'	*dakkina*	*dækkəpu*	*dækkə*

TABLE 1. Verb forms for Independent and Relative Clauses

(2) NONPAST Form

mamə [mage   wæḍə   kərənə]   laməyaṭə   pain gæhuwa
1SG   1SG-GEN work   do-NPST-REL child-DAT kick-PST
'I kicked the boy who does my work.'

Example (2) demonstrates the nonpast, non-finite verb form used in relative clauses. The verb *kərənə,* 'do,' is the nonpast relative verb form of *kərənəwa.* The relative clause precedes the head noun, *laməyaṭə,* 'child,' which is in the dative case as required by this particular main verb, *pain gæhuwa,* 'kick.'

(3) PAST 1 Form

laməya [ohu   epa   kərəpu]   redi   heduwa
child   3M.SG hate   do-PST-REL clothes wash-PST
'The boy washed the clothes which he hated.'

Example (3) uses the more common of the past relative forms, those ending in the *-pu* suffix. Again, the relative clause precedes the head noun, *redi,* 'clothes,' which is the object of the main verb, *heduwa,* 'wash.'

(4) PAST 2 Form

miniha [nammə   amətəkə keruwə]   laməyaṭə   kata kərənəwa
man   name   forget   do-PST-REL child-DAT speak-PRES
'The man speaks to the boy whose name he forgot.'

Example (4) uses the less common form of the past relative verb, built upon the dative object required by the main verb of the sentence, *kata kərənəwa,* 'speak.'

Two elicited examples suggested a variation between the Past 1 and Past 2 forms listed above based on the grammatical relation of NPrel. In example (5) below, the Past 2 form corresponds with NPrel acting as a subject of the relative clause, while in example (6), the Past 1 form is used with NPrel acting as an object of the relative clause.

(5) NPrel as Subject

miniha [[tamanwə hæpuwə]   balla aiti]   laməyaṭə   kata kərənəwa
man   self-ACC bite-PST-REL dog   own-PRES-REL   child-DAT speak-PRES
'The man speaks to the child whose dog bit him.'

Example (5) contains both a relative clause and an additional relative clause embedded within the first, as indicated by the brackets. This construction is common for expressions of

explicit ownership and will be discussed in further detail later in this paper. It is important to note for the current discussion only the grammatical relations of the NPrel in comparison with the form of the past relative used. In this example, *hæpuwə*, 'bite,' a Past 2 form, coincides with NPrel as its subject.

(6) NPrel as Object

miniha	[[taman	hapəpu]	ballage	aitikaarə]	laməyatə	kata kərənəwa
man	self	bite-PST-REL	dog-GEN	owner	child-DAT	speak-PRES

'The man speaks to the child whose dog he bit.'

Example (6) also contains an embedded relative clause. Again, it is important only to note that the Past 1 relative form *hapəpu*, 'bite,' co-occurs with NPrel as its object.

However, this distinction does not hold true in all cases. In example (7) below, a Past 1 form occurs with NPrel as its subject, not its object as in example (6).

(7) NPrel as Subject

[pussawə	hapəpu]	ballatə	Mamə	kæməti
cat-ACC-ANIM	bite-PST-REL	dog-DAT	1SG	like

'I like the dog that bit the cat.'

In this example, despite the use of a Past 1 form, NPrel is its subject. NPrel is coreferential with *ballatə*, 'dog,' which does the biting in the relative clause. Further data will need to be collected in order to make a more informed attempt at explaining the variation between these two past verb forms.

A few verbs show an unusual past relative form. For instance, the past relative forms for 'fall', 'become', and 'die' are respectively, *wæticcə*, *mæriccə*, and *weccə*. It is uncertain with which past relative form these forms correspond as additional past relative forms for these verbs have not been elicited.

4. THE GAP STRATEGY. Sinhala expresses NPrel, the element in the relative clause that is coreferential with the head noun, by leaving it out altogether, or 'gapping' it. The omitted word along with the verb form marks the clause as a relative clause, not an independent one. The grammatical relation of the omitted or 'gapped' word, NPrel, can then either be retrieved through context or through suffixes on the expressed argument.

(8) NPrel as Subject

mamə	[-----	mage	wædə	kərənə]	laməyatə	pain gæhuwa
1SG	GAP	1SG-GEN	work	do-NPST-REL	child-DAT	kick-PST

'I kicked the boy who does my work.'

Example (8) illustrates a typical relative clause where NPrel is the subject of the clause. Putting aside the main clause, *mamə laməyatə pain gæhuwa*, 'I kicked the child,' leaves the incomplete fragment, *mage wædə kərənə*, 'who does my work.' The verb form and the missing argument mark this as a dependent, relative clause, built on the dative object, *laməyatə*, 'child.' Because of the verb-final word order, it is ambiguous at first whether the expressed argument is the subject or object of the transitive, relative verb, *kərənə*, 'do.' However, this argument is

not typically agentive enough to be the subject of this verb. Therefore, this clause lacks a subject, and NPrel, being coreferential with the highly agentive noun, *lamǝya*, 'child', fits logically into this 'gap.'

    (9) NPrel as Object

       [ohu   ------   kǝrǝpu]   wæḍǝ   apahasui

       3M.SG   GAP   do-PST-REL   work   difficult

       'The work he did was hard.'

In example (9) NPrel is the object of the relative clause. Once again, removing the main clause, *wæḍǝ apahasui*, 'the work is hard,' the fragment that is left is incomplete. In Sinhala animate objects are marked with a suffix, -*wǝ*. The lack of this suffix on the pronoun in the relative clause, *ohu*, 'he,' indicates that it is in the nominative case and therefore the subject of *kǝrǝpu*, 'do.' This verb, which is typically transitive, thus lacks an object. NPrel is coreferential with *wæḍǝ*, 'work,' in this sentence, a prototypical object, especially for this particular verb. Therefore NPrel is the object of the relative clause.

However, case recovery is not always so clear, especially when there are no case markings present, as in the case of inanimates acting on one another, or when a sentence has two possible agents and one object.

    (10) Ambiguous Relative Clause

       ?[kolla   [tæægǝ   dunnǝ]   kellǝṭǝ   pain gæhuwa

       boy   gift   give-PST-REL   girl-DAT   kick-PST

       'The boy kicked the girl to whom he gave the gift.' or 'The boy kicked the girl who gave him the gift.'

Example (10)'s ditransitive structure makes the case of NPrel ambiguous. As indicated above, NPrel could be construed as either the subject or the indirect object of the relative clause. The interpretation depends on whether *kolla*, 'boy,' is read as the subject of the main clause or as the subject of the relative clause, with the subject of the main clause then being implied. If *kolla* is the subject of the main clause, then the relative clause lacks an agentive subject to fit its verb. NPrel fills preferentially the subject role, more often leaving the direct object or in this case, the indirect object, to be supplied by context. Therefore, a Sinhala speaker would assume the subject of the relative clause, NPrel, to be the same as the head noun, *kellǝṭǝ*, 'girl,' a typical agent. *Kolla*, would then be the implied indirect object of the relative clause as the only remaining noun that would logically fit this role. Under this assumption, the girl would be the subject of the relative clause, the one giving the gift. However, if *kolla* is the subject of the relative clause, then only the indirect object of *dunnǝ*, 'give,' is missing from the clause, and therefore NPrel, coreferential with *kellǝṭǝ*, must be the missing indirect object. In this interpretation, the boy would be the subject of the relative clause, the one giving the gift.

    (11) Unambiguous

       [tæægǝ   dunnǝ]   kellǝṭǝ   kolla   pain gæhuwa

       gift   give-PST-REL   girl-DAT   boy   kick-PST

       'The boy kicked the girl who gave him the gift.'

Example (11) changes the word order of example (10) from SOV to OSV, eliminating the possibility of interpreting *kolla* as the subject of the relative clause. Therefore, the relative clause lacks a subject, and NPrel must assume this role. Again, *kolla* becomes the implied indirect object of *dunnə*.

(12) Unambiguous
    kolla [taman tæægə dunnə]   kellətə   pain gæhuwa
    boy   self   gift   give-PST-REL girl-DAT kick-PST
    'The boy kicked the girl to whom he gave the gift.'

Example (12) adds the pronoun *taman,* 'self,' to the relative clause. *Taman* is a reflexive pronoun and can only refer to an explicitly stated antecedent. This pronoun clearly refers to the subject of the main clause, because NPrel is always gapped. Therefore, with *kolla,* 'boy,' accounted for as the subject of the relative clause, and *tæægə,* 'gift,' as the direct object, only *kellətə,* 'girl,' is left as a logical indirect object.

In cases with two animates acting on one another, the accusative case suffix, *-wə*, clears away ambiguity. Animate direct objects in Sinhala are marked with this suffix, clearly distinguishing them from subjects. Therefore, with two animate objects acting on one another, this suffix clears away any ambiguity by its presence or absence on the overt argument.

(13) NPrel as Subject
    [pussawə hapəpu]   ballətə   mamə   kæməti
    cat-ACC   bite-PST-REL dog-DAT   1SG   like
    'I like the dog that bit the cat.'

In example (13) the accusative object of the relative clause, *pussa,* 'cat,' is marked with the suffix *-wə.* With the direct object accounted for, NPrel must be the subject of this relative clause in order to complete it.

(14) NPrel as Object
    [pussa hapəpu]   ballətə   mamə   kæməti
    cat   bite-PST-REL dog-DAT   1SG   like
    'I like the dog that was bitten by the cat.'

In example (14) *pussa,* 'cat,' is not marked with the accusative suffix. However, since *pussa* is animate and lacks the accusative suffix, or any other suffix marking it as one of the other cases, it must be in the unmarked, nominative case. Therefore NPrel must be the accusative object of the relative clause.

Clauses involving two inanimate objects acting on one another also create ambiguity that cannot be resolved with the use of the animate accusative suffix. However, there is a inanimate agentive suffix, *-yen*, that may sometimes be used to make the meaning clear.

(15) NPrel as Subject

    mamə [meeseyə samətəla kərəpu]   peṭṭiyə issuwa
    1SG   table   flat   do-PST-REL box   lift-PST
    'I picked up the box that flattened the table.'

In example (15) there is only one explicitly stated argument, *meeseyə*, 'table,' in the relative clause. As mentioned previously, NPrel appears to fill the subject role preferentially. Therefore, with NPrel as the subject, *meeseyə* must be the direct object of the transitive verb, *samətəla kərəpu*, 'flatten.'

(16) NPrel as Object

    [meeseyen   samətəla kərəpu]   peṭṭiyə mamə issuwa
    table-AGENT flat   do-PST-REL box   1SG   lift-PST
    'I picked up the box that the table flattened.'

In example (16) the overt argument *meeseyə,* 'table,' carries the suffix *-yen*. This suffix marks the argument as the subject of this relative clause, leaving NPrel to be the object of the clause.

This agentive suffix appears to be similar to the instrumental suffix, as in *polisiyen*, 'police,' in the instrumental case. It also resembles the locative suffix meaning 'from', as in *ambə gediyen,* 'from the mango.' However, it should be noted that this construction was difficult for the consultant to use. For instance, he was unable to produce the same paradigm around the objects *rupəwahiniyə,* 'TV', and *potə,* 'book'.

5. KEENAN'S AND COMRIE'S RELATIVIZATION HIERARCHY. Keenan and Comrie constructed a hierarchy of argument types on which languages form relative clauses. They found an order of elements that if a language can form a relative clause on one argument type, then it can form relative clauses on all of the types to the left on the hierarchy.

SUBJ>DIRECT OBJ>INDIRECT OBJ>OBLQ>POSSESSOR

(Keenan and Comrie 1979:333-351).

Sinhala can form relative clauses on all of the elements with some trouble with the last argument type, possessors.

(17) Relativized Subject

    arə [mawə dækkəpu]   miniha
    DEM 1SG-ACC see-PST-REL man
    'That is the man who saw me.'

(18) Relativized Direct Object
 lamǝya [ohu ændǝpu] redi heduwa
 child 3M.SG wear-PST-REL clothes wash-PST
 'The boy washed the clothes that he wore.'

(19) Relativized Indirect Object
 miniha [taman tæægi dennǝ] lamǝyatǝ kata kǝrǝnawa
 man self gift-PL give-NPST-REL child-DAT speak-PRES
 'The man speaks to the boy to whom he gives gifts.'

(20) Relativized Oblique
 mee [laŋgǝ tibunǝ] hootǝle ekǝ mage muǝl dawǝsǝ gattǝkeruwa
 HES close.by exist-PST-REL hotel one 1SG-GEN first day spend-PST
 'Um, I spent my first day at a hotel that was close by.'

In addition to canonical subjects, Sinhala can also form relative clauses using dative subjects. However, the case of NPrel does not affect the head noun in any way, nor is it expressed explicitly.

(21) Dative Subject
 lamǝyatǝ gedǝrǝ wædǝ matak una
 child-DAT home work remember-PST
 'The boy remembered the home work.'

(22) NPrel as Dative Subject
 [gedǝrǝ wædǝ matak wecca] lamǝyawǝ mamǝ dannǝwa
 home work remember-PST-REL child-ACC 1SG know
 'I know the boy who remembered the homework.'

In example (22), *lamǝya,* 'child,' takes the accusative suffix *-wǝ* required by the main clause, leaving no trace of the dative case of NPrel.
 Sinhala only creates relative clause on possessors when the possession is either inherent or explicit ownership.

(23) Head Noun as Inherent Possessor
 kolla [bottam kædunǝ] kamisǝ heduwa
 boy button-PL break-PST-REL shirt wash-PST
 'The boy washed the shirt whose buttons were broken.'

In example (23) NPrel refers to the shirt, *kamisǝ,* but its ownership of the buttons must be inferred based on the relationship of the part to the whole.

(24) Head Noun as Inherent Possessor
 kolla [balla Mæriccǝ kellǝtǝ kata keruwa
 boy dog die-PST-REL girl-DAT speak-PST
 'The boy spoke to the girl whose dog was dead.'

In example (24) the relationship between NPrel, *kelləṭə*, 'girl,' and the subject of the relative clause, *balla*, 'dog,' is more subtle. However, since the verb in the relative clause is intransitive, and the ownership of dogs as pets by children is so salient, NPrel as a possessor is the most reasonable interpretation.

(25) Head Noun as Inherent Possessor
    [nyanəwəntə putta innə]   miniha   welendek
    intelligent   son   exist-PRES-REL man   merchant
    'The man whose son is intelligent is a merchant.'

In example (25) the highly salient relationship of kinship between father and son is implied.

Other relative clauses built on possessors can be formed when the relationship between owner and property is overtly expressed.

(26) Declarative Clause Expressing Ownership
    balla   kollaṭə   aiti
    dog   boy-DAT   own
    'The boy owns the dog.'

(27)   miniha [balla aiti]   laməyaṭə kata keruwa
     man   dog   own-PRES-REL child-DAT speak-PST
     'The man spoke to the boy who owns the dog.'

Example (27) is a typical relative clause where NPrel is the subject and the ownership is explicitly stated in the verb *aiti*, 'own.'

(28)   [[hayiyen duhənə]   balla aiti]   miniha   welendek
     fast   run-PRES-REL dog   own-PRES-REL man   merchant
     'The man who owns the dog that runs fast is a merchant.'

In example (28) there are two relative clauses, one within the other, or 'embedded.' Read without the second relative clause, *balla aiti miniha welendek*, the sentence reads 'the man who owns the dog is a merchant.' The relative clause explicitly expresses the ownership of the dog by the man. With the addition of the second relative clause built on the object of the first relative clause, an approximation of a relativized possessor is formed. In idiomatic English, this sentence would read, 'The man whose dog runs fast is a merchant.' However, Sinhala lacks the possessive relative pronoun, 'whose,' to express the concept of ownership in a relative clause without resorting to embedded relative clauses using the verb *aiti*, 'own,' or implied ownership.

5. HEADLESS RELATIVE CLAUSES. In addition to the normal prenominal relative clauses, a few headless clauses were elicited.

(29) [redi   hodənə]   (kena) Nuwanwə taraha æwisuwə
    clothes wash-PRES-REL one   N.-ACC   anger   induce-PST
    'The one washing the clothes made Nuwan angry.'

(30) [ballǝṭǝ   pain gahanǝ] (puḍgǝleya) [mamǝ   dækkǝpu]   minihai
     dog-DAT kick-PRES-REL person   1SG   see-PST-REL man-FOC
     'The person kicking the dog was the man I saw.'

In both numbers 29 and 30, the sentences were judged grammatical by the consultant with the head nouns in parentheses omitted.

6. CONCLUSION. As a SOV language and in accordance with Greenberg's word-order correlates, Sinhala uses relative clauses that precede the head noun being modified. Sinhala creates prenominal relative clauses using the 'gap' strategy and non-finite verb forms. Ambiguity in the relative clause due to the SOV word order is avoided through the use of case suffixes. Despite the lack of relative pronouns, Sinhala still relativizes on possessors using embedded relative clauses or context. Through a combination of these strategies, Sinhala proves its versatility, allowing speakers the freedom to relativize and thus modify all types of arguments.

REFERENCES

PAYNE, THOMAS E. 1997. *Describing Morphosyntax.* Cambridge: Cambridge University Press.
KEENAN, EDWARD L. and BERNARD COMRIE. 1979. Data on the noun phrase accessibility hierarchy. *Language* 55.333-351.

Rice University
Department of Linguistics, MS 23
6100 Main Street
Houston, TX 77005-1892
zeb@rice.edu

COMPLEMENTATION IN COLLOQUIAL SINHALA: OBSERVATIONS ON THE BINDING HIERARCHY

BENJAMIN WHEELER

Rice University

1. INTRODUCTION. This paper briefly examines the relationship between Sinhala's complementation system and Givón's (1980) Binding Hierarchy. In what ways does Sinhala's complementation system correspond to the Binding Hierarchy and in what ways does it diverge from it? What can examining their relationship reveal about the Binding Hierarchy itself? In my work, I will adopt Noonan's definition of complementation as 'the syntactic situation that arises when a notional sentence or predication is an argument of a predicate' (Noonan 1985:41). When a predicate can take a complement clause as an argument, it is termed a COMPLEMENT-TAKING PREDICATE (CTP). For the purposes of this paper, I will only be analyzing examples of complementation in which the complement clause acts as the object of the predicate. My data shows that Sinhala generally conforms to the Binding Hierarchy, but that certain CTPs exhibit behavior that deviates from it.

1.1. COMPLEMENTS IN SINHALA. Sinhala has five different complement types. The first complement type simply involves the juxtaposition of two clauses, which I will call the JUXTAPOSED-CLAUSE COMPLEMENT. Examples 1 and 2 give instances of the juxtaposed-clause complement.

(1) laməya kukulawə mærənəwa mamə dækka
     child   chicken-ACC kill-NONPST 1SG   see-PST
     'I saw the child kill/killing the chicken'

(2) rošini redi   hodənəwa   maṭə   æhuna
     rošini clothes wash-NONPST 1SG-DAT hear-PST
     'I heard Rošini washing clothes'

Example 1 should be considered a complement because the clause *laməya kukulawə mærənəwa* 'the child kill/killing the chicken' acts as an object of the matrix clause *mamə dækka* 'I saw.' Similarly in 2, *rošini redi hodənəwa* 'Rošini washing clothes' acts as the object of the matrix clause *maṭə æhuna* 'I heard.'

These complement clauses can be considered examples of a SENTENCE-LIKE COMPLEMENT TYPE (s-like). Noonan defines a sentence-like complement clause as one in which 'the predicate has the same syntactic relation to its subject and its other arguments that it has in syntactic main clauses' (Noonan 1985:49). Most languages include not only s-like complement types, but other complement types, called NON-SENTENCE-LIKE COMPLEMENTS (non-s-like), in which the subject does not have the same syntactic relations with its predicate as it would in a main clause. One feature that distinguishes s-like complements from non-s-like complements is the verb forms with which they occur. The juxtaposed-clause complement type takes FINITE VERB FORMS—meaning that the verb is fully inflected for TAM and is used in main clauses. In 1, for instance, the complement clause *laməya kukulawə mærənəwa* 'the child killing the chicken' is a grammatically acceptable clause on its own. Some of the other complement types in Sinhala, however, occur with NON-FINITE VERB FORMS. Non-finite verb forms are not fully inflected for

TAM and are used in subordinate clauses. Complement clauses using these verb forms could not stand alone as independent clauses, and therefore could be considered less s-like.

The second complement type is composed of two clauses joined by the complementizer *kiyəla,* which I will call the *kiyəla* COMPLEMENT.

(3) mamə dækka laməya kukulawə mærəla kiyəla

    1SG   see-PST child   chicken-ACC kill-CON COMP

    'I saw that the child had killed the chicken'

(4) rošini redi   hodənəwa   kiyəla maṭə   æhuna

    rošini clothes wash-NONPST COMP   1SG-DAT hear-PST

    'I heard that Rošini was washing clothes'

Again, these should be considered examples of complementation because the kiyəla clause acts as an argument of the predicate in the matrix clause. Complement clauses using *kiyəla* are s-like and occur with finite verb forms.[1]

Another complement type in Sinhala involves two clauses linked by the complementizer *bawə,* which I will refer to as the *bawə* COMPLEMENT.

(5) ohuge   bahareawət hoyənə   gaman horek tamange kææmə

    3M.SG -GEN wife   look-REL.NONPST while robber himself food

    horəkan kərəpu   bawə ohu   dækka

    steal   do-REL.PST COMP 3M.SG   SEE-PST

    'While looking for his wife the man saw that the robber had stolen his food'

(6) mamə   toppi horəkan kərənə   bawə tirənəya-keruwa

    1SG   hat-PL steal   do-REL.NONPST COMP decide do-PST

    'I decided to steal the hats'

These two are examples of complementation because the *bawə* clause acts as an argument of the predicate in the matrix clause. As in the examples above, *bawə* normally occurs with non-finite verbs. Very rarely *bawə* occurs with finite verbs, such as in 7.

(7) mamə ohu wiləṭə   giya bawə hoyaa-gatta

    1SG   3M.SG lake-DAT go-PST COMP discover take-PST

    'I discovered he went to the lake'[2]

Like kiyəla complements, the bawə complement in 7 occurs with a finite verb form, so that *ohu wiləṭə giya* 'he went to the lake' could be an independent clause. In their grammar of Sinhala, Gair and Paolillo report that *bawə* is 'restricted in use to factual or knowledge contexts' (Gair and Paolillo 1997:53). From my data, this appears to be true, as long as we consider *dakinəwa*

---

[1] Although converbs are not generally considered to be finite verb forms (see 3), in Sinhala they appear to be able to function in this way.

[2] The same sentence could be made with ekə, but the verb form would have to be *giyə*, the relative past form. I did not have enough time to test out all the verb forms that *bawə* complements can use.

'see,' *tirənəyə kərənəwa* 'decide,' and *balaaporoṭṭu wenəwa* 'hope/expect' to belong to 'factual or knowledge contexts.'

The complementizer *ekə* is also used to link two clauses, as in 8 and 9:

(8) mamə redi   hodənə   ekə iwərə keruwa
     1SG   clothes wash-REL.NONPST NOM FINISH do-PST
     'I finished washing the clothes'

(9) mamə wæḍe kərənə   ekə wælækuwa
     1SG   work do-REL.NONPST NOM prevent-PST
     'I avoided doing the work'

Again these are examples in which the complement clause acts as an argument of the predicate in the matrix clause. I will call this complement type the *ekə* COMPLEMENT. The complementizer *ekə* only occurs with relative past and relative non-past verb forms so the complement clauses could not function as independent clauses (i.e. this complement type is less s-like). There is evidence that *ekə* is a nominalizer, as it can take the postposition *gænə*.

(10) waňdura toppi horəkan kərəpu   ekə gænə dukaa unaa
     monkey hat   steal   do-REL.PST NOM about sad   become-PST
     'The monkey was sad about stealing the hats.'

In all three of the examples, the subject of the matrix clause and the subject of the complement clause are the same entity. Because the subjects are co-referential, it is only necessary to identify the subject one time (this is sometimes called EQUI-DELETION). However, the subject of the matrix clause and the subject of the complement clause do not need to be co-referential, as in 11.

(11) æyə wiləṭə yanə   ekə mamə hoyaa-gatta
     3F.SG lake-ACC go-REL.NONPST COMP 1SG discover-take-PST
     'I discovered that she was going to the lake.'

Lastly, Sinhala uses an infinitival verb form and no complementizer in what I will term the INFINITIVE COMPLEMENT.

(12) oyaa redi   hodannə   awašay
     2SG   clothes wash-INF   necessary
     'It is necessary that you wash the clothes.'

(13) oyaaṭə wiləṭə   yannə puluwan
     2SG-DAT lake-DAT go-INF can
     'You can go to the lake.'

(14) reenu ballawə mærennə æriya
     reenu dog-ACC die-INF   let-PST
     'Reenu let the dog die.'

The infinitive phrase in all three of these examples is acting as an argument of the matrix clause, and therefore should be considered a complement. 12 and 13 are examples of sentences in which the subject of the matrix clause and the subject of the complement clause are co-referential, but 14 has different subjects for the two clauses. The infinitive form is a non-finite form, and therefore this complement type should be considered less s-like. Example 14 gives further evidence for this complement type as less s-like. In this sentence, the subject of the complement clause *balla* has been raised to the object of the matrix clause and therefore carries the accusative *-wə*. Therefore, the predicate of the complement clause, *mærennə,* does not have normal syntactic relations with its subject because *balla* is not in the nominative case as it would be in a main clause.[3]

It is interesting to note the different orders in which the clauses in the various complement types appear, described in Figure 1.

Order Name	Order
C+M	COMPLEMENT CL. + MATRIX CL.
M+C	MATRIX CL. + COMPLEMENT CL.
Embedded	MATRIX CL. [$Subj_{matrix}$ + COMPLEMENT CL. + $PRED_{matrix}$]

FIGURE 1. Word Order of Complement Types

Unfortunately, at this point, meaning differences embodied in these different word orders and their pragmatic/discourse functions remain unclear. However, the data does suggest that certain complement types prefer certain word orders. The juxtaposed-clause complement appeared only in the C+M order, but there were very few examples of this complement type. The kiyəla complement appeared in all three word orders, but seemed to prefer the M+C order. The bawə complement appeared equally in the C+M and the embedded word order, but did not appear in the M+C order, while the ekə complement appeared only in the embedded order. Lastly, the infinitive complement appeared in all three orders but strongly favored the embedded word order.

1.2. GIVÓN'S BINDING HIERARCHY. In his article 'The Binding Hierarchy and the Typology of Complements,' Givón (1980) establishes the relationship between 'the semantic structure of complement-taking verbs and the syntactic structure of their complements' (Givón 1980:333). He argues that one can establish a hierarchy that systematically describes this relationship and that this hierarchy is cross-linguistically robust. In terms of the semantic structure, the complements are arranged over three different overlapping semantic scales—epistemic attitude, emotive attitude, and implicativity (Givón 1980:368). Each of these factors bifurcate into high and low categories—weak epistemic, strong epistemic, low emotive, high emotive, strong-attempt and implicative. The syntactic hierarchy codes for four factors—degree of structural integration, degree of freedom of action, degree of freedom of the agent, and use of complementizing subordinators (Givón 1980:371). Givón claims that the semantic categories are represented iconically in the structure of complement clauses. According to the hierarchy, CTPs with weak epistemic attitude will take complements with free clauses (i.e. the

---

[3] Another analysis of this clause is possible where these constructions are actually auxiliaries and therefore are single clauses. In any case, these would be placed at the far end of the Binding Hierarchy. These predicates are so 'bound,' in other words, that they are a single clause.

complement-clause and the main clause are clearly distinguishable and independent from one another) whereas implicative CTPs will tend to occur in complements which are more integrated into the main clause, such as infinitive complements and nominalizations. Below is a reproduction of Givón's chart.
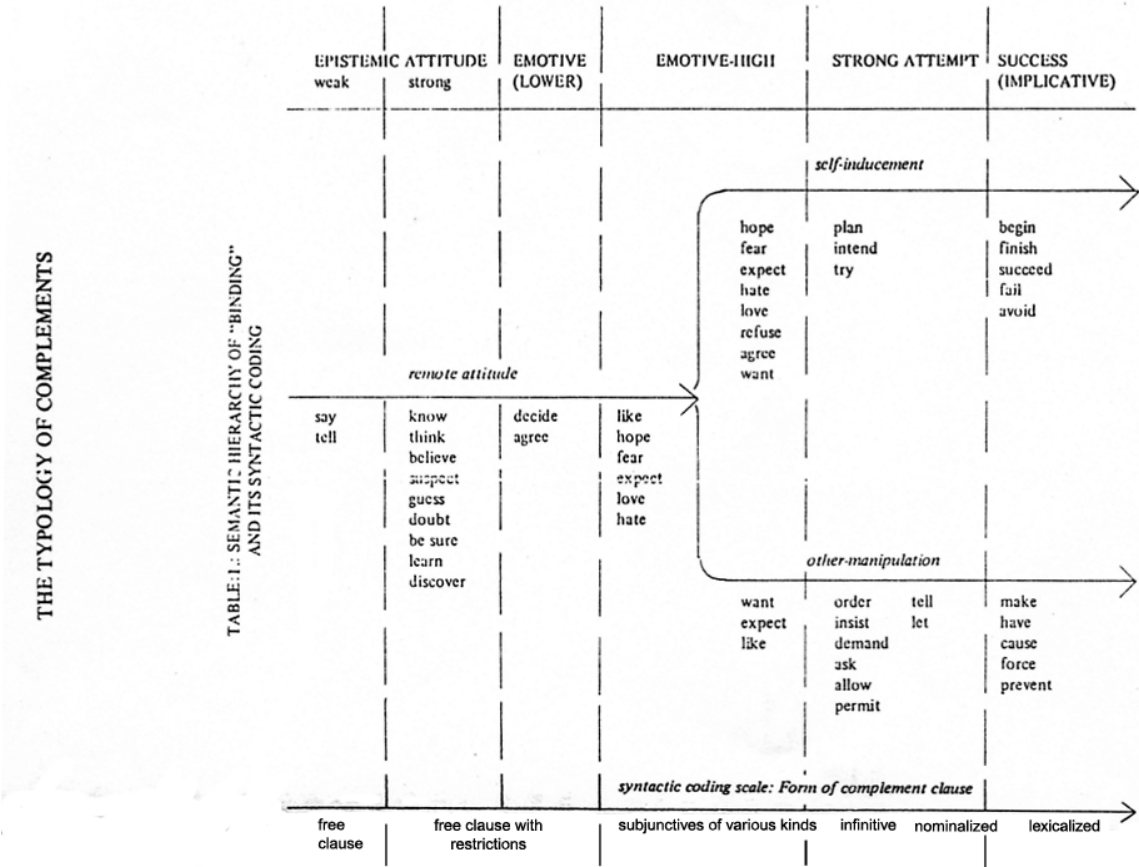

FIGURE 2. Givón's Binding Hierarchy, taken from Givón 1980 p. 369

In examining the Binding Hierarchy and Sinhala I had two main research questions:
  1) What correlations and differences are there between Sinhala's complement system and the Binding Hierarchy's theoretical system?
  2) What can these correlations and differences reveal about the Binding Hierarchy as a whole?

  2. THE BINDING HIERARCHY IN SINHALA. To study the Binding Hierarchy in Sinhala, I elicited data for thirty CTPs. For each CTP I looked at which types of complements each predicate could take. The results are listed in Figure 3. The CTPs are sorted first according to Givón's semantic scales and next according to the complement types with which they can occur. Structurally I have organized the complement clauses left to right from most independent to least independent (from free clauses to bound clauses).
  At the far left I put the juxtaposed-clause complement because the matrix clause and the complement clause undergo no structural integration, and the verb in the complement clause

can have its own independent TAM marking. Furthermore, the juxtaposed-clause complement tends to favor word orders where the complement clause and the matrix clause are distinct. Unlike some of the complement types farther down on the scale, this complement type occurs with finite verb forms. Putting this complement type above the kiyəla complement may seem at first to be a contradiction to Givón's claim of iconicity because there is no complementizer to separate the clauses. However, it could be argued that the complementizer acts as a kind of subordinator, marking one clause as predicating another, and therefore this complement type should be considered to be more independent than the kiyəla complement. The same would not be true for complement types farther down on the scale because of their restrictions and degree of integration into the matrix clause.

Next I have placed the kiyəla complement. As in the juxtaposed-clause complement, the matrix clause and the complement clause in the kiyəla complement type are independent. There is no evidence of structural integration and the verb in the complement takes independent TAM markings and finite verb markings. In addition, this verb occurs only in word orders where the matrix clause and the complement clause are clearly distinguishable.

After the kiyəla complement, I have placed the bawə complement type. The verbs in this complement clause can occur in both finite and relative verb forms, but they strongly prefer relative verb forms. The use of relative verb forms affects the independence of the matrix clause—it could not stand alone as an independent clause. Last, the bawə complement occurs both in word orders where the matrix and the complement clauses are distinct and where the complement clause is embedded in the matrix clause. Therefore, the bawə complement is subject to more structural integration than the kiyəla complement or the juxtaposed-clause complement.

To the right of the bawə complement I put the ekə complement. The ekə complement is even less independent than the bawə complement because it can only occur with relative verb forms. In addition, it only occurs in the embedded word order, which shows that it is less structurally independent than the complement types above it on the scale. On the very end I put the infinitive complement because it occurred with only one verb form, which does not take independent TAM markings. The TAM of the complement clause is therefore determined by the TAM marking in the matrix clause. Like the ekə complement, the infinitive complement strongly prefers the embedded word order.

The data suggests that Sinhala generally conforms to Givón's hierarchy. For instance, the weak epistemic verb *kiyənəwa* 'to say or to tell' can only occur with the kiyəla complement, as shown in Ex. 15.

(15) nuwan redi   heduwa   kiyəla sarat kiyənəwa
      nuwan clothes wash-PST COMP   sarat say-NONPST
      'Sarat says that Nuwan washed the clothes.'

On the opposite side of the scale, the implicative, other-manipulation CTP *kriyə kiyənəwa* 'to cause' can only take the infinitive complement.

(16) ohu   gaha mærennə kriya keruwa
      3M.SG tree die-INF   cause do-PST
      'He caused the tree to die.'

Verb	Gloss		Semantic Scale	0	kiyəla	bawə	ekə	inf
kiyənəwa	say, tell		Epistemic-Weak	-	g	-	-	-
kæhægæhuwa	shout		Epistemic-Weak	-	g	-	-	-
dakinəwa	see		Epistemic-Strong	g	g	g	-	-
æhenəwa	hear		Epistemic-Strong	g	g	-	-	-
hitənəwa næhæ	doubt		Epistemic-Strong	-	g	-	-	-
hoyaagatta	discover		Epistemic-Strong	-	g	g	g	-
dannəwa	know		Epistemic-Strong	-	g	g	gwg	-
hitənəwa	think		Epistemic-Strong	-	g	-	gwg	gcs
matəkə tiyənəwa	remember		Epistemic-Strong	g	g	g	g	gcs
tirənəyə kərənəwa	decide		Emotive-Low	-	g	g	gwg	gcs
dukai	sad		Emotive-High	-	g	-	gwg	gcs
santosai	happy		Emotive-High	-	g	-	gwg	gcs
bəyai	afraid		Emotive-High	-	g	-	gwg	gcs
dukaa unaa	regret		Emotive-High	-	g	-	gwg	-
kalpəna kərənəwa	imagine/dream		Emotive-High	-	g	-	gwg	-
balaaporoṭṭu wenəwa	hope/expect		Emotive-High	-	g	g	-	gcs
kæməti	like		Emotive-High	-	g	-	-	gcs
aasay	love		Emotive-High	-	-	-	-	gcs

oone	want		Emotive-High	-	-	-	-	gcs
hædənəwa	try	si	Strong Attempt	-	-	-	-	gcs
æhuwa	ask		Strong Attempt	-	-	-	gwg	g
arinəwa	let		Strong Attempt	-	-	-	-	g
dunna	allow	om	Strong Attempt	-	-	-	-	g
puluwan	able to		Strong Attempt	-	-	-	-	g
awašay	to be necessary		Strong Attempt	-	-	-	-	g
wælækuwa	prevent	om	Implicative	-	g	-	g	-
næwætuwa	stop	om/si	Implicative	-	g	-	g	-
iwərə kərənəwa	finish	si	Implicative	-	-	-	g	gcs
balə kərənəwa	force	om	Implicative	-	-	-	-	g
kriya kərənəwa	cause	om	implicative	-	-	-	-	g

g=grammatical, -=ungrammatical, gwg=grammatical only with gænə, gcs=grammatical only if the subject of complement clause and main clause correspond, si=self-induced, om=other-manipulation

FIGURE 3. Sinhala Complementation Data

Furthermore, note that the CTPs at the top of the table cannot take the infinitive complement. Similarly, most of the CTPs at the bottom of the scale cannot take the kiyəla complement.

The mid-range of the hierarchy involves quite a bit of overlap and requires some explanation. Overlap across complement types is not surprising—Givón notes in his paper that the various scales overlap each other. Some CTPs can use both kiyəla and ekə complement types, but in order for the clause to be acceptable, ekə must be followed by the postposition *gænə* 'about.' The CTP *dannəwa* 'to know' is one example of this phenomenon.

(17) mamə dannəwa   waňdura toppi issuwa   kiyəla
    1SG   know-NONPST monkey hat-PL steal-PST COMP
    'I know that the monkey stole the hats.'

(18) waňdura kehel   geḍi wələṭə kæməti bawə mamə dannəwa
    monkey bananas CL   for   like   COMP 1SG   know-NONPST
    'I know that the monkey likes bananas.'[4]

(19) eyaa gedərə   yanə   ekə gænə dannəwa
    3M.SG home-DAT go-REL.NONPST COMP about know-NONPST
    'He knows about going home.'[5]

(20) *eyaa gedərə   yanə   ekə dannəwa
    3SG:M home-DAT go-REL.NONPST COMP know-NONPST

The CTPs that can only occur with *ekə gænə* are distributed mainly in the emotive portion of the semantic scale. The restricted use of *ekə* with the postposition is not unique to Sinhala. For instance, we see the same pattern in the English verb *know*.

(21) I know about going home.
(22) *I   know going home.

Certain CTPs can only occur with the infinitive if the subject for the main clause and the subject for the complement clause are co-referential. An example of such a CTP is *hitənəwa* 'to think,' which can also occur with *kiyəla* and *ekə gænə*.

(23) mamə ohu wiləṭə   giya   kiyəla hituwa
    1SG   3SG:M lake-ACC go-PST COMP   think-PST
    'I thought that he went to the lake.'

(24) mamə wiləṭə   yanə ekə gænə hituwa
    1SG   lake-DAT go   COMP ABOUT think-PST
    'I thought about going to the lake.'

---

[4] Note that *kæməti* is a quasi-verb which does not take any TAM markings. There are a large number of these quasi-verbs that are CTPs. As they do not behave differently than the regular verbal CTPs, I did not treat them differently in the data.

[5] Restrictions on the subject in this construction need to be investigated. It is not clear whether the subjects of the two clauses must be co-referential.

(25) mamə wilətə   yannə hituwa
     1SG   lake-ACC go-INF think-PST
     'I thought about going to the lake.'

(26) *mamə ohu wilətə yannə hituwa

This restriction with infinitives does not apply lower on the hierarchy. For instance, the verb *balə-kərənəwa* 'to force' does not have the same restriction.

(27) mamə nuwantə   wæḍe kerannə balə-kərənəwa
     1SG   nuwan-DAT work do-INF   force do-NONPST
     'I will force Nuwan to do the work.'

This is an example of what Givón would call an OTHER-MANIPULATION IMPLICATIVE CTP. The semantics of the complement itself therefore may restrict it so that it requires an explicit subject in the complement clause. Thus, 'I forced myself to do the work' may require (as in English) a reflexive pronoun. This needs to be investigated further because I did not check this in my elicitation sessions. Still, we can see the progression of the hierarchy is generally preserved with infinitive complements—the top of the hierarchy cannot take infinitive complements, the middle can take the infinitive complement when the subjects of the main clause and the complement clause are co-referential, and the bottom can take infinitive complements when the subject is different. Similarly, kiyəla complements can occur with the epistemic and emotive complements, but cannot occur with the strong attempt and implicative complements. The bawə complements only occur with a very limited number of CTPs. Last, the ekə complements do not occur at all at the top of the chart, occur with *gænə* in the middle of the chart, and appear alone towards the bottom of the chart. Thus we see different parts of the hierarchy patterning similarly in terms of the complement types they can take.

2.1. COMPLICATIONS IN THE DATA. There are, however, some CTPs which disrupt the tidy progression of the hierarchy. *Kiyənəwa* is used both in the sense of 'say' and in the sense of 'tell' so that it can be used to describe an indirect order. Givón puts *tell* both in the weak epistemic category at the top of the chart and in the strong-attempt, other-manipulation category at the bottom of the chart. We might then expect to find a point lower on the structural scale to code for indirect orders. Instead, we find that it can only be used with *kiyəla*, even for indirect orders, as in 28.

(28) mamə kiwwa nuwantə   wæḍe kərannə kiyəla
     nuwan say-PST nuwan-DAT work do-INF   COMP
     'I told Nuwan to do the work.'

From this example we might think that the verb forms within the complement clause are restricted to the infinitive, but it turns out that other forms may be used with this CTP as well:

(29) mamə nuwanṭə   kiwwa oyaa wæḍe kərənəwamay   kiyəla
    nuwan nuwan-DAT say-PST 3SG:M work do-NONPST-EMPH   COMP
    'I told Nuwan that he will do the work'

The quasi-verb *kæmeti* 'to like' and the verb *balaaporoṭṭu-wenəwa* 'to hope/expect' show another deviation from the binding hierarchy. These verbs may take both kiyəla and infinitive complement types. One would expect that they would also be able to take ekə complement types, as is true for the other emotive CTPs on the chart that take both of these complement types. Instead, we find that these CTPs skip over sections of the hierarchy rather than overlapping them.

(30) mamə kæməti waňdura toppi issuwa   kiyəla
    1SG   like   monkey hat-PL steal-PST COMP
    'I like it that the monkey stole the hats'

(31) mamə toppi ussənnə kæməti
    1SG   hat-PL steal-INF like
    'I like to steal hats.'

(32) *mamə kæməti waňdura toppi ussənə   ekə
    1SG   like   monkey hat-PL steal-REL.NONPST COMP

(33) mamə ohu wiləṭə   yanəwa   kiyəla balaaporoṭṭu-wenəwa
    1SG   3SG:M lake-DAT go-NONPST COMP   hope become-NONPST
    'I expect that he will go to the lake.'

(34) mamə wiləṭə   yannə balaaporottu-wenəwa
    1SG   lake-DAT go-INF hope become-NONPST
    'I hope/expect to go to the lake.'

(35) *mamə wiləṭə   ohu yanə   ekə balaaporottu-wenəwa
    1SG   lake-DAT 3M.SG go-REL.NONPST COMP hope become-NONPST

Two of the implicative verbs are also not where we would expect them to be on the hierarchy—*wælækuwa* 'to prevent/avoid' and *næwætuwa* 'to stop.'[6] Given that these are other-manipulation implicative verbs, we would expect them to occur in structurally bound complement clauses. The other verbs on this end of the semantic scale occur with ekə and/or infinitive complements. As it turns out, these verbs can only occur with *kiyəla* and with *ekə*.

(36) mamə waňdura toppi usənəwa   kiyəla wælækuwa
    1SG   monkey hat-PL steal-NONPST COMP prevent-PST
    'I prevented the monkey from stealing the hats.'

(37) mamə waňdura toppi ussənə   ekə wælækuwa
    1SG   monkey hat   steal-REL.NONPST COMP prevent-PST
    'I prevented the monkey from stealing the hats.'

---

[6] The present tense forms of these verbs are unknown.

(38) mamə laməya wiləṭə   yanəwa kiyəla næwætuwa
      1SG   child lake-DAT go   COMP stop-PST
      'I stopped the child from going to the lake.'

(39) mamə laməya wiləṭə   yanə   ekə næwætuwa
      1SG   child lake-DAT go-REL.NONPST COMP stop-PST
      'I stopped the child from going to the lake.'

The language consultant described the sentences using *kiyəla* as having a different meaning than the sentences using *ekə,* but the exact meaning difference remains unclear. In any case, the meaning difference between the two suggests that when CTPs can take different complement types, the use of one complement type over another is not simply subject to free variation. Rather, speakers may choose different complement types depending on the semantic, pragmatic and discourse variables.

3. CONCLUSION. This paper has described, in brief, the complement system of Sinhala and the ways in which it conforms to and deviates from Givón's Binding Hierarchy. From the data, CTPs in Sinhala tend to distribute along the Binding Hierarchy as Givón predicts—CTPs that are epistemic or emotive tend to take complement types that are more biclausal in nature while CTPs that are implicative tend to have complement clauses more tightly bound into the main clause. In Sinhala, this is reflected syntactically in the type and range of verb forms a complement type allows (finite vs. non-finite) and the preferred word order of a given complement type (biclausal versus embedded). However, my data also reveals CTPs that deviate from the Binding Hierarchy by taking complement types that would not be predicted from their semantic classification. This data does not necessarily undermine the validity of the Binding Hierarchy, but it does have implications for the use of the Binding Hierarchy. Based on language data, the Binding Hierarchy should be understood as a general pattern and not as a predictive formula or universally applicable rule. Although a typological hierarchy, like Givón's binding hierarchy, can show trends across languages, actual language data is complex, and will never conform completely to theory.

More work needs to be done on Sinhala's complementation system. The differences in meaning as well as pragmatic and discourse functions of the various permissible word orders warrants further study. As yet we do not have a full understanding of the exact syntactic relations in these constructions, and the meaning differences between the various complement types are still unclear. Furthermore, most of the data in this study was gathered through elicitation sessions rather than through data of language-in-use. As is widely recognized, language-in-use data often varies considerably from elicited data. Making grammaticality judgments in context-free environments is difficult. Unfortunately, our limited corpus of language-in-use provided few examples of complementation and was restricted to the storytelling genre. For a truly comprehensive study of Sinhala's complementation system and the Binding hierarchy one would want to include data from spontaneous discourse from a variety of genres as well. [7]

---

[7] Many thanks to Nissanka Wickremasinghe for his patience and dedication during long data elicitation sessions.

REFERENCES

GAIR, JAMES W. and JOHN C. PAOLILLO. 1997. *Sinhala.* München: Lincom Europa.

GIVÓN, TALMY. 1980. The Binding Hierarchy and the Typology of Complements. *Studies in Language.* 4.3.333-377.

NOONAN, MICHAEL. 1985. Complementation. *Language Typology and Syntactic Description: Complex Constructions* Vol. 2, ed. by Timothy Shopen, 42-140. Cambridge, UK: Cambridge University Press.

Rice University
Department of Linguistics, MS 23
6100 Main Street
Houston, TX 77005-1892