

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 it 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 may 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) <i>kumbi-ya</i> <i>harək-a</i> <i>maalu-wa</i> <i>laməy-a</i>	(stem) <i>kumbi</i> <i>harak</i> <i>maalu</i> <i>lamai</i>	'ant' 'cow' 'fish' 'child'
I1	Inanimate	(stem + -ə) <i>raum-ə</i> <i>æs-ə</i> <i>huləjg-ə</i> <i>taruw-ə</i> <i>suli-yə</i> <i>liñd-ə</i>	(stem) <i>raum, rauj</i> <i>æs</i> <i>huləj</i> <i>taru</i> <i>suli</i> <i>lij</i>	'circle' 'eye' 'wind' 'star' 'current' '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) <i>taata</i> <i>aaci</i> <i>duwə</i> <i>raaləhaami</i> <i>ætinni</i>	(stem + -la) <i>taata-la</i> <i>aaci-la</i> <i>duwa-la</i> <i>raaləhaami-la</i> <i>ætinni-la</i>	'father' 'grandmother' 'daughter' 'police officer' 'female elephant'
A3 ¹	Animate	(stem) <i>iiri</i> <i>birində</i> <i>ætinni</i>	(stem + -n) <i>iirii-n</i> <i>birinda-n</i> <i>ætinnii-n</i>	'sow' 'wife' '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) <i>duwə</i> <i>mawə</i> <i>piya</i>	(stem + -wəru) <i>duu-wəru</i> <i>mau-wəru</i> <i>piya-wəru</i>	'daughter' 'mother' '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) <i>walah-a</i> <i>makər-a</i> <i>put-aa</i> <i>leen-a</i> <i>riləw-a</i> <i>hiwəl-a</i> <i>næədæə-ya</i>	(stem + gem + -u) <i>walass-u</i> <i>makər-u</i> <i>putt-u</i> <i>leenn-u</i> <i>rila-u</i> <i>hiwall-u</i> <i>næədæə-yu</i>	'bear' 'dragon' 'son' 'squirrel' 'rhesus monkey' 'wolf' 'relative'
A6	Animate	(stem + -a) <i>kurull-a</i> <i>waluur-a</i> <i>mu-wa</i> <i>puus-a</i> <i>ukun-a</i> <i>gowi-ya</i>	(stem + -o) <i>kurull-o</i> <i>waluur-o</i> <i>mu-wo</i> <i>puus-o</i> <i>ukun-o</i> <i>gowi-yo</i>	'bird' 'boar' 'deer' 'cat' 'louse' 'farmer'
A7 ²	Animate	(stem + -i) <i>gææn-i</i>	(stem + -u) <i>gææn-u</i>	'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 næw-ə	(stem + dgm + -u) walaakul-u deedun-u akun-u kañd-u næ-u	‘cloud’ ‘rainbow’ ‘lightning’ ‘mountain, hill’ ‘ship’
I3	Inanimate	(stem + -ə) raʔ-ə lookə-yə æł-ə	(stem + -əwal) raʔ-əwal look-əwal æł-əwal	‘country, nation’ ‘world’ ‘stream, brook’
I4	Inanimate	(stem + gem + -ə) wæss-ə pinn-ə diyə æll-ə	(stem + -i) wæs-i, wæh-i pin-i diyə æł-i	‘rain’ ‘dew’ ‘waterfall’
I5 ⁴	Inanimate	(stem + -ə) gaŋg-ə	(stem + -aa) gaŋg-aa	‘river’

TABLE 3. Examples of nouns with replacive patterns⁵

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/kañdu*). 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’ *wæssə/wæhi*), 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 ællə/diyə æli* contains the word for ‘stream,’ which is *æłə* (with two possible plural forms *æłə* or *æłəwal*).

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) laməy-a maawə dækka
 child-SG.DEF 1SG.ACC see.PST
 ‘The child saw me.’
- (2) laməy=ek maawə dækka
 child=SG.IND 1SG.ACC see.PST
 ‘A child saw me.’
- (3) lamai maawə dækka
 child.PL 1SG.ACC see.PST
 ‘The children saw me.’ or ‘Children saw me.’
- (4) maŋ gaŋg-ə dækka
 1SG river-SG.DEF see.PST
 ‘I saw the river.’
- (5) maŋ gaŋg=ak dækka
 1SG river=SG.IND see.PST
 ‘I saw a river.’
- (6) maŋ gaŋg-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 be 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*, *-wəru*, *-u*, *-o*, *-əwal*, *-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*, *-ə*, 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	<i>laməy-a</i>	<i>laməy=ek</i>	<i>lamai</i>	'child'
	<i>birind-ə</i>	<i>birind=ak</i>	<i>birind-an</i>	'wife'
	<i>duwə⁶</i>	<i>duw=ek</i>	<i>duwə-la,</i> <i>duu-wəru</i>	'daughter'
	<i>noona</i>	<i>noona kenek</i>	<i>noona-la</i>	'wife'
	<i>dostərə</i>	<i>dostərə kenek</i>	<i>dostərə-la, dostərə-wəru</i>	'doctor'
	<i>minih-a</i>	<i>minih=ek</i>	<i>miniss-u</i>	'man'
Inanimate	<i>buuru-wa</i>	<i>buuru=wek</i>	<i>buuru-wo</i>	'donkey'
	<i>ætinni</i>	<i>ætinni=yek</i>	<i>ætinni-in</i>	'f. elephant'
	<i>meesə-yə</i>	<i>meesə=yak</i>	<i>meesə</i>	'table'
	<i>gal-ə</i>	<i>gal=ak</i>	<i>gal</i>	'stone'
	<i>raṭə</i>	<i>raṭ=ak</i>	<i>raṭə-wal</i>	'country, nation'
	<i>dawəs-ə</i>	<i>dawəs=ak</i>	<i>dawas</i>	'day'
<i>pot-ə</i>	<i>pot=ak</i>	<i>pot</i>	'book'	
<i>gaṅg-ə</i>	<i>gaṅg=ak</i>	<i>gaṅg-aa</i>	'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, *rajjakenek* '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 *-wəru* 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 gaŋgaa kiipəyak dækka
 1SG river.PL a.few.IND see.PST
 'I saw some (unspecified) rivers.'
- (10) maŋ gaŋgaa tunak dækka
 1SG river.PL three.IND see.PST
 'I saw three rivers.'
- (11) maŋ gaŋgaa 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) maṅ 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	=tə	=tə	=ek(o)tə ⁸	=əkətə	=tə ⁹	wələtə
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 =ektə. However, *duwə* ‘daughter’ takes =ekoṭə.

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:

mæssā buuruwage æṅṅee wəhuwa
 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 lamaŋgeŋ 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 gəhuwa
 3SM stick.PL PL.INAN.INST donkey.SG.ANIM.DEF.DAT hit.PST
 ‘He hit the donkey with sticks.’
- d ohu kootu kiipəyəkiŋ buuruwaṭə gəhuwa
 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 gəhuwa
 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.ANIM	using	cart.SG.INAN.DEF	pull-PST

‘He pulled the cart using donkeys.’

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

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'	<i>lamə-ya=ge</i>	<i>lamə-ya=geŋ</i>	<i>lamə-ya</i>	<i>lama-iŋ=ge</i>	<i>lama-iŋ=geŋ</i>	<i>lamai</i>
'dog'	<i>ball-a=ge</i>	<i>ball-a=geŋ</i>	<i>ball-a</i>	<i>ball-aŋ=ge</i>	<i>ball-aŋ=geŋ</i>	<i>ball-o</i>
'man'	<i>minih-a=ge</i>	<i>minih-a=geŋ</i>	<i>minih-a</i>	<i>miniss-uŋ=ge</i>	<i>miniss-uŋ=geŋ</i>	<i>miniss-u</i>

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'	<i>olu=we</i>	<i>olu=weŋ</i>	<i>olu-wə</i>	<i>olu wələ</i>	<i>olu wəliŋ</i>	<i>olu</i>
'knife'	<i>pihi=ye</i>	<i>pihi=yeŋ</i>	<i>pihi-yə</i>	<i>pihi wələ</i>	<i>pihi wəliŋ</i>	<i>pihi</i>
'table'	<i>mees=e</i> (<i>meesə=ye</i>)	<i>meesə=yeŋ</i>	<i>mees-e</i> (<i>meesə-yə</i>)	<i>meesə wələ</i>	<i>meesə wəliŋ</i>	<i>meesə</i>

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' *raɕə* (which is *raɕəwal* in the plural but *raɕə 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' *kootə*, which is *kootu* in the plural and *kootu wəliŋ* 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=ɛŋ*, 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 (<i>buuruwa</i>)	Animate (<i>buuruwek</i>)	Animate (<i>buuruwo</i>)
ACC	=wə (<i>buuru-wa=wə</i>)	=ekwə (<i>buuru=wekwə</i>)	=wə (<i>buuru-waŋ=wə</i>)

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 *=tə* 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 (<i>buuruwa</i>)	Inanimate (<i>pæænə</i>)	Animate (<i>buuruwek</i>)	Inanimate (<i>pæænak</i>)	Animate (<i>buuruwo</i>)	Inanimate (<i>pææŋ</i>)
DAT	=tə (<i>buuru-wa=tə</i>)	=tə (<i>pææn-ə=tə</i>)	=ek(o)tə (<i>buuru-wektə</i>)	=əkətə (<i>pææn-əkətə</i>)	-Vn=tə (<i>buuru-wan=tə</i>)	wələtə (<i>pææŋ wələtə</i>)

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 *buuruwa*tə is easily segmented into the noun root *buuru* ‘donkey,’ the singular marker *-wa*, and the dative marker =tə). For singular indefinite, =tə 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 =tə represents dative case. For plurals, =tə 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 =tə. However, *wələ* alone is the LOC/GEN marker. It clearly does not represent that here. In this case, it seems that the segmentability of *wələ*tə 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 =tə 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 (<i>buuruwa</i>)	Inanimate (<i>pæænə</i>)	Animate (<i>buuruwek</i>)	Inanimate (<i>pæænak</i>)	Animate (<i>buuruwo</i>)	Inanimate (<i>pææŋ</i>)
LOC/GEN	=ge (<i>buuruwage</i>)	=e (<i>pææne</i>)	=ekge (<i>buuruwekge</i>)	=əkə (<i>pæænəkə</i>)	=Vŋge (<i>buuruwanje</i>)	wələ (<i>pææŋ wələ</i>)

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 *wəlatə* in the previous section in that the form appears segmentable, but the segmentation produces homophony within the paradigm (*wəlatə* represents LOC/GEN in some forms and plural inanimate in others, and *-ə* represents LOC/GEN in 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 *wəlatə* could be segmented into *wəl-* for inanimate plural and *-ə* 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 throughout¹² (though it may be preceded by other material). Table 13 shows the INST/ABL forms with examples, using *buuruwa* ‘donkey’ for animate forms and *pæænə* ‘bread’ for inanimate forms.

Case	Singular Definite		Singular Indefinite		Plural	
	Animate (<i>buuruwa</i>)	Inanimate (<i>pæænə</i>)	Animate (<i>buuruwek</i>)	Inanimate (<i>pæænək</i>)	Animate (<i>buuruwo</i>)	Inanimate (<i>pææŋ</i>)
INST/ABL	=geŋ (<i>buuruwageŋ</i>)	=eŋ/iŋ (<i>pææneŋ</i>)	=ekgeŋ (<i>buuruwekgeŋ</i>)	=əkikŋ (<i>pæænəkikŋ</i>)	=Vŋgeŋ (<i>buuruwageŋ</i>)	wəlŋ (<i>pææŋ wəlŋ</i>)

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 =ək changing to =ək (due to the previously mentioned alternation between [ə] 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 wəl- 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 phonological rule or noun class. For example *potə* ‘book’ and *atə* ‘hand’ are both inanimate singular nouns and similar in phonology, but one takes the -iŋ and the other takes -eŋ (*potēŋ* ‘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	1
ACC	2	1	--	--	2	1	--	--	2	1	--	--	--
DAT	2	1	2	1	2	1	2	1	2	1	2	2	2
LOC	--	--	1.5	1	--	--	1.5	1	--	--	2	2	2
GEN	2	1	1.5	1	2	1	1.5	1	2	1	2	2	2
INST	--	--	1.5	1	--	--	1.5	1	--	--	2	2	2
ABL	2	1	1.5	1	2	1	1.5	1	2	1	2	2	2

TABLE 14¹³. 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 languages 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