Focus position in SOV languages

Susan C. Herring
University of Texas at Arlington

John C. Paolillo
University of Texas at Arlington

1. Introduction: focus and word order*

Many languages manifest a preference for placing 'focused' information —
that is, information to which the speaker/writer wishes to accord particular
salience — in a special position in the sentence, commonly referred to as
'focus position'. In English, especially in written English, where intonation is
not available as a cue to signal saliency differences, the primary focus
position is claimed to be located at or near the end of the sentence (Halliday
1967; Leech and Short 1981). There are plausible cognitive and functional
motivations for placing salient information in sentence-final position: people
better retain in short-term memory that which they have heard or read most
recently, a tendency which takes on functional significance when the final
focused element becomes the topic of the subsequent discourse, as is often the
case. It should not be surprising, therefore, that sentence-final focus is not
restricted to English, but has also been documented in other languages of the
SVO type (e.g. French, Mandarin Chinese), and indeed in languages of
different word order types as well.

The present study is concerned with the location of primary focus posi-
tion in languages whose basic word order is SOV. Verb-final languages pose
a unique challenge to the cross-linguistic tendency towards sentence-final
focus, since in many such languages, there is a requirement that the finite verb
appear in absolute final position; that is, postposed orders such as (X)VO and
(X)VS are excluded, except in cases where the postposed element is an

intonationally-backgrounded 'afterthought'. Given that focus of the sort we are concerned with crucially involves constituents other than the finite verb, the question naturally arises as to where focus position in SOV languages is located.

1.1 *Two conflicting views*

Readers familiar with the literature on word order and pragmatic focus may feel that this question has already been answered. In fact, however, two different answers have been proposed which taken together make conflicting claims. The most widely-held view is that in SOV languages which have a strict verb-final constraint, the position of greatest focus is immediately *preverbal*. Such languages allow rightward movement of constituents for focusing effect up to — but not beyond — the finite verb. This pattern has been described for rigid SOV languages such as Turkish (Erguvanı 1984), Kartvelian (Harris 1981), Japanese (Kuno 1978; cited in Kim 1988), and Korean (Kim 1985). Kim (1988) generalizes these observations as the 'linear order focus hypothesis', given in (1):

(1)   *Linear Order Focus Hypothesis*
      If L is a rigid verb-final language in its basic word order, the rhematic focus of a sentence of L is most likely to be in the position immediately preceding the finite verb. (Kim 1988:150)

Later in the same article, Kim neutralizes the distinction between rigid and non-rigid verb-final languages, provided that they are 'harmoniously head-final' languages of Greenberg's (1966) Type XXIII:

(2)   If a language has a harmonious head-final property, the information flow principle will not apply beyond the verbal head of the sentence. (Kim 1988:162)

According to this view, then, the primary focus position in SOV languages is immediately preverbal, while postverbal focus is predicted never to occur.

The alternative view holds that the tendency towards *sentence-final focus* noted at the outset is a language universal, implicitly present in languages of all word order types. As such, it seeks to manifest itself in SOV languages, resulting in violations of strict verb-finality. Hetzron (1975) argues at length for the universality of sentence-final focus, citing evidence of focused ele-

ments occurring *after* the finite verb in otherwise strict SOV languages such as Amharic and Somali. Clearly, these two views make conflicting predictions regarding the preferred focus position in languages of the SOV type: is it immediately preverbal, or is it final?

1.2 *Methodological considerations*

At this point, a methodological objection might be raised. Languages employ a variety of focus types, some of which could conceivably occupy different positions within the sentence; precisely what is intended here by the use of the term 'focus'? It is noteworthy that the claims of Hetzron and Kim are based on different focus types. Hetzron's observations derive from a cross-linguistic survey of 'presentational' focus constructions[1] — that is, constructions used to 'present' referents that are new or otherwise informationally salient, and about which the speaker typically intends to comment in the subsequent discourse. The examples discussed by Hetzron include locative inversions (e.g. 'In the house (is) a man') and clefts (e.g. 'What I saw was a pink elephant'). Kim's analysis, on the other hand, concentrates on subject WH- question words and the responses they elicit. Given these differences, it is legitimate to ask whether focus position varies depending on focus type — if, for example, it could be demonstrated that WH- focus were preverbal, and presentational focus final, then the two conflicting views might be reconciled.

A second methodological concern involves the kinds of data considered. The data in both Hetzron's and Kim's studies are isolated sentences (or simple question-response sequences) drawn from secondary sources or constructed by the authors or their informants; no actual discourse or contexts of use are analyzed. Indeed, one of the justifications advanced by Kim for choosing WH-words is that they are *inherently* focused; as a consequence, "the researcher [is] relatively free from the necessity of confirming whether or not a given subject NP in the preverbal position is indeed serving as the focus of the sentence in a specific discourse context" (1988:151). However, the results of our investigation reveal that discourse variables influence the positioning of WH- words (see Sections 2.2, 3.2), and that it is not possible to identify certain types of presentational focus without taking the immediate discourse context into account (Section 3.1). The fact that neither Hetzron nor Kim analyzed actually occurring discourse, therefore, potentially skews their investigative results. Given considerations such as these, it is obvious that no definitive answer yet

exists to the question of what constitutes focus position in SOV languages — further systematic, discourse-based research is required.

1.3 *Aims and methods of the present investigation*

The present investigation analyzes the relationship between focus and linear word order in two SOV languages, Sinhala (Indo-Aryan) and Tamil (Dravidian).[2] Both are harmoniously head-final type XXIII languages: each has postpositions rather than prepositions, and genitives and adjectives precede the nouns they modify. According to the predictions advanced in (1) and (2) above, focus position (at least for WH- words) in both languages should be preverbal, and final focus should never occur. Indeed, Tamil is one of the languages cited by Kim in support of the linear order focus hypothesis, and therefore any Tamil evidence which runs counter to these predictions would constitute a direct challenge to Kim's claim. In contrast, the universalist hypothesis predicts that Sinhala and Tamil should make use of final focus, at least in the presentation of informationally salient referents.

The methodology employed in this study involves the analysis of both presentational focus and WH- words in a corpus of oral and written narratives. The Sinhala corpus is made up of four oral narratives and nine short written narratives, for a total of 1154 finite clauses. The Tamil corpus consists of six oral narratives and 11 short written narratives, for a total of 1018 finite clauses.

The oral narratives were tape-recorded by the authors in Sri Lanka and in Tamil Nadu, South India, several years ago. The Sinhala oral narratives were related informally by Buddhist monks; all are folk tales involving Buddhist themes. The Tamil oral narratives were told by a variety of narrators, and include real-life accounts as well as folk tales and Hindu myths.

The written narrative texts include one children's story from each language, as well as narrative selections from pedagogical texts for adult second-language learners. Both children's stories were intended for native-speaker students at a first or second grade level. The second-language texts are the product of a single, native-speaker author in each language; the Sinhala selections are folk tales, and the Tamil selections are classical literary narratives retold in simplified, informal prose.[3] The written data represent a relatively straightforward, colloquial style, which because of its pedagogical orientation can be considered to reflect prescriptive norms of modern usage.

We analyzed the position of focus words in these data both quantitatively and qualitatively, with a view toward their functional correlates in discourse. The following section (Section 2) describes the results of our analysis for presentational and WH- focus in Sinhala. Section 3 then compares these findings with the results of a similar analysis for Tamil.

2. Sinhala

2.1 *Presentational focus in Sinhala*

Presentational focus is prototypically associated with a cluster of features, including the status of the referent as 'new' information, its resumption as 'given' or thematic information in the clauses which immediately follow, and its salience or importance relative to the ongoing discourse. Of these, newness of mention is the most consistent correlate of presentational focus, despite the fact that a very small number of 'presented' referents in our data are in fact reintroduced after having been mentioned previously.[4] We therefore analyzed new mentions with respect to their position within the sentence: do they favor immediate preverbal position, and if so, to what extent?

The results of our analysis show that new mentions in Sinhala are by no means exclusively restricted to preverbal position, but rather are distributed across initial (first and second),[5] preverbal, and final positions. However,
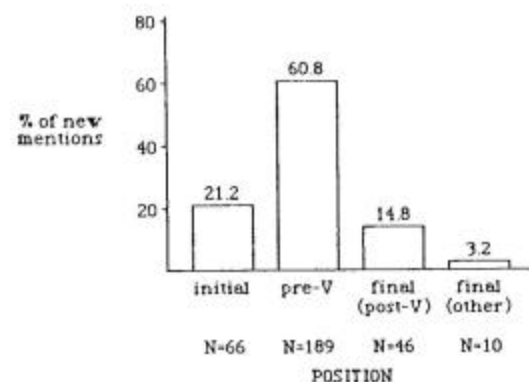

Figure 1. Distributions of new mentions in Sinhala (all grammatical roles)

about three times more new mentions (60.8%) appear in preverbal position than they do in initial (21.2%) or final positions (18.0%).[6] This is depicted graphically in Figure 1.

The high incidence of new mentions in preverbal position in Sinhala would appear to lend strong support to the linear order focus hypothesis. At the same time, one cannot rule out the possibility that new mentions occur there for other reasons, e.g. because they are grammatical objects, or are otherwise syntactically conditioned to appear directly before the finite verb. This is a hypothesis which can easily be tested. A closer examination of the data reveals that although new Os are found virtually exclusively in preverbal position, new Ss and obliques statistically favor this position as well.[7] The location of new mentions according to their grammatical role is summarized in Table 1.[8]

The fact that the overwhelming majority (94.5%) of new Os in the corpus occur in preverbal position, which is the unmarked position for direct objects in Sinhala, is not surprising, and should not be taken as evidence for preverbal focus. That is, the figure of 60.8% in Figure 1 above implies a higher degree of preverbal focus than can justifiably be claimed to exist. In order to obtain a more representative statistical measure, we calculated the percentage of non-Os (i.e. S and oblique arguments) in each of the three major positions. This calculation produces a somewhat more balanced distribution: 31.6% for initial (first and second) positions, 44.5% for preverbal position, and 19.1% and 4.8% for final positions — that is, nearly half of all non-O new mentions appear immediately preverbally, while the remainder are distributed across the other sentential positions. The adjusted distribution of new mentions in Sinhala is shown in Figure 2. These results suggest that there is a tendency — albeit not an overwhelmingly strong one — for new referents to be presented in immediate preverbal position, irrespective of their grammatical role.

Table 1. *Position of Sinhala new mentions by grammatical role*

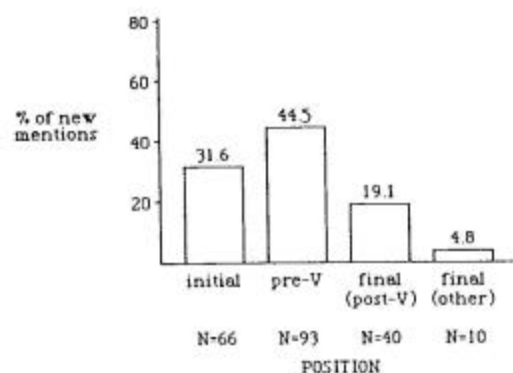|	S	O	Obl	Total
initial	31.6%	0%	34.3%	N= 66
preverbal	**44.7%**	**94.5%**	**49.3%**	N=189
final	23.7%	5.5%	16.4%	N= 46
Total:	100%	100%	100%	N=301
	(N=106)	(N=102)	(N=93)	

Figure 2. *Distribution of non-object new mentions in Sinhala*

An example of preverbal presentation from the narrative data is reproduced in (3). [9]

(3)   Sinhala (written folk tale)
       *oya gaňg-en   e-goḍa   eka pœtta-k-a   **loku kœlœœœwa-k***
       that river-INSTR that-bank one side-INDEF-LOC **large forest-INDEF**
       tibunaa.
       exist-PAST
       'On the far bank of that river was **a large forest**.' (lit. '**a large forest** was')

In this example, the locative adverbial 'on the far bank of that river' is in initial position, followed by the subject 'a large forest'. If the inverse order had been employed, the focus would have been on the location of the forest; as the sentence actually appears, the existence of the forest is accorded greater salience. This sentence occurs near the beginning of the narrative as part of a sequence of sentences which contain the verb *tibunaa* 'exist', and which describe the physical setting (there was a river, a forest, a village, etc.). In this sequence, all of the new referents occur in preverbal position.

In our attempts to establish preverbal focus for Sinhala, we have thus far said little about what is in fact a rather remarkable discovery in our analysis — the occurrence of presented referents in postverbal position. Although these account for only 14.8% (N=46) of new referents, the fact that they occur at all is an explicit violation of principle (2) above, which states that new informa-

tion will not be introduced beyond the verbal head of the sentence. These instances are distributed relatively evenly across all of the narratives in the sample, both spoken and written. What is their function, and how does it differ from that of preverbal focus?

Postverbal new mentions are strongly associated with the presentational function. Not only are they new information, but they tend overwhelmingly to be 'important' referents which are carried over as themes in the following discourse. Example (4) below is from the same story as example (3), and also introduces a forest.[10] This time, however, the forest is not simply one of a set of background features of the general narrative orientation, but rather functions as the location in which the dramatic events of the story begin. Corresponding to its functional importance, the subject NP 'a forest' appears in the more salient postverbal position:

(4)   Sinhala (written folk tale)
      gaňg-en   me-goḍa-t   tibunaa **kælæœwa-k.**
      river-INSTR this-bank-also be-PAST **forest-INDEF**
      'On this bank of the river also was **a forest.**'

By the use of a marked word order, the author effectively draws the reader's attention to this particular forest, which then becomes the theme of the following sentence:

(5)   oya kælæœ-we hiṭiyaa   **nariy-ek.**
      that forest-LOC live-PAST **jackal-INDEF**
      'In that forest lived **a jackal.**'[11]

Example (5) contains yet another instance of postposed presentational focus. The jackal introduced here is a main character in the narrative, and thus is accorded positional salience. Postposing is employed in introducing the only other main character — a crocodile — as well.[12]

There appears to be a functional distinction between preverbal and postverbal presentation in Sinhala. Postverbal presentation is characteristic of the introduction of referents or features of the setting which are central to the narrative plot. Of the 13 narratives in the corpus, the Sinhala children's story illustrates this pattern most clearly. The story tells of a young boy who learns to read and write, but is sad because he has no stories to read. His writing stylus and palm-leaf notebook, seeing his condition, go off on a late-night journey, looking for a story to bring back to the boy. They meet an assortment of

creatures along the way, but are unsuccessful until, at the world's end, they encounter Sarasvati, the goddess of learning, who summons for them all the pundits of the world, each of whom writes a story for the boy. The boy reads the stories and becomes a wise pundit himself. In this narrative, only four nominal referents are presented postverbally: the boy, the stylus, the palm-leaf book, and Sarasvati. These are clearly the central participants of the narrative. Secondary participants, such as the boy's mother and father, the boy's teacher, the various creatures encountered by the stylus and the palm-leaf book, and the pundits, are introduced in immediate preverbal position.

Given the existence of a postverbal focusing strategy of this type, is it reasonable to maintain that the primary focus position in Sinhala is preverbal? In terms of sheer numbers, preverbal focus predominates: more than three times the number of new referents are introduced preverbally than postverbally. Yet of the two strategies, postverbal presentation appears to be more specialized in its focus function.

In order to quantify this observation, we calculated the percentage of overall mentions in each of the three major positions which are new. If a particular sentential position were found to contain mostly new mentions, we could then conclude that the position was specialized in the presentational function, the assumption being that the position containing the highest concentration of new mentions would be salient in this function to users of the language.

The results of this analysis reveal a profile very different from that shown by simple distribution (cf. Figure 2). Athough initial position contains a significant number of new mentions (N=66), this number is completely overshadowed by the number of mentions which are old or accessible information (N=1245),[13] resulting in a density of new mentions of only 5.3% However, this is not surprising when we consider that initial position in subject-initial languages is conventionally reserved for thematic and/or topical information (Herring 1990). Preverbal position, on the other hand, contains relatively fewer old mentions, although the ratio of old and accessible to new is still nearly five-to-one (percentage of new mentions = 21.4%; see Figure 3 below). If initial and preverbal positions are compared, preverbal position is significantly more specialized for the presentational function than is initial position. However the issue of real importance — that is, the comparison between preverbal and final positions — remains to be considered.

Final position in Sinhala also contains elements other than new mentions. These are of two unrelated types. First, there is the so-called 'afterthought' construction, which bears a formal resemblance to the presentational construction in that both allow the postposing of nominals in all case roles, without requiring any other modifications to the morphosyntax of the sentence. The two are functional complements, however, in that information appearing after the finite verb in an 'afterthought' construction is typically contextually given, or, less commonly, new (but not crucial) information added on to an otherwise functionally complete proposition.[14] Afterthought postposings are characteristic of both written and spoken narratives, as illustrated in examples (6)-(7).

(6)   Sinhala (written folk tale)

ee særee gemba ara kanda uḍa-ṭa   gihin   kææ gaha-nḍa
then   frog   that hill   top-DAT   go-PPLE shout-INF
paṭan gatta aliya-ṭa-t   hoňḍaṭa-ma watura tibaha-yi.
start-PAST   elephant-DAT-also well-EMP   water   thirsty-PRED.
gemba-ge sadde æhi-laa   aliyaa   bohoma amaaruw-en
frog-GEN   sound hear-PPLE elephant very   difficulty-INSTR
nægga   ara kanda-ṭa.
climb-PAST that hill-DAT

'Then the frog, having gone to the top of that hill, began to croak. The elephant also was very thirsty. Having heard the sound of the frog, the elephant climbed up with great difficulty, **that hill**.'

(7)   Sinhala (oral folk tale)

eyaa keree   mokak-da? eewaa tæn   tæn   wala
he   do-PAST-EMP what-Q   those   place-place-LOC
*walee dæmma, **parissan karanna, ee kaale   tibba***
bury-PAST   **protect-INF**   **that time   place-PAST-ADJ**
widiya-ṭa.
way-DAT

'What did he do? He buried them in various places, **for safe-keeping, the way they kept things at that time**.'

In (6), the postposed nominal 'that hill' is contextually given, having been mentioned two sentences earlier. In (7), the postposed elements — an infinitive phrase and a complex nominal — have not been previously mentioned, but neither do they contain information that is surprising or of particular significance to the narrative. They are added on as 'afterthoughts', i.e. for the sake of

specifying additional detail to an already essentially complete utterance. In sentences with postposed presentation, in contrast, the postposed element is a vital part of the sentence's meaning; it could be moved preverbally, but the sentence would make no sense without it, given the context in which it occurs.

In distinguishing between afterthought postposing and postverbal presentation in the narrative data, we made use of three criteria. First, the entation in the narrative data, we made use of three criteria. First, the postposed NP in the presentational construction is usually indefinite, while that in the afterthought construction is definite. Second, the presentational construction has an intonational peak coinciding with the postverbal constituent, whereas the corresponding intonational peak in the afterthought construction coincides with the verb itself, the postverbal material being 'backgrounded'. Finally, afterthought nominals, being non-central, are rarely if ever elaborated upon in the clauses which follow.[15] In short, although the two strategies make use of the same word order, they are sufficiently distinct in other respects that the possibility of confusion between them rarely arises.

The remaining type of postposing in Sinhala bears a functional relationship to presentational focus, yet is formally distinct. This is the so-called 'emphatic transformation' (Gair 1970), in which a special emphatic form of the finite verb is used. The 'focus' in such constructions is prototypically contrastive, as in (8) below:

(8)   Sinhala (written folk tale)

*mama kiya-nna ya-nne **lankaaw-e**   **kataawa-k   nemey,***
I   tell-INF   go-EMP **Sri Lankan-GEN story-INDEF NEG**
indiyaaw-e kataawa-k.
Indian-GEN story-INDEF

'It is **not a Sri Lankan story** that I am going to tell, but **an Indian story**.'

At the same time, like the English *it*-cleft construction often used to translate it, the Sinhala emphatic construction has a broader functional range than the expression of simple contrast, including non-contrastive assertion and stylistic emphasis, as illustrated in example (9):

(9)   Sinhala (oral folk tale)

balan iňda-laa mahawsada   panḍitayoo dæn naagana hema
wait   be-PPLE Mahawshadha Pandit   now bathe-PPLE etc.

aawa.   *æwilla*   **dæn baḍaginn-e-y**   *in-ne.*

come-PAST   come-PPLE   now   **hunger-LOC-FOC**   be-PRES-EMPH

'Having waited, Mahawshadha now bathed himself and came. Having come, he is now **very hungry**.' (lit. 'it is **in hunger** that he now is')

In this example, no contrast is implied; the focus construction simply emphasizes the state of being hungry.

The emphatic form of the verb, in addition to appearing in emphatic focus constructions, is also grammatically required in negated sentences and with WH- words.[16] As a consequence, the emphatic construction is extremely frequent in modern colloquial Sinhala: 10.5% of all finite clauses in the narrative corpus have the verb in the emphatic form, and the frequency is even higher in conversation, which makes greater use of contrastive assertions, negation, and questions.

The focused nominal in an emphatic construction may appear either before or after the verb. In the Sinhala narrative corpus, 28.8% of emphatics involve an initial or immediately preverbal focused element, while the remaining 71.2% are postposed, as in (8) above.[17] Thus the emphatic construction alone does not appear to require a particular word order; by virtue of its characteristic verbal morphology, the construction is always 'focused' regardless of where in the sentence the focused element occurs.

Taking into consideration postverbal mentions of all three types — presentational, afterthought, and emphatic — we arrive at a percentage of 44.7% of the total that are new. This represents a degree of specialization that is more than twice that for preverbal position. The averaged percentage is misleading, however, in that the three postposed types exhibit widely divergent degrees of specialization. The percentage of 'afterthoughts' that are new information is extremely low: 7.7%, a percentage which reflects the *negative* correlation between afterthoughts and new mentions. Postposed emphatics, on the other hand, are new one-half of the time. While this figure would seem to suggest that information status is irrelevant to the placement of focused nominals in the emphatic construction, such is not the case. At 50.0% new, postposed emphatics are nearly four times more likely to be first mentions than are preverbal emphatics, which are new only 13.3% of the time. This argues further for an association between postverbal position and the presentational function, even if newness alone cannot account for all of the instances of postposed mentions in emphatic constructions.[18]
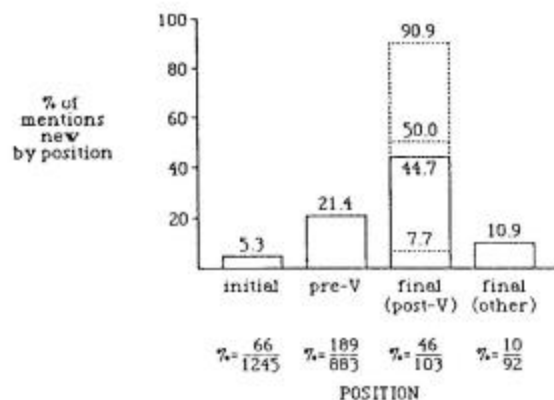

Figure 3. Specialization of positions for new mentions in Sinhala

Excluding afterthoughts and emphatics from the count, we are left with a percentage of 90.9% of simple postposings which introduce new referents — a very high percentage indeed! The relative degrees of specialization of each sentential position for new mentions are summarized in Figure 3.

Figure 3 shows that in terms of the specialization of each position for new mentions, sentence-final position, rather than immediately preverbal position, predominates. The Sinhala situation thus supports the predictions of Hetzron regarding presentational focus. In so doing, it explicitly contradicts one of the predictions made by Kim — namely, that in head-final SOV languages, focused information will not appear after the finite verb.

2.2 WH- focus in Sinhala

Thus far we have shown that postverbal position is more specialized for presentational focus than is preverbal position. The crosslinguistic study on which the 'linear order focus hypothesis' is based did not consider presentational focus, however, but rather WH- words. It is at least theoretically possible that WH- words in Sinhala pattern according to the principles given in (1) and (2) above — that is, that they appear preferentially in immediate preverbal position, and do not appear postverbally. In this section, the use of WH- words in the Sinhala narrative corpus is considered.

Table 2. Position of Sinhala WH-words by grammatical role

	S	O	Obl	Adv	Total:
initial	50.0%	17.6%	100%	36.4%	N=14
preverbal	25.0%	**47.1%**	0%	27.3%	N=12
final	25.0%	35.3%	0%	36.4%	N=11
Total:	100%	100%	100%	100%	N=37
	(N=4)	(N=17)	(N=5)	(N=11)	

WH- questions are relatively infrequent in narrative. Out of 1154 finite clauses, only 48 (4.2%) are WH- questions; of these, three are WH- words appearing alone as interjections, and eight contain WH- words in final position in nominal and adjectival predicate constructions, which we exclude from further consideration since they do not contain a finite verb. The remaining WH- words are of four grammatical types: subject, object, oblique (morphologically marked for an oblique case, e.g. 'with whom?', 'from where?'), and Adverb ('where', 'when', 'how'). The distribution of WH- words in Sinhala is summarized in Table 2.

It is immediately obvious upon examining Table 2 that WH- words in Sinhala do not favor preverbal position. Even object WH- words show only a slight statistical preference for preverbal position (47.1%), despite this being the canonical location for objects (compare this percentage with the percentage of presented object NPs in preverbal position (94.5%) in Figure 1). Subjects and adverbials, on the other hand, are distributed relatively evenly across the three positions,[19] while obliques cluster in initial position. Ironically, the highest overall concentration of WH- words is in initial position — the position in the sentence conventionally considered to be *least* favored for focus functions.[20] What accounts for this distribution?

The explanation, we suggest, is partially discourse-based. More than one-third (35.1%) of the WH- questions in the Sinhala narratives are rhetorical in nature. These may be of the 'classical' type, which implies a universally quantified negative response ('Where would a bird learn stories?' = 'A bird wouldn't learn stories anywhere'), or alternatively, of the 'thematicizing' type,[21] which leads into the next narrative event ('What did he do then? He buried all his possessions'). WH- words in classical rhetorical questions

(CRQs) appear exclusively in initial or preverbal positions, in contrast with genuine information-seeking questions, which may place the WH- word in final position. This contrast is illustrated in examples (10)-(11):

(10) Sinhala (written folk tale)   [CRQ]
    dæn pænikirilli-ṭa   harima duka-yi. ara aliya-t   ekka
    now hummingbird-DAT very   sad-PRED. that elephant-also with
    *harima taraha-y.   eet **moka-da** kara-nne?*
    very   angry-PRED. still **what-Q**   do-PRES-EMP
    'Now the hummingbird is very sad. She is very angry with that elephant. But **what** can she do?' (i.e. 'She can't do anything')

(11) Sinhala (written folk tale)   [info Q]
    awasarayi deeway-an wahansa, ara maagal-ee wele-nḍa
    permission lord-ACC   HON-VOC that mat-LOC   dry-INF
    *daa-la   tiye-nne   **monawaa-da?***
    put-PPLE be-PRES-EMP **what-Q**
    'Excuse me, your highness, **what** has been set to dry on that mat?'

In the information-seeking question in (11), posed by a character in a folk tale, the WH- word follows the finite verb, which is in the emphatic form, as is grammatically required. In (10), however, the question is rhetorical — in the given context, we understand it to mean 'she can't do anything' — and the WH- word appears *before* the emphatic form of the finite verb. Classical rhetorical questions comprise only 14.3% of WH- words in initial position, however, and thus make only a minor contribution to the higher incidence of WH- words in that location.

Thematicizing rhetorical questions (TRQs), on the other hand, typically situate the WH- word in postverbal position (e.g. 'Then he did *what*?'; see

Table 3. Position of Sahala WH-words by rhetorical type

	info Q	CRQ	TRQ	Total
initial	**50.0%**	**40.0%**	0%	N=14
preverbal	29.2%	**60.0%**	25.0%	N=12
final	20.8%	0%	**75.0%**	N=11
Total:	100%	100%	100%	N=37
	(N=24)	(N=5)	(N=8)	

also example (7)), where they are closest to the narrative 'answer' which follows. Rhetorical questions of this type account for more than half (54.5%) of the postverbal WH- words in the Sinhala data, a concentration which significantly influences the overall incidence of WH- words in that position. Table 3 shows the positioning of WH- questions by rhetorical type. The most obvious conclusion to be drawn from these data is that the discourse functions of WH- words can and do affect their positioning in the sentence, and hence must be taken into account. True information-seeking questions statistically favor the placement of WH- words in initial position, CRQ's favor initial or preverbal[22] (but not final) position, and TRQ's favor final (but not initial) position. In all, this evidence offers no support whatsoever for the prediction that WH- words in Sinhala will occur preferentially in preverbal position — indeed, preverbal focus is *weaker* for WH- words overall than it is for new referent presentation.

2.3 *Focus and the rigid verb-final constraint*

The Sinhala findings presented thus far argue against the generalization that the 'linear order focus hypothesis' applies to all head-final languages. Sinhala is an otherwise well-behaved head-final language, yet it has final focus for presentation and thematicizing rhetorical questions, and initial focus for other question types. In his original formulation of the hypothesis (see (1) above), however, Kim appears to suggest that the restriction against focus postposing, and the resultant specialization of preverbal position for focus, might be a characteristic of rigidly verb-final languages only. If it could be demonstrated that Sinhala were a non-rigidly verb-final language, its lack of conformity to the predicted pattern might be explained; at the very least, it would not violate the prediction in (1).

In order to arrive at a measure of verb-finality for Sinhala, we sorted all of the finite clauses in the corpus into the categories 'verb-final' ((X)V), 'verb non-final' (VX), and 'verbless'.[23] The results are summarized in Table 4.

Table 4 shows that three-quarters of all verbal clauses are verb-final, which is to say that one-quarter of clauses containing a finite verb do *not* have the verb in final position.[24] Although verb-final utterances are statistically preferred overall, Sinhala cannot be considered a rigidly verb-final language. This fact may account for the possibility of focusing nominals by moving them to the end of the sentence, beyond the finite verb.

3. Tamil

Tamil, like Sinhala, is a head-final type XXIII language. Unlike Sinhala, however, Tamil is rigidly verb-final. In the 11 written Tamil narratives analyzed, each and every verbal clause has the finite verb in final position; that is, there is 100% verb-finality.[25] Oral narratives, on the other hand, allow some postposing, although the percentage of verb-final verbal clauses is still extremely high (92.9%). The combined figures regarding verb-finality in Tamil are shown in Table 5. Given its high degree of verb-finality, Tamil is thus an ideal language in which to test the linear order focus hypothesis.

Kim (1988) includes Tamil in his list of languages which "show a noticeable disposition to put their WH- words in immediately preverbal position" (154). At the same time, he admits to gathering his data from a single source, Ganesan (1975), in which "there are not enough data to test the hypothesis in terms of the position of subject WH- words in ... Tamil" (153). He does however cite Ganesan's remarks concerning general focus and word order in Tamil:

> In his comparative study of Tamil and Hindi, Ganesan (1975) remarks that the unmarked word order of Tamil sentences is typically SOV, "however the order may be OSV if the subject is to be put in focus (p.456)," as in *putakam na:n etutte:n* [sic] 'I took the book' [lit. 'book I took' - SCH]. He also notes that in bitransitive sentences the dative NP usually precedes the accusative NP but this order may be reversed "if the dative object is in focus

Table 4. Verb-finality in Sinhala

	(X)V	VX	verbless
% all finite clauses	66.5%	21.4%	12.0%
% verbal clauses only	75.6%	24.4%	—

Table 5. Verb-finality in Tamil

	(X)V	VX	verbless
% all finite clauses	83.0%	4.0%	13.0%
% verbal clauses only	95.4%	4.6%	—

(p. 457)." Tamil adverbs seem to behave in a similar fashion. According to Ganesan, temporal as well as spatial adverbs occur in three optional positions: sentence-initial, medial, and immediately preverbal, but those adverbs occurring in the immediately preverbal position always receive a special focus interpretation (p. 459). (Kim 1988:153-4).

In the following sections, we evaluate how well the claim that Tamil has preverbal focus fits the distribution of presentational and WH- focus in actual narrative usage.

3.1 Presentational focus in Tamil

In our analysis of presentational focus in Tamil, the same procedure was followed as for Sinhala. That is, new mentions were identified and classified according to position in the sentence, and a graph showing the distribution of new mentions by position was plotted. The new mentions were then further broken down by grammatical role, and new Os in preverbal position excluded. Last, the relative degree of specialization of each position for new mentions was calculated.

The Sinhala data did not display significant differences in focus strategies between the written and the oral texts, since both were of a colloquial variety. However, preferred focus strategies differ in the written and spoken Tamil texts. We therefore present the findings for each modality separately in the discussion that follows.
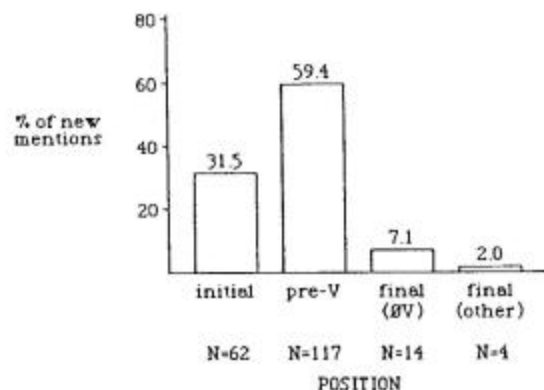

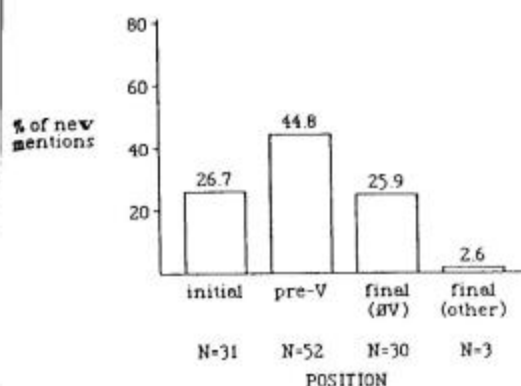*Figure 4. Distribution of new mentions in written Tamil (all grammatical roles)*

Figure 5. Distribution of new mentions in spoken Tamil (all grammatical roles)

The overall distribution of new mentions in Tamil is shown in Figures 4 and 5. As these figures show, preverbal new mentions in both written and spoken Tamil predominate numerically over new mentions in other sentential positions. The written texts have relatively more new mentions in preverbal position than the oral texts, however, and new mentions are less likely in the written texts than in the oral to appear in sentence-final position, i.e. in verbless (ØV) constructions (see below).[26] Tables 6 and 7 give the breakdown of new mentions by grammatical role.[27] New Os overwhelmingly favor preverbal position in both the written and spoken data, as was found in Sinhala. If we exclude Os, however, the two varieties display rather different patterns. New mentions in written Tamil show a preference for initial position when subjects, and preverbal position when obliques, suggesting a distributional pattern based on grammatical role (S-initial, O-preverbal), with a preference for preverbal Obliques (cf. Ganesan's observations on 'dative objects'). Spoken Tamil, in contrast, prefers to introduce new Ss in *final* position, and divides Obliques equally between initial and preverbal positions — a system in which neither grammatical conditioning nor preverbal focus plays an obvious role. The adjusted distribution of subject and oblique new mentions is represented graphically in Figure 6 (for written Tamil) and Figure 7 (for spoken Tamil).

Comparing Figure 6 with Figure 2, we see that written Tamil shows no greater degree of preverbal focus than does Sinhala. And spoken Tamil manifests no preference for preverbal focus at all — new mentions are evenly

Table 6. Position of written Tamil new mentions by grammatical role

	S	O	Obl	Total:
initial	44.8%	4.4%	34.0%	N= 62
preverbal	40.6%	**95.6%**	**66.0%**	N=115
final	14.6%	—	—	N= 14
Total:	100%	100%	100%	N=191
	(N=96)	(N=45)	(N=50)	

Table 7. Position of spoken Tamil new mentions by grammatical role

	S	O	Obl	Total:
initial	26.9%	4.8%	50.0%	N=31
preverbal	28.4%	**95.2%**	50.0%	N=51
final	**44.8%**	—	—	N=30
Total:	100%	100%	100%	N=112
	(N=67)	(N=21)	(N=24)	

distributed across the three major positions. These results are surprising, given that Tamil is more rigidly verb-final, and hence is predicted to have stronger preverbal focus according to the hypothesis in (1).

This is not to say that preverbal focus is non-existent in Tamil. On the contrary, the fact that many new Ss and obliques appear preverbally can be taken as evidence that the position has focus value.[28] Preverbal focus is most apparent in written Tamil; as for example in (12):

(12)   Tamil (written historical narrative)
     anta nāṭṭ-il   **oru aracan** *iru-nt-āṇ*
     that   country-LOC   **one king**   be-PAST-3SG:MASC
     'In that country, (there) was **a king**.' (lit. '**a king** was')

As in the Sinhala example in (3), the referent is presented here as the subject of an intransitive predicate, and is situated between an initial locative adverbial and the finite verb, in this case, the verb *iru* 'to be'.

The Tamil children's story in the corpus makes systematic use of preverbal presentation. In this story, a puppy becomes separated from its master, a young boy, during the drive back from a mountain vacation resort. The puppy, although frightened, is determined to find its way back to the boy, and
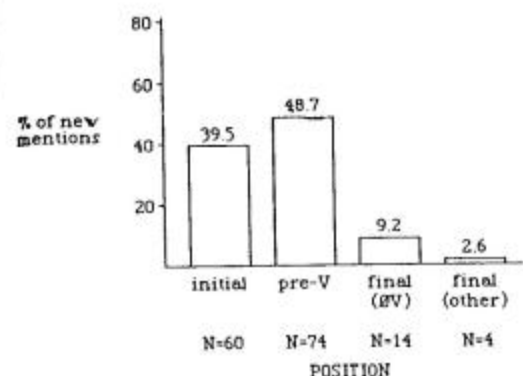

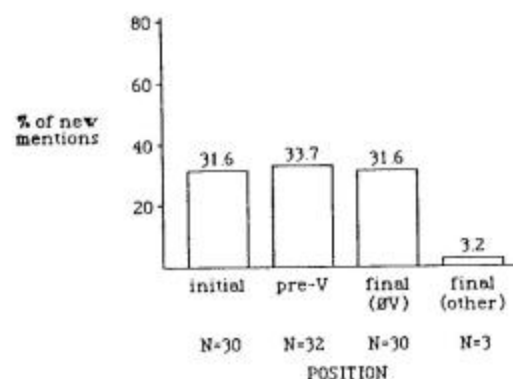*Figure 6. Distribution of non-object new mentions in written Tamil*

Figure 7. Distribution of non-object new mentions in spoken Tamil

sets out along the road. In the chapter analyzed, it encounters a series of larger animals — an adult dog, a mother bear and her two cubs, a lame horse — who befriend the puppy and offer it food, shelter, and advice along the way. Each of these animals is introduced preverbally, as the subject of an intransitive verb following a locative adverbial, as in example (13):

(13)   Tamil (written children's story)
     Pātai ōratt-il   **oru kutirai** *mēyntu-koṇṭiru-nt-atu.*
     path   side-LOC   **one horse**   graze-CONT-PAST-3SG:N
     'By the side of the path, **a horse** was grazing.'

Interestingly, however, the most central characters in this story — the puppy and the boy — are introduced not preverbally but as subjects in initial position, as in (14), the first mention of the boy.

(14) Tamil (written children's story)
avar-utaiya makan katirvēl nāy-kkuṭṭiy-iṭam anp-uṭan
he-GEN   son   Katirvel dog-child-LOC   affection-with
paḻaki-va-nt-āṇ.
spend:time-come-PAST-3SG:MASC
'**His son Katirvel** used to spend time being affectionate with the puppy.'

This strategy is characteristic of main character presentation in the Tamil written narratives more generally, as reflected in the higher percentages of initial Ss in the first column in Table 6.

Preverbal focus is also employed to some extent in the oral narratives, especially in introducing non-central participants. Example (15) below is structurally and functionally analogous to written Tamil example (12) and Sinhala example (3):

(15) Tamil (oral folk tale)
Oru kāṭṭ-ile   vantu, inta **niṟaiya mirunkañ-kaḷ** iru-nt-atu.
one forest-LOC TOP   this **many   animal-PL**   be-PAST-3SG:N
'In a forest, this — there were **many animals**.' (lit. '**many animals were**')

Important referents, on the other hand, are introduced via a strategy which allows the new referent to appear in sentence-final position. Unlike in Sinhala, this is not accomplished by moving the referent to the right of the finite verb, since in Tamil postverbal position is generally reserved for intonationally backgrounded given or accessible referents.[29] Instead, the Tamil presen-tational construction omits the finite verb, resulting in an utter-ance type in which a nominal referent occupies absolute sentence-final posi-tion 'by default'. We refer to this phenomenon as 'verbless presentation'.[30]

As with the Sinhala postposing strategy, a verblessly presented constitu-ent in Tamil is almost invariably new information, which is treated as given in the utterances that follow. This is illustrated in example (16), which immedi-ately follows the utterance in (15) in the text of the narrative:

(16) Tamil (oral folk tale)
At-ile   **oru periya ciñkam**. Cinkatt-ai kaṇ-ṭā   ellā ellā
that-LOC **one big   lion**.   lion-ACC   see-COND all all
animals-ukkum payam.
animals-DAT   fear
'Among them, **a big lion**. Whenever (they) saw the lion, all the animals were afraid.'

The lion in this example is the villain of the story, and hence a central participant. The hero — a rabbit — is also introduced verblessly:

(17) Appuṟam mūṉāvatu nāḷ vantu **oru muyal**. Cens
then   third   day TOP   **one rabbit**. chance
va-nt-atu.
come-PAST-3SG:N
'Then on the third day, **a rabbit**. (It's) turn came.'

In contrast, the other animals — both collectively and individually — are introduced in preverbal position, as are other secondary participants in the tale. In this respect, verbless presentation in Tamil functionally parallels presenta-tional postposing in Sinhala, in that both specialize in the introduction of thematically important referents.

Verbless presentation is found in written Tamil as well, where it typically functions to set the physical or temporal scene:

(18) Tamil (written historical narrative)
Anta ūr-il   **oru tiruviḻā**. Tōli   talaivi-ṭam, "talaivi,
that town-LOC **one festival** friend heroine-LOC heroine
nāṅkaḷ tiruviḻā pār-kka pō-kiṟ-ōm. Nīy-um vā-yeṉ"
we   festival see-INF go-PRES-1PL you-and come-why
en-ṟ-āḷ
say-PAST-3SG:FEM
'In that town, **a festival**. Her friend said to the heroine, "Heroine, we're going to see the festival. Why don't you come along?"'

Verbless presentation exhibits its own unique formal characteristics. Unlike in Sinhala, the presented NP must always be in the subject, or morpho-logically unmarked, case. The actual grammatical role played by the referent in the ongoing discourse is indicated by the case marking used when it is

resumed in the following clause (e.g. accusative in (16) and genitive in (17); it may also be resumed inside of a quote, as in (18)). As for the 'missing' verb in such constructions, it need not be a simple verb of existence, as (17) attests; its interpretation relies heavily on the content of the immediately following utterance. Verbless presentation in Tamil is thus a grammatical strategy which can properly be understood *only* through the analysis of connected discourse.

The verbless construction, although accounting for only 13.7% of new mentions in the Tamil corpus overall, is highly specialized for new referent introduction. Out of 30 instances in the oral narratives, fully 88.2% contain new mentions or mentions reintroduced after a long hiatus. This figure subsumes two verbless construction types. The first, participant presentation as illustrated in (16) - (18), introduces referents that are new 89.3% of the time — more than five times the degree of specialization of preverbal position (17.5%). The second type is the 'dative subject' possessor construction, which occurs both with and without a finite verb. An example of the verbless variant is found in the last sentence in (16) — the expression *ellā animals-ukkum payam* literally translates as 'to all the animals, fear'. This clause introduces the notion of 'fear' into the discourse, although 'fear' is of course not a narrative participant in the same sense as is 'the lion' or 'the rabbit'. Dative subject constructions are primarily used in narrative to describe the existence (or coming into existence) of a physical or emotional state, and hence are semantically conditioned. At the same time, such constructions, *when verbless*, tend strongly to present states that are being mentioned for the first time: 83.3% of verbless dative subject constructions in the oral texts describe 'new' states of fear, anger, hunger, etc., as opposed to only 40.0% of those in which a finite verb appears. Although the dative subject construction and verbless presentation are grammatically distinct (i.e. by virtue of the presence or absence of a possessor NP in the dative case) and only partially congruent in function, it is nevertheless significant that the absence of a finite verb correlates with the presentational function in both.

The same tendencies are present in written Tamil, although to a weaker extent. Final verbless constructions, at 60.9% new, are considerably more specialized for new information than any other sentence position. Presentational-type constructions contain 64.7% new mentions, and dative-subject constructions contain new mentions exactly half the time. Interestingly, the majority (78%) of verbless constructions in the written data are not presentational but rather of the dative subject type, while the inverse is true in the oral
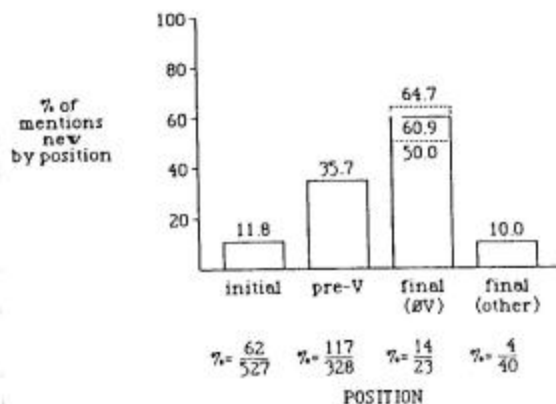

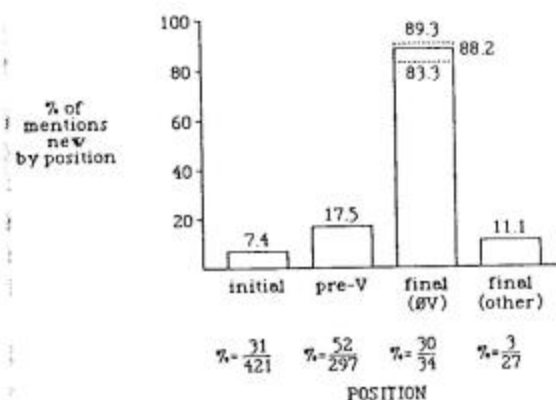Figure 8. Specialization of positions for new mentions in written Tamil

Figure 9. Specialization of positions for new mentions in spoken Tamil

data (87% presentational), reinforcing the view that verbless presentation is predominantly an oral strategy. Figures 8 and 9 represent the relative degree of specialization of each position for new mentions.

At this point, the reader is invited to consult Figure 3 again and to compare it with Figures 8 and 9. Despite differences of degree in written and spoken Tamil involving the verbless construction, the same pattern is evident in both Tamil varieties as is found in Sinhala: Final position is more specialized for new mentions than preverbal position, which in turn is more specialized than initial position. This is so, even though the final presentational strategy in each language is quite different — postposing in Sinhala, and verblessness in Tamil. The common thread appears to be that both provide a means of placing the presentationally focused constituent in sentence-final position, thereby circumventing the verb-final constraint. Thus despite its more rigid degree of verb-finality, Tamil does not support the predictions of the linear order focus hypothesis any better than does Sinhala, at least not for presentational focus.

3.2 *WH- focus in Tamil*

We come at last to the positioning of WH- words in Tamil, the area in which our data stand to confirm or disconfirm the linear order focus hypothesis most directly. From the 7.6% of finite clauses in the Tamil corpus containing WH- words (N=77), we excluded twelve instances appearing in final position in nominal predicate constructions, as was done in the previous analyses. Table 8 shows the distribution of the remaining tokens according to grammatical role and sentential position for the spoken and written data combined.[31]

The strong correlation (90.5%) evident in Table 8 between grammatical objects and preverbal position should by now be a familiar result. More

Table 8. *Position of Tamil WH-words by grammatical role*

	S	O	Obl	Adv	Interj	Total:
initial	**69.2%**	9.5%	0%	48.0%	75.0%	N=26
preverbal	30.8%	**90.5%**	100%	44.0%	0%	N=35
final	0%	0%	0%	8.0%	25.0%	N= 3
Total:	100%	100%	100%	100%	100%	N=64
	(N=13)	(N=21)	(N=1)	(N=25)	(N=4)	

surprising is the correlation (69.2%) between subject WH- words and initial position, a pattern which, taken together with preverbal object WH- words, once again suggests grammatical conditioning, rather than a pragmatically motivated distribution. The distribution of WH- adverbs is also compatible with this view: adverbs are not assigned a fixed position in Tamil grammar, nor do the WH- adverbs in our data favor a single location, but rather are roughly evenly split between initial and preverbal positions. This is clearly not the distribution we would expect if WH- words, as inherently focused elements, were pragmatically conditioned to appear preverbally.

Although the above results were obtained in a manner consistent with Kim's methodological approach — i.e. by treating all WH- words as equally 'focused', in isolation from their contexts of use — in carrying out our analysis, we observed that the position of WH- words in the Tamil data was influenced by discourse functions. More than one-half of the WH- questions in the Tamil narratives are rhetorical; these include CRQs, TRQs, and WH- interjections (whereby the speaker expresses incredulity and/or disapproval without the expectation of a literal 'answer'). The three rhetorical question types are illustrated in (19) - (21).

(19) Tamil (written children's story)   [CRQ]
Ēṉ *ippaṭi naṭuṇku-kiṟ-āy? Nay-kuṭṭi-kaḷ-ai   nāṉ*
why like:this tremble-PRES-2SG dog-child-PL-ACC I
cāppiṭu-vat-illai eṉṟu cirittu-koṇṭ-ē   coll-iṟṟ-u.
eat-FUT:NZR-NEG QUOT laugh-CONT-EMPH say-PAST-3SG:N
'"**Why** are you trembling so? I don't eat puppies," [the bear] said laughingly.' (i.e. "There's no reason for you to tremble so...")

(20) Tamil (oral folk tale)   [TRQ]
*Uṭaṉē   teṉālirāmaṉ **eṉṉa** paṇṇ-iṉ-āṉ*
immediately Tenaliraman **what** do-PAST-3SG:MASC
mīti   pāl elām nallā āṟa va-ccu, avaṉ-um avaṉ
remainder milk all well boil put-PPLE he-and   his
kuṭumpam-um uṭkār-ntu, ellātt-aiy-um kuṭi-cciru-cc-u.
family-and   sit-PPLE   all-ACC-and drink-PERF-PAST-3SG:N
'And then **what** did Tenaliraman do? Having boiled all the milk, he and his family sat down and drank it all up.'

(21) Tamil (oral folk tale)   [WH- interjection]

"Tay-ē.. nāṉ uṉṉ-uṭaiya paktaṉ-āka iru-kkiṉravaṉ.
mother-VOC I   you-GEN   devotee-ADV be-one:who:is
Ippaṭi nāḷ kaṇakk-āka, māta kaṇakk-āka, āṇṭu
like:this day measure-ADV month measure-ADV year
*kaṇakk-āka vantu-koṇṭiru-kkir̄-ēṉ-ē; eṉ-akku **eṉṉa***
measure-ADV come-CONT-PRES-1SG-TAG I-DAT **what**
nī aruḷ paṇṇa-maṭṭ-ēṉ ñ-kir̄-ēy.
you grace do-FUT:NEG-1SG say-PRES-2SG

'O goddess! I am your devotee. Day after day, month after month,
year after year I come to you like this, right? And **what**, you refuse
to grant me your grace?!'

The location of WH- words in Tamil according to rhetorical function is
summarized in Table 9.[32] Unlike Sinhala WH- words, these data display a
noticeable tendency towards preverbal focus, especially in TRQs and infor-
mation questions. This would seem to suggest that Tamil WH- words can
indeed be analyzed as favoring preverbal focus, but only when their rhetorical
functions are taken into account.

Upon closer examination, however, even this claim is problematic. Many
of the question words in preverbal position are grammatical objects, creating
an overall bias in favor of preverbal position — particularly in the case of
TRQs — which does not accurately reflect the pragmatic word order prefer-
ences of the language. That is, while it is a potentially interesting fact of usage
that the vast majority of thematicizing rhetorical questions in Tamil involve an
object WH- word (see e.g. example (20)), this distribution cannot be used to

Table 9. Position of Tamil WH-words by rhetorical type (all grammatical roles)

	info Q	CRQ	TRQ	Interj	Total:
initial	3.3%	58.3%	20.0%	60.0%	N=21
preverbal	**62.5%**	41.7%	**80.0%**	0%	N=32
final	4.2%	0%	0%	40.0%	N= 3
Total:	100%	100%	100%	100%	N=56
	(N=24)	(N=12)	(N=15)	(N=5)	

Table 10. Position of Tamil WH-adverbs by rhetorical type

	info Q	CRQ	TRQ	Total:
initial	33.3%	**71.4%**	50.0%	N=11
preverbal	**58.3%**	28.6%	25.0%	N=10
final	8.3%	0%	25.0%	N= 2
Total:	100%	100%	100%	N=23
	(N=12)	(N=7)	(N=4)	

argue for preverbal WH- focus in Tamil, since the occurrence of objects in
preverbal position could equally well be explained by grammatical condition-
ing. The occurrence of initial subjects in these data is similarly suspect; only
adverbs are neutral in their distribution, as indicated in Table 8. What, then, of
the distribution of WH- adverbs alone? This is shown in Table 10.[33]

These data show that WH- adverbs in information questions display a
preference (58.3%) for preverbal position. In contrast, CRQs (and TRQs as
well, although the number of tokens involved is not significant) prefer to
situate the WH- adverb in initial position. When the WH- questions in the data
are analyzed separately according to rhetorical function, therefore, it becomes
possible to perceive a pattern which supports, at least in part, the claims of the
linear order focus hypothesis — that is, WH- adverbs in information questions
appear to prefer preverbal position. However the claim that preverbal position
is favored for WH- words *independent* of their grammatical role or rhetorical
function is not supported by the Tamil data, nor is Kim's larger claim
extending his predictions to pragmatic 'focus' more generally.

4. Implications and conclusions

In our discussion thus far, we have been primarily concerned to draw atten-
tion to the ways in which the placement of focused elements in Sinhala and
Tamil provide counter-evidence to the claim that focus position in SOV
languages is immediately preverbal. Our findings also enable us to venture
several positive observations regarding the expression of focus in Sinhala and
Tamil, and to hypothesize about the typological correlates of focus positions
in SOV languages more generally.

In the course of our analysis, we discovered intriguing evidence in both languages of sentence-final presentational focus, true to the spirit of Hetzron's (1975) claim. Sinhala focuses presented referents by moving them to the right of the finite verb, in violation of the 'verb-final constraint' characteristic of head-final SOV languages. Tamil circumvents the verb-final constraint by omitting the finite verb, leaving the focused nominal in sentence-final position. However, the Tamil findings suggest that Hetzron's claim, inasmuch as it is formulated in terms of presentative *movement*, needs to be modified to accomodate cases where no movement is involved.[34] Future discourse-based studies of other SOV languages may well turn up evidence of as yet unidentified and potentially even more inventive sentence-final presentational strategies, especially if colloquial spoken usage is considered.[35]

The fact that such differences exist may well reflect a typological dimension. A possible explanation for the differences between Sinhala and Tamil presentational focus is that Tamil is more rigid in its word order, and in particular, in its degree of verb-finality. By allowing greater word order freedom, Sinhala makes it possible to employ different word orders for discourse-pragmatic ends, as is manifested both in its predilection for postverbal presentation, and rhetorically-controlled WH- word placement. Tamil, on the other hand, displays an overall preference for placing subjects initially and objects preverbally, although the position of adverbs is more free. (Note that in omitting the finite verb, Tamil is not violating any word *order* constraint: the resultant structure typically has the form *(Adv) S*).[36] Rigidness of verb-finality, then, may be related to the degree to which a language exercises pragmatic as opposed to grammatical control of word order, which in turn determines the placement of focused elements.

At the same time that we have identified a tendency toward final focus, our findings also provide limited support for the notion of preverbal focus. Both languages 'present' new referents preverbally more often than in any other position, and preverbal position is relatively more specialized for the presentational function than is initial position. In addition, there is some statistical evidence that Tamil favors preverbal position for WH- adverbs in information questions. The problem, however, lies in interpreting these patterns. A precise statistical criterion for preferred focus position has never been proposed — is something a 'preference' if it holds in 60% of the cases? 40%? 20%? Clearly the relative frequencies of other positions must be factored into the analysis as well, along with their degree of functional specialization. Yet what of cases where these differing measures produce conflicting results? Preverbal focus position in our data is preferred for presentation in terms of sheer numbers, but final position is more specialized in the presentational function; which is more basic?

We suggest that these two focus positions are employed in different functions and for different effect, at least in the languages analyzed here. Final position, as a less frequent and hence more marked strategy, is favored for the presentation of central, salient participants. Preverbal position, on the other hand, is used for secondary participants, as well as for a variety of other elements, only some of which are accorded a focus interpretation. This pattern is found in the oral narratives in Tamil, and in both the oral and written Sinhala data. As for WH- focus, Sinhala regularly employs postverbal position for thematicizing rhetorical questions (TRQs), a strategy which is iconic with their function of introducing narrative events in the clauses which immediately follow. Conversely, both languages prefer early sentence positions for WH-words in classical rhetorical questions (CRQs), a reflection perhaps of the non-literal (and hence less focused?) nature of questions of this type.

One obvious conclusion to be drawn from these observations is that 'focus' cannot be treated as a homogeneous category in studies of word order variation. Although our findings do not support the hypothesis that presentational focus is final, and WH- focus preverbal, we have nevertheless discovered systematic differences in the preferred placement of the two focus types in each language. Moreover, WH- words pattern differently according to their rhetorical function, suggesting that it is not enough to distinguish WH- focus from other focus types — one must identify rhetorical WH- subtypes as well. Any typological study of focus phenomena which fails to take these distinctions into account is destined to generate oversimplistic generalizations which make inappropriate predictions for some focus types. The claim that immediately preverbal position is *the* focus position in SOV languages is one such generalization.

Notes

1. The actual term used by Hetzron is 'presentative'; we employ the more widely-used term 'presentational' here.

2. Although genetically unrelated, Sinhala and Tamil are spoken in close proximity to one another, and have a long history of contact. A number of grammatical borrowings have made their way into Sinhala from Tamil, including, most notably for the purposes of the present paper, an emphatic or 'contrastive' focus construction (Gair 1986; see also fn.30).

3. The second-language texts are *Advanced Spoken Sinhala Reader*, by W.S. Karunatillake (Kelaniya University, Sri Lanka), and *Tamil Madhu: Sangam Mudal Bharati Varai*, by Kausalya Hart (University of California, Berkeley).

4. Less than 10% of all referents which appear as the focus of a presentational construction were previously introduced.

5. In order to simplify our exposition, which primarily evaluates preverbal as opposed to postverbal focus, we conflate first and second positions into a single 'initial' position. Filled second position slots are infrequent in these data; the majority of 'initial' mentions (85% in Sinhala; 81% in Tamil) appear in (absolute) first position. Moreover, the density of new mentions in first and second positions is similarly low for the corpus overall — 7% of mentions in first position in Sinhala are new, as compared with 4% of those in second position, and the Tamil figures are 9% and 10% respectively.

6. This figure combines the two 'final mention' categories. The first contains postposed mentions which appear after a finite verb. The 'other' category is reserved for sentences in which no finite verb appears, and includes instances of ellipsis as well as nominal predication.

7. The category labelled S includes subjects of both intransitive and transitive predicates. The vast majority of new Ss in the languages considered here are intransitive subjects, however, in keeping with the tendency noted by Du Bois (1987) for speakers to avoid introducing new mentions as transitive subjects. (For an analysis of new mentions by transitivity for Tamil, see Herring (1989).)

8. Excluded from the total are final mentions in the verbless 'other' category, as such instances cannot be used to argue for a particular position vis-à-vis the finite verb.

9. Abbreviations in glosses follow standard practice (see list on p. ix). In addition, the following abbreviations are used: PPLE 'past participle', NZR 'nominalizer', and 3SG:N 'third person singular neuter'.

10. There are two forests in the story, one on either bank of the river.

11. Unlike in English, the postposed nominal in constructions of this sort need not be the subject of an intransitive predicate, but rather may occur in any thematic or case role. Postposed Os and obliques are relatively less frequent than postposed Ss overall, however, a skewing which we attribute to pragmatic principles governing the preferred argument roles for the introduction of new information, rather than to any constraint on the syntax of presentational focus constructions per se.

12. Although the Sinhala examples cited in this discussion are drawn from colloquial written narratives, the tendency toward postverbal focus is equally strong — indeed, somewhat stronger — in the oral narratives.

13. We use the term 'accessible information' to refer to information which the speaker assumes the hearer can infer or otherwise access (Du Bois (1987); cf. Chafe's (1987) 'semi-active' category, and Prince's (1981) 'inferrables'). In our data, accessible mentions are those conventionally associated with a previously mentioned referent, as in e.g. the mention of a body part ('his hand') after the individual as a whole ('a man') has been introduced, mention of 'the villagers' after we have already learned of the existence of 'a village', and '(his/her/its) mother', *except* in cases where the hand, villagers, or mother are themselves thematic narrative participants.

14. By 'functionally complete', we intend 'sufficiently specified to make sense in the discourse context in which it occurs'. Sinhala and Tamil are both zero anaphora languages, and thus completeness cannot be determined by the presence or absence of syntactic arguments alone.

15. In keeping with this last observation, it follows that distinguishing between instances of presentational and afterthought postposing requires taking the discourse context into account.

16. This is the case for true information-seeking questions. In rhetorical WH- questions, a non-emphatic form of the verb is sometimes used.

17. These figures take into account only adverbs and non-WH- lexical NPs — that is, only those elements that could conceivably qualify as new mentions. If we include WH- words and clausal constituents in the count, the distribution shifts: 56.1% preverbal, and 43.9% postverbal.

18. An avenue for further research would be to analyze the functions of the postposed emphatics in the subsequent discourse. If Hetzron's claims are correct, we would predict that postposed mentions would be carried over as thematic material in the immediately following clauses, while preverbal emphatic mentions would not.

19. The occurrence of two subject WH- words in initial position, as compared with one in each of the other positions, cannot be construed as sufficient evidence of a preference for placing subject WH- words in initial position.

20. See Herring (1990) for discussion. An exception is Givón (1990), who argues that initial position is universally preferred for contrastive focus.

21. These terms are from Herring (1991).

22. Of the three instances in the preverbal category, however, two are repetitions of a formulaic expression, *(Api) monawa karanna da* 'Whatever shall we do?' (lit. 'we what to do?'), in which the order of elements is fixed.

23. The verbless category in Sinhala includes nominal/adjectival predication and ellipsis.

24. In addition to postposed new referents, afterthoughts, and emphatics, clausal constituents and quotes may also appear postverbally in Sinhala. This accounts for the higher percentage of postposings than is indicated by the number of postposed elements in Figure 3.

25. This exceptionless adherence to verb-finality reflects the normative, pedagogical orientation of the written texts in our corpus. Word order in modern Tamil short stories and novels is somewhat less strict (Herring, To appear).

26. An analogous distinction exists between Colloquial and Literary Sinhala. Presentational postposing appears in the colloquial language, but is lacking in more literary varieties (Paolillo 1992). Since the present corpus contains only colloquial written and spoken texts, however, this distinction does not surface in our analysis.

27. Excluded from the total in Table 6 are the four instances of final mentions in the 'other' category, along with two adverbial constituents which, for grammatical reasons, could not have appeared in any but preverbal position. Excluded from Table 7 are three instances of final mentions in the 'other' category, and two adverbial constituents.

28. Where a single NP appears preverbally, its position was coded as 'initial' if it was the grammatical subject, and 'preverbal' if it was the object or an oblique argument of the verb. Thus all instances of 'preverbal subjects' involve cases where an S appears immediately before the verb, and is preceded by some other element.

29. In this respect, Tamil is consistent with the claims of Kim for rigid verb-final languages. An example of a postposed given referent is the following:

   (i) Tamil (oral real-life account)

      Appuṟam   vantu,   oru   leṭar   pōṭ-ṭ-ā.
      afterwards TOP   one letter   put-PAST-3SG:FEM
      *Añke   yār-ukkum   paṭi-kka   teriy-ātu,   **anta leṭar-ai.***
      there no:one-DAT   read-INF   be:known-NEG   **that letter-ACC**
      'Afterwards, she sent a letter. No one there could read it, *the letter.*'

   Such postposings can be afterthoughts (i.e. repairs or clarifications) or conventionalized 'antitopics' with specific discourse-level functions (see Herring (to appear) for a detailed treatment of this phenomenon).

30. Tamil also has a cleft construction (on which the Sinhala emphatic construction was calqued (Gair 1986)) which allows movement of nominals to the right of the (nominalized) verb. However, there are no instances of this construction in our data.

31. Because of the smaller numbers of tokens in this analysis, we did not separate the spoken and written data. However, no obvious patterned differences between the two varieties emerged on the basis of the examples available.

32. For the purposes of this analysis, eight WH- elements expressing indefinite quantification (*eṉṉa* NP-*um* 'whatever NP'; *eppaṭiyō* 'somehow or the other', etc.) were excluded from the data.

33. The difference between the total number of WH- adverbs here and in Table 8 is a consequence of excluding indefinite quantifiers (see fn.32), two of which are adverbial.

34. It is also possible that Hetzron's claim should be modified to include only SVO and SOV languages, since it is for languages of these types that the most convincing evidence has been adduced. Hetzron's claim that sentence-final presentational focus is 'universal' must be considered suspect, since his data contain no examples from verb-initial languages (save for one Arabic example in which the verb — an optional copula — is deleted). In fact, the available evidence suggests that focus position in verb-initial languages is immediately before or after the finite verb, rather than sentence final (see sources in Herring (1990)).

35. Verbless presentation is never mentioned in grammatical descriptions of Tamil, nor is it accepted as fully grammatical by native speaker informants when examples are presented as isolated sentences.

36. See examples (16)-(18).

References

Chafe, Wallace L. 1987. "Cognitive constraints on information flow." *Cohesion and Grounding in Discourse* ed. by R. Tomlin, 21-51. Amsterdam/Philadelphia: John Benjamins.

Du Bois, John W. 1987. "The discourse basis of ergativity." *Language* 63.805-55.

Erguvanı, Eser E. 1984. *The Function of Word Order in Turkish Grammar*. Berkeley: University of California Press.

Gair, James W. 1970. *Colloquial Sinhalese Clause Structures*. The Hague: Mouton.

Gair, James W. 1986. "Sinhala focussed sentences: naturalization of a calque." *Structure, Convergence and Diglossia* ed. by Bh. Krishnamirti, C.P. Masica and A.K. Sinha, 147-67. New Delhi: Motilal Banarsidass.

Ganesan, S.N. 1975. *A Contrastive Grammar of Hindi and Tamil*. Madras: University of Madras.

Givón, T. 1990. *Syntax: A Functional-Typological Introduction*, Vol. 2. Amsterdam/ Philadelphia: John Benjamins.

Greenberg, Joseph H. 1966. "Some universals of grammar with particular attention to the order of meaningful elements." *Universals of Language* ed. by J.H. Greenberg, 73-113. Cambridge. Mass.: MIT Press.

Halliday, M.A.K. 1967. "Notes on transitivity and theme in English: Part 2." *Journal of Linguistics* 3.199-244

Harris, Alice. 1981. *Georgian Syntax*. London: Cambridge University Press.

Herring, Susan C. 1989. "Verbless presentation and the discourse basis of ergativity." *Proceedings of the Twenty-fifth Annual Chicago Linguistics Society, Parasession on Language in Context*, 123-35.

Herring, Susan C. 1990. "Information structure as a consequence of word order type." *Proceedings of the Sixteenth Annual Berkeley Linguistics Society*, 163-74.

Herring, Susan C. 1991. "The grammaticalization of rhetorical questions in Tamil." In *Approaches to Grammaticalization* Vol. 1 ed. by E.C. Traugott and B. Heine, 253-284. Amsterdam: John Benjamins.

Herring, Susan C. To appear. "Afterthoughts, antitopics and emphasis in Tamil." In *Theoretical Approaches to Word Order in South Asian Languages* ed. by M. Butt, T. King, and G. Ramchand. Stanford: Center for the Study of Language and Information.

Hetzron, Robert. 1975. "The presentative movement, or why the ideal word order is V.S.O.P." *Word Order and Word Order Change* ed. by C.N. Li, 347-388.

Kim, Alan H. 1985. "The functions of linear order in Korean syntax." *Harvard Studies in Korean Linguistics*, 154-167. Cambridge: Harvard University.

Kim, Alan H. 1988. "Preverbal focusing and type XXIII languages." *Studies in Syntactic Typology* ed. by M. Hammond, E. Moravcsik and J.R. Wirth, 147-169. Amsterdam/

Kuno, Susumo. 1978. *Danwa no Bunpo* (The Grammar of Discourse). Tokyo: Taishukan.
Leech, Geoffrey N. and Michael H. Short. 1981. *Style in Fiction: A linguistic approach to English fictional prose*. London: Longman.
Paolillo, John C. 1992. *Functional Articulation in Diglossia: A case study of grammatical and social correspondences in Sinhala*. Ph.D dissertation, Stanford University.
Prince, Ellen F. 1981. "Toward a taxonomy of given – new information." *Radical Pragmatics* ed. by P. Cole. New York: Academic Press.

Word order at the noun phrase level in Japanese

Quantifier constructions and discourse functions[*]

Alan Hyun-Oak Kim
Southern Illinois University at Carbondale

1. Introduction

In Japanese sentences, arguments have considerable freedom in taking positions without changing their grammatical relations (GFs), a phenomenon which is known as scrambling. The reordering of arguments from one syntactic position to another in a clause or sentence is not, however, completely free. In actuality it is subject to a number of restrictions. First, no elements occur after the predicating verb, so as to keep the verb stationary at the clause-final position (as per the verb-final constraint).[1] Second, when arguments are presented without explicit case-marking, as in the case of casual conversations, such arguments appear in the unmarked basic S-IO-DO-V order. Third, as studies indicate (e.g., Kuno 1973b; Desző 1982; Kim 1988), syntactic positions, particularly the sentence-initial position and the immediately preverbal position, are generally reserved for two distinct discourse functions — Topic and Focus, respectively. Lastly, scrambling is predominantly a main clause phenomenon. It takes place typically at the level of matrix clauses. This may be because the main clause is the domain which is most susceptible to pragmatic influences from discourse contexts.

An NP is another syntactic level which is regarded as highly susceptible to reordering because it very often contains argument-like constituents and a rich array of adjectival modifiers within the structure. However, one may