

THE LEXICAL CATEGORY AUXILIARY IN SINHALA¹

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

¹ The author wishes to thank Nissanka Wickremasinghe for bearing with a seemingly endless series of questions about his native language. I am also indebted to Robert Englebretson, all students of the 2004–2005 field methods class at Rice, and the organizers and participants of the 2005 UCSB workshop on Sinhala linguistics. I received many helpful comments on the present paper. The remaining inadequacies are, of course, my own. Thanks are due to Rice University and the German Academic Exchange Service (DAAD).

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
<i>yannə</i>	go	FUTURE	INF
<i>dennə</i>	give	PERMISSIVE	INF
<i>patan gannə</i>	start take	INCEPTIVE	INF
<i>wennə</i>	become	EQUATIVE	CLAUSE
<i>nəhə</i>	not	NEGATION	FOCUS PHRASE, AUX
<i>bəhə</i>	impossibly	EPISTEMIC	INF, AUX
<i>əti</i>	definitely	EPISTEMIC	INF, VERB PHRASE, AUX
<i>puluwaŋ</i>	possibly	EPISTEMIC	INF, AUX, CLAUSE
<i>kəməti</i>	like	AFFECTION	INF, CLAUSE
<i>kanəgatu</i>	sorry	REGRET	INF
<i>bayə</i>	afraid	FEAR	INF, CLAUSE
<i>oone</i>	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 (*bəhə*, *nəhə*, etc.), and STANCE ELEMENTS (*kəməti*, 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 auxiliatation, 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 auxiliatation, 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 auxiliatation 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 auxiliatation 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 auxiliatation 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 gedyak kəwwa
 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 maṭə epel gedyak 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 kaɖaŋ wəʃennə 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 *pənnə* ‘jump’ in conjunction with *yannə* has acquired the meaning ‘escape’:

- (4) e. gamba botələyen pənəla 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 maʃə epel gediyak kannə dunna
 3SG 1SG-DAT apple CL-IND eat-INF PRM-PST
 ‘He let me eat an apple.’
 b. ohu maʃə 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 gedyak 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 behaviour 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 bohoma 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 gedyak 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 auxiliiation, 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 *nəhə* 'not' and *bəhə* '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 decategorized 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 *nəhə* '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. laməyek nəhə
 child-IND NEG
 'There is no child.'
- b. eyaaṭə laməyek nəhə
 she- DAT child-IND NEG
 'She has no child.' (lit. There is no child to her.)

Evidence for a verb-like status of *nəhə* 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
 næti *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* *wəḍe* *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 oonəhə
 1SG-DAT apple CL-IND need-NEG
 ‘I don’t need an apple.’
- j. maṭə epel gediyak oone unnəhə
 1SG-DAT apple CL-IND need become-PST-NEG
 ‘I didn’t need an apple.’

Third, it is problematic to view *nəhə*, *nəti*, 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 *nəhə* 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 *bəhə* ‘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 *bəhə* does not code deontic modality, it is confined to epistemic meaning. The element is morphologically almost identical to *nəhə*, 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 *nəhə*, *bəhə* has an alternate form. The alternative form *bəri* occurs in subordinate clauses, as shown in 12d.

- (12) a. ohu horek wennə bəhə
 3SG robber-IND EQ-INF IMPOSS
 ‘He can’t be a robber.’
- b. ohuṭə tibaha wennə bəhə
 3SG-DAT thirsty EQ-INF IMPOSS
 ‘He can’t be thirsty.’
- c. ohu wəde kərannə bəhə
 3SG task do-INF IMPOSS
 ‘He can’t possibly do the work.’
- d. ohu horek wennə bəri hinda ohu hire yanne nəhə
 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, *bəhə* shows some parallels with *nəhə* 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 *əti* ‘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, *əti* requires a copula in predicate nominals, just like *bəhə*. Example 13b illustrates how *əti* can modify regular non-finite verb phrases. What sets *əti*

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 jiiṭ 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 *kəməti* ‘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, *kəməti*, *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 *kæmæti* ‘like’ can take nominal arguments, and clausal and infinitive complements. These are illustrated in 15(a-c). The subjects of sentences with *kæmæti* are in the nominative. If *kæmæti* 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 kæmætii*
 1SG animal.PL like-ASS
 ‘I like animals.’
- b. *ballat ekə sellam kərənə ekəṭə ɬamay laməya kæmætii*
 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 kæmætii*
 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 kæmæti nəhæ*
 1SG animal.PL like NEG
 ‘I don’t like animals.’
- e. *redi hodəpu ekəṭə ɬamay mamə kæmæti une*
 clothes wash- PST.REL COMP ASS 1SG like EQ-PST.FOC
 ‘It was washing clothes that I liked.’
- f. *mamə satuntə kæmæti wennə yanne nəhæ*
 1SG animal.PL like EQ-INF FUT-FOC NEG
 ‘I will not like animals.’

Gair and Paolillo (1997:26) present evidence that *kæmæti* inflects for syntactic context in the same way that lexical verbs and the epistemic elements *æti* and *nəhæ* 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 | <i>kæmæti</i> | <i>kæmæti</i> |
| CONDITIONAL | <i>kæmættot</i> | <i>kæmæti unot</i> |
| CONCESSIVE | <i>kæmættat</i> | <i>kæmæti unat</i> |
| FOCUS | <i>kæmætte</i> | <i>kæmæti une</i> |

In summary, *kæmæti* 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 *kæmæti* 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. maṭə bayai
 1SG-DAT fear-ASS
 'I am afraid.'
- e. mamə gederə yannə bayə nəhə
 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. maṭə epel gedyak oone
 1SG-DAT apple CL-IND need
 'I need an apple.'
- b. redi hodənə ekəṭə ṭamay maṭə onee
 clothes wash-REL.PRES COMP ASS 1SG-DAT need
 'It is washing clothes that I want.'
- c. oyaatə 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. maṭə epel gedyak oone wey
 1SG-DAT apple CL-IND need EQ-FUT
 'I will need an apple.'
- e. maṭə epel gedyak 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, *kəməti*, *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. maṭə epel gedyak oonəhə
 1SG-DAT apple CL-IND need-NEG
 'I don't need an apple.'

In summary, despite a number of morphological and syntactic differences, *kæməti*, *baya*, *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 auxiliatation.

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 (*baya*) or adjectives (*kanagatu*) 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.
- BENEVIESTE, EMILE. 1968. Mutations of linguistic categories. *Directions for historical linguistics: 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