

# Word order and focus in the Sandawe irrealis

Helen Eaton

School of Linguistics & Applied Language Studies, The University of Reading

*Abstract.* This paper reports on some original research conducted on Sandawe, a Khoisan language spoken in Tanzania.<sup>1</sup> Firstly, a review is presented of previous research on the issue of word order and focus in irrealis sentences in this language and the different positions of the main Sandawe scholars are contrasted with each other. This review is followed by a comparison of the linguistic informants used in these studies of Sandawe with those used in the author's research. A description of a sentence anagram task recently carried out by the author follows on from this. The rationale for the task is presented and the procedure discussed. An explanation of the categories of analysis is given. The results from the sentence anagram task are then discussed. In the concluding sections, the findings are reviewed and briefly compared with those for realis sentences. Finally, a provisional structural analysis is proposed for Sandawe.

## 1. Introduction

Dalgish (1979) describes Sandawe as a free word order language. This position is contested by Kagaya (1990) and Elderkin (1989, 1994), who both claim that there are restrictions on word order relating to inflection and focus. The research to be reported here aims to address the issue of whether Sandawe is a free word order language and to consider to what extent, if any, focus plays a role in determining word order.

If focus does play such a role, should Sandawe be described as a *discourse configurational* language? This is a language type 'in which primary sentence articulation is motivated by discourse-semantic, rather than theta role or case, considerations' (Kiss 1995:3). By Kiss's definition, one of the properties of discourse configurational languages concerns how focus is realised. In such languages, the '(discourse-)semantic function 'focus', expressing identification, is realized through a particular structural relation (that is, by movement into a particular structural position)' (1995:6).

---

<sup>1</sup> All the data discussed here was collected by the author during a field trip to Tanzania in 2000. My main informant during this time was Nestori Michaeli of Magambua, Usandawe. I would like to record here my thanks to him and to Daniel and Elisabeth Hunziker of SIL International, without whom this research would not have been possible.

The focus referred to in the above quotation is *identificational focus*, the function of which can be defined as follows:

An identificational focus represents a subset of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate phrase actually holds (Kiss 1998:245).

This type of focus may be contrastive, but need not be so. It is to be contrasted with *information focus*, which expresses new, non-presupposed information (to be defined more precisely below). The following examples illustrate the difference between these two types of focus:

- (1) It was *a fish* the woman cooked.
- (2) The woman cooked *a fish*.

The first sentence is an example of identificational focus and is expressed in English by a cleft sentence. We understand that the woman cooked only a fish and nothing else. In the second sentence, ‘a fish’ has information focus and is marked by means of intonation. It is an appropriate answer to a question such as, ‘What did the woman cook?’ Sentence (2) leaves open the possibility that the woman also cooked other items and thus does not express exhaustive identification in the way that sentence (1) does.

Kiss (1998:249) shows how in Hungarian identificational focus must be realised by the movement of the relevant constituent into a preverbal position. In contrast, information foci typically appear in situ, in a postverbal position. Later Kiss (1998:259) states that information focus ‘does not have a designated structural position in the sentence’ in languages, although many languages show a preference for new information to follow old. The validity of this statement for Sandawe will be considered in the light of the data to be presented.

The type of focus to be discussed in relation to Sandawe in this paper is information focus. I follow Lambrecht’s (1994:213) functional definition of focus as ‘the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition’. Non-contrastive focus can be elicited by WH-questions and contrastive focus by yes/no questions, as the examples below show:

- (3) What did the woman cook?  
The woman cooked *a fish*. *Non-contrastive*

- (4) Did the woman cook a chicken?  
 No, the woman cooked *a fish*. *Contrastive*

Lambrecht (1994:290) views contrastiveness not as a category of grammar, but as a gradient concept which 'arises from particular inferences which we draw on the basis of given conversational contexts'. The two examples above represent opposite ends of this contrastiveness gradient.<sup>2</sup>

For the purposes of this paper, the data to be considered will be restricted to irrealis sentences, although some comparison with realis sentences will also be made. The limited scope of this paper makes such a restriction necessary. The irrealis has been chosen as it the less complex of the two sentence types and can thus be treated in more detail within the confines of a paper of this length. A discussion of the irrealis also provides a good departure point for a consideration of the realis. Can the analysis proposed here for the irrealis be extended easily to handle the realis?

## ***2. Previous research on word order and focus in the Sandawe irrealis***

In the following two sections the previous research on word order and focus in the Sandawe irrealis is reviewed. In the third section a comparison is made between the language informants used in the previous research and those used by the author in the research to be reported in this paper.

### *2.1 Word order*

The *realis/irrealis* distinction is one of mood. It has been defined as 'the grammaticalised expression of the location of an event or state in either the real world or in some hypothesised, but not real, world' (Elliott 2000:81). Sandawe marks reality status formally by the use of a portmanteau morpheme indicating the realis/irrealis distinction together with the person, gender and number of the subject of the verb. I follow Elderkin (1986:133, after Hagman 1977:41 for Nama) by referring to this morpheme as a *pgn morpheme*. Since tense is not marked in the language, one of the chief functions of the *pgn* morpheme is to distinguish between non-future and future events, as first noted by Elderkin (1989:28). The following examples from my data illustrate the use of different *pgn* morphemes:

---

<sup>2</sup> In contrast, for Kiss (1998:267), contrastiveness is a feature of identificational focus, as is exhaustiveness. The identificational foci of different languages have either or both of these features.

- (5) O,, Ỹ aY4gi ˘ · gc4aVU4' h<] 4aY4-gUR  
 woman fish cook-3f.sg. realis pgn  
 'The woman cooked a fish'
- (6) O,, Ỹ aY4gi ˘ · gc4aVU4' h<] 4aYR-gi ,  
 woman fish cook-3f.sg. irrealis pgn  
 'The woman will cook a fish'
- (7) O,, Ỹ aY4gY4ö' \_c4\_c4' · X c˘ acR-U' ·  
 man chicken buy-3m.sg. realis pgn  
 'The man bought a chicken'
- (8) O,, Ỹ aY4gY4ö' \_c4\_c4' · X c˘ acR-] ·  
 man chicken buy-3m.sg. irrealis pgn  
 'The man will buy a chicken'

The only inflection pattern which is acceptable for an irrealis sentence is one in which the verb is inflected and no other constituent is inflected (Dalgish 1979:279). In contrast, in realis sentences other non-subjects may be inflected with the pgn morpheme, subject to restrictions on word order and focus, which will not be discussed here (see Eaton, forthcoming).

Dalgish (1979:279) claims that in irrealis sentences word order is free, although 'SOV word order is statistically the most common'. This view that all six logically possible orders of S, O and V are acceptable for an irrealis declarative, is also expressed by Kagaya (1990:6).

Elderkin (1991:109) does not explicitly state which word orders he considers acceptable for the irrealis, but he does claim that 'the basic clause allows variation in word order; less is possible with the future'. By 'future' Elderkin (1989:28) is referring to one use of the irrealis. He (1989:29) analyses the irrealis as having a 'nominal clause' as its basis. This nominal clause consists of two noun phrases in juxtaposition, the second of which is suffixed with a pgn morpheme. For the irrealis, the NP<sub>2</sub> consists of a VP followed by a pgn morpheme. Elderkin (1989:122) takes the inability of the declarative morpheme /-[ UR/ to follow the object in an irrealis sentence to be evidence that the object and the verb form one constituent in this sentence type. It is possible to infer from this that Elderkin would not accept irrealis word orders which do not preserve the uninterrupted sequence OV. In which case, only SOV and OVS are to be accepted. However, this view is not made explicit in Elderkin's work.

## 2.2 Focus

The preceding section has shown how there is no universal view on the extent to which word order may vary in the Sandawe irrealis. An equally contentious issue concerns whether this word order variation is influenced in any way by focus.

Early on in his paper, Dalgish (1979:274) makes the following statement:

The word order in Sandawe is, to a great extent, free, in that meaningful lexical elements may appear in any order without significantly affecting topic, focus, definiteness, etc. [and] without necessarily implying greater emphasis in one way or another.

Dalgish does not say on what basis he is making this claim. He does not give any examples of sentences in context and makes no further mention of focus. It is of course possible, if unlikely, that Sandawe word order is free to the extent that Dalgish claims, but without any evidence to support it, such a claim is hard to believe.

Like Dalgish, Kagaya studied Sandawe word order and inflection without reference to discourse factors. He observes some restrictions on the word order of Sandawe and concludes that word order is ‘not completely free’ in the language (Kagaya 1990:11). The aim of his work is to account for *how* Sandawe word order and inflection patterns are restricted, in both the realis and the irrealis. He does not concentrate on *why* a particular constituent is chosen to be inflected or *why* a particular word order is used. However, he does comment in a footnote that in *realis* sentences, the word that has the pgn morpheme suffixed to it receives more ‘emphasis’ than the other words (Kagaya 1990:2). If the translation ‘emphasis’ is taken to be equivalent to focus, Kagaya’s analysis of the realis data suggests that the inflection placement for a sentence is chosen first, on the basis of focus considerations, and then the word order is dependent on this. He does not comment on whether ‘emphasis’ is a factor in the choice of word order for irrealis sentences.

In contrast with Dalgish and Kagaya, Elderkin worked with Sandawe texts as well by the more unnatural method of asking an informant questions. He criticises the other two researchers for claiming that Sandawe has a tendency towards free word order on the basis of data elicited from informants in a ‘textless vacuum’ (Elderkin 1994:1). Elderkin (1994:12) analyses some transcribed stories in his 1994 paper and comes to the conclusion that, ‘It is not true that Sandawe is a free word order language’. Although none of Elderkin’s published work deals in depth

with focus in Sandawe, he has made several important observations on this subject, with respect both to the realis and the irrealis.

To begin with the realis, Elderkin (1989:27) refers to any constituent which is suffixed with either the nominative marker or a pgn morpheme as a ‘marked’ constituent. These two types of morpheme have a function in common, namely giving the constituent to which they attach ‘some sort of prominence in the information structure of the clause’ (Elderkin 1986:108). Elderkin (1994:5) also makes some observations about the relationship between word order and information structure in Sandawe. He remarks that fronting is a means that the language employs to give a constituent discourse prominence. Elderkin (1989:92) uses the terms *focus* and *emphasis* with ‘the meaning as defined by the formal phenomena to which they refer’. By his interpretation a focused constituent is one which is fronted, inflected and given a higher pitch level, whereas an emphasised constituent is one which remains in situ, is uninflected, but has a higher pitch level.

With respect to the irrealis, Elderkin (1991:110) notes that when there is no overt object, the subject can be made an ‘information point’ with tone levels. The low tone pattern on the verb contrasts with the high tone pattern of the preceding focused constituent.<sup>3</sup> The examples below from my data illustrate this phenomenon:

- (9) ‘Is *Ts’inkwa* going to cook?’ *Subject focus*
- Ũ #U4                      bURã              h<˘ ÚaY˘ · R-gi ,  
 no                      Nam                      cook-3f.sg. irr.  
 ‘No, *Nam* is going to cook’

In sentence (9) the focused subject is immediately followed by the low-toned verb. Thus the subject is given prominence in the utterance as it alone has the highest pitch. In contrast, the focused verb in the next two examples occurs with a high tone pattern and in the second of the two examples, the verb precedes the subject:

- (10) ‘Is *Nam* going to *sweep*?’ *Verb focus*

<sup>3</sup> A similar phenomenon is noted by Bearth (1987:33, 1992:82) for Toura, a Southern Mande language. In this language, a lowered tone on certain verb forms marks information focus on the *postverbal* constituent.

- (a) Ũ #U4̃ · bURã · h<̃ ØaYR-gi , ·  
 no Nam cook-3f.sg. irr.  
 ‘No, Nam is going to *cook*’
- (b) Ũ #U4̃ · h<̃ ØaYR-gi , · bURã ·  
 no cook-3f.sg. irr. Nam  
 ‘No, Nam is going to *cook*’

If an overt object is included in an irrealis sentence it is not possible to focus the preceding subject by lowering the tone pattern of the object. This has been noted by Elderkin (1991:111). He explains:

The reason is clearly one of interpretation strategy: to have the subject NP one tone level higher than the following NP which is the object of the verb, would produce a tonal relationship between them which would be interpreted as two NPs in an associative construction.

Thus the following sentence, from my own data, means ‘She will cook Nam’s fish’ and not ‘*Nam* will cook a fish’:

- (11) bURã · gc̃ aVŨ · h<̃ ÚaỸ · R-gi , ·  
 Nam fish cook-3f.sg. irr.  
 ‘She will cook Nam’s fish’  
 \* ‘*Nam* will cook a fish’

According to the data collected from my informant, the subject *may* be focused by tone levels in such a sentence, but a change of word order is necessary:

- (12) ‘Will *Ts’inkwa* cook a fish?’
- Ũ #U4̃ · bURã · h<̃ ÚaỸ · R-gi , ·  
 gc4aVU4̃ ·  
 no Nam cook-3f.sg. irr. fish  
 ‘No, *Nam* will cook a fish’

If the word order is SVO then it is possible to lower the tone pattern of the verb following the focused subject and thereby use the same focusing mechanism observed above in intransitive sentences. The subject and object occur with their basic tone patterns, and are separated by the verb,

so they are not understood as forming an associative construction with the meaning ‘Nam’s fish’. As was discussed above, it is possible to conclude that Elderkin’s analysis of the object and the verb as one constituent in the irrealis rules out the possibility of the word order SVO. In which case, the possibility of realising subject focus as in (12) above is also rejected.

### *2.3 A comparison of the informants used by author and those used in previous research on the Sandawe irrealis*

Four informants were used for the author’s research, all of whom are native speakers of Sandawe living in Usandawe.<sup>4</sup> Most of the work was done with one informant, Nestori Michaeli, who has worked with linguists from SIL International for several years and is experienced in the procedure of language elicitation. Further data was collected from two male Sandawe speakers and one female speaker. All four informants are competent in Swahili, but speak mainly Sandawe in their everyday lives.

Dalgish (1979:274) names two main informants, Mr. J.G.D. Wagine and Ms. R. Duma, both from villages in Usandawe. He carried out his research whilst at the University of Dar es Salaam and does not mention whether his informants were living in the city at the time of his research. Dalgish mainly used English during the elicitation process, with some Swahili.

Elderkin (1989:14) worked with several informants between 1974 and 1983, all of whom were based in Dar es Salaam,. He also made a short visit to Usandawe in 1982, where he made some recordings of spoken Sandawe.

Kagaya (1990:11) had only one informant for his data. He (1990:2) names this informant as Mr. Fredrick Duma, who was also Kagaya’s informant for a later piece of work (1993). In the preface to this second piece of work, the author (1993:i) briefly describes his informant’s linguistic background:

Mr. Fredrick Duma [...] was born in 1955 at Kurio village, Kwa-Mtoro ward, Kondoa District of Dodoma Region, and brought up there until 1973. Then, he moved to Dodoma, and Dar es Salaam for his study. He can speak Swahili, Nyaturu and English besides his mother tongue, Sandawe.

---

<sup>4</sup> *Usandawe* is the Swahili term for the Sandawe speaking area of Tanzania. It does not correspond exactly to any official district.



The data collection took place over thirteen years after Mr. Duma moved away from Usandawe.

The judgments of these various informants used in previous research could have been affected by the length of time they had spent away from Usandawe. In addition, no mention is made of whether the informants were regularly using Sandawe, other than in language elicitation sessions. The ability to make fine judgments on matters of focus may be affected by a lack of regular language use with other native speakers.

### ***3. The sentence anagram task***

The rationale, procedure and categories of analysis for the sentence anagram task are considered in the next section. This is followed by a presentation and discussion of the results of the task.

#### *3.1 The rationale for the task*

A sentence anagram task was designed and carried out with the main informant in order to investigate whether word order in the Sandawe irrealis is affected by focus.

In a sentence anagram task a subject is required to form a sentence by manipulating a set of sentence fragments. It is a technique more commonly used with the language-impaired, but it is also a useful tool to gather information on sentence construction in speakers of unimpaired language.

The informant was required to produce several sentences from the same set of sentence fragments, but in response to different question prompts from the investigator, designed to elicit responses with varying types of focus, as will be explained below. This task ensured that the informant was constrained in the sentences he could produce and the focus of the sentences was carefully controlled.

12 different target sentences were chosen and up to 22 different questions were used for each target. In this way it was possible to compare the responses for individual target sentences on the one hand and those for each type of focus on the other hand. For example, from the data set of 186 sentences, 12 sentences were elicited with contrastive subject focus. Each of those 12 sentences is an example of a different target sentence. 13 of the 186 sentences were examples of target sentence A. Each of those 13 sentences is an example of a different type of focus.

If word order in the Sandawe irrealis is affected by focus, we would expect to see variation in word order in the set of the 13 examples of target

sentence A as these items were elicited with different types of focus. In contrast, we would expect to see similarities in the word order of the set of 12 sentences elicited with contrastive subject focus as they were all elicited with the same type of focus.

### *3.2 The task procedure*

The task procedure was explained to the informant in Swahili and he was given several practice items using different sentences so that he could become accustomed to what was required of him before the task proper began. For each target sentence a set of cards transcribed with the required lexical items was prepared and mixed together. As each question was asked the informant was handed the corresponding set of cards. He was then asked to respond to the question by forming a sentence using all the cards in whatever order he preferred. Only one constituent was written on each card so that the informant had maximum freedom in the order of constituents for the sentences he formed.

Before the task began, it was explained to the informant that he was required to form only the first sentence pattern which came to mind and not to spend time thinking of alternative ways to answer the question. Once he had formed his answer and was satisfied with it, he read out the sentence he had formed from the cards twice and whistled the tone pattern for it once.<sup>5</sup> The elicitation sessions were taped on audiocassette so that the tone of the sentence responses could be analysed later. During the task itself I transcribed the sentences which the informant produced in an abbreviated form, such as 'SOV<sup>H</sup>', next to the relevant question as a back up to the tape.

The constituents of the 12 target sentences were carefully selected to ensure only one sensible sentence could be made from each set of constituents, with the exception of the constituents for targets K and L, which were deliberately chosen for their potential ambiguity (see below). The questions were put to the informant in a randomised order and it was ensured that neither the same target sentences nor examples of the same type of question were ever ordered consecutively.

Sandawe is not yet a written language, but the informant is able to read the language transcribed phonetically, having worked with SIL International linguists for several years. The words on the cards were not written with their tone patterns marked since it has been claimed that the tone pattern of an irrealis verb varies according to whether it is focused

---

<sup>5</sup>The informant has been taught to whistle the tone patterns of sentences which he produces and this facilitates an accurate transcription.

(see above, section 2.2). Writing the tone in for the words used might have affected the informant's choice of sentence pattern. None of the words used in the experiment is a member of a tonal minimal pair.

### 3.3 Categories of analysis

The 12 target sentences chosen for the investigation are given in the appendix to this paper. These 12 targets can be grouped into pairs. The members of each pair differ in the choice of lexical items, but not in the type of construction expected. The six types of target constructions are shown below:

*Table 1 Content of target sentences*

<i>Target sentence</i>	<i>Constituents</i>
AB	SOV
CD	SVAdvP
EF	SOVAdvP
GH	SOVPP
IJ	SOV (with heavy object)
KL	SOV (reversible subject and object)

The first four target types simply contain different combinations of constituents. Targets I and J contain a heavy object (a possessive construction) to see if this has any effect on how the sentence is focused. Targets K and L have been chosen to observe the effects of potential ambiguity on sentence patterns. In both sentences the subject and object are of the same gender and are both semantically capable of being either the subject or the object. Targets I, J, K and L can be compared with targets A and B, which are simple SOV sentences.

Questions were prepared to elicit each target construction with different constituents in focus and with different types of focus. An example of this is given below:

- (13) A(2)(i) Who will buy the cow?  
           <sup>^</sup> <opi will buy the cow.

The target construction is shown in bold type. By this question it is elicited with non-contrastive subject focus. This is indicated by the key A(2)(i), which shows which target construction it is by the letter, which constituents are in focus by the first number and what the type of focus is

by the second number. The following categories of constituents in focus were investigated:

1	sentence <sup>6</sup>	11	subject and AdvP
2	subject	12	object and AdvP
3	object	13	verb and AdvP
4	verb	14	subject, object and verb
5	subject and object	15	subject, object and AdvP
6	subject and verb	16	subject, verb and AdvP
7	object and verb	17	object, verb and AdvP
8	sentential polarity <sup>7</sup>	18	possessive modifier
9a	tense/aspect (cont.)	19	possessive head
9b	tense/aspect (perf.)	20	postpositional phrase (PP)
10	adverbial phrase (AdvP)		

Two types of focus were considered in the sentence anagram task:

- (i) non-contrastive focus
- (ii) contrastive focus

Non-contrastive focus was elicited by WH-questions and contrastive focus by yes/no questions, as the following examples show:

- (14) D(2)(i) ‘Who will sweep tomorrow?’  
 ‘K’ats’awa will sweep tomorrow’  
*Non-contrastive subject focus*
- (15) D(2)(ii) ‘Will ^ oka sweep tomorrow?’  
 ‘No, K’ats’awa will sweep tomorrow’  
*Contrastive subject focus*

Not all the questions were used for eliciting both types of focus. Questions 8, 9a and 9b were not used for eliciting non-contrastive focus

<sup>6</sup> Sentence focus is the type of focus in which ‘the domain of the “new information” extends over the entire proposition, including the subject’ (Lambrecht 1994:14). All 12 target sentences used in the sentence anagram task were elicited with non-contrastive sentence focus by the question, ‘What happened?’.

<sup>7</sup> For example, D(8)(ii): ‘K’ats’awa won’t sweep tomorrow, will she?’  
 ‘No, K’ats’awa *will* sweep tomorrow’

Sandawe contradicts a question such as the one above by using the word /Ũ #U4/. A card with this word written on it was given to the informant each time a polarity focus question was asked. The term *polarity focus* is taken from Lambrecht (1994:236).

and all those questions focusing more than one constituent were not used for eliciting contrastive focus. The reason for this is that these questions are extremely contrived. For example, a suitable question for contrastive subject and verb focus would be, ‘Will *ˆ*oka cook tomorrow?’, with the rather unnatural reply, ‘No, *K’ats’awa* will *sweep* tomorrow’.

The difference between types (9a) and (9b), which focus tense and aspect, is illustrated below:

- (16) L(9a)(ii) ‘Is Gawa beating up Degera now?’  
‘No, Gawa *will* beat up Degera *later*’

- (17) L(9b)(ii) ‘Has Gawa already beaten up Degera?’  
‘No, Gawa *will* beat up Degera *later*’

It was necessary to hand the informant a card transcribed with the word /*ç*Y4#/, meaning ‘later’, for questions (9a) and (9b), except with targets C, D, E and F, which already contain an AdvP specifying future time. This is because the irrealis does not always express future time. In Sandawe it is possible to respond to the two questions above in the affirmative with an irrealis sentence and not contradict the time reference of the questions. The addition of the AdvP /*ç*Y4#/ makes it clear that the irrealis is expressing future time.

### 3.4 Results and analysis

In total, 186 sentences were elicited during the sentence anagram task. The word orders shown in the results are summarised below, beginning with the relative order of the subject and the verb:

*Table 2 Relative order of S and V*

<i>Word order</i>	<i>Number of occurrences</i>
SV	186
VS	0
<i>Total</i>	186

In all the sentences in the data set the subject preceded the verb. The relative order of the object and the verb was not as fixed as that of the subject and the verb, as is shown in the next table:

*Table 3 Relative order of S, O and V*

<i>Word order</i>	<i>Number of occurrences</i>
SOV	143
SVO	17
Other	0
<i>Total</i>	160

11% of sentences with a subject, object and verb occurred in the order SVO. Neither the presence of a heavy object (targets I and J) nor the reversible nature of the subject and object (targets K and L) made a difference to the word order of the sentences produced.

The position of the adjunct phrases AdvP and PP relative to the verb showed more variation than the position of the subject and object:

*Table 4 Relative order of AdvP/PP and V*

<i>Word order</i>	<i>Number of occurrences<sup>8</sup></i>
AdvP/PP V	41
V AdvP/PP	75
<i>Total</i>	116

The relationship between all these orders and focus will now be considered.

The table below summarises the word order positions of focused non-verbs, excluding those which are focused together with other constituents:

*Table 5 Position of single focused constituents relative to the verb*

<i>Constituent in focus</i>	<i>Preverbal occurrences</i>		<i>Postverbal occurrences</i>	<i>Total</i>
	<i>Immediately</i>	<i>Non-imm.</i>		
S	5	19	0	24
O	20	0	0	20
AdvP/PP	31	5	0	36
<i>Total</i>	56	24	0	80

All 80 single focused constituents are found in a preverbal position. The first row of this table does not reveal anything new as *all* subjects occurred

<sup>8</sup> The two items which contain a preverbal AdvP and a postverbal PP are included in both rows of the table. Two further items, namely the two occurrences of SOPPAdvPV, are counted twice in the first row.

preverbally, regardless of whether they were focused. The second row is more interesting since 11% of all objects were found in the postverbal position, but none of the single focused objects occurred in this position. In fact, none of the 17 SVO orders contained a focused object together with another focused constituent either.

The third row of the table is more interesting still. All 36 single focused AdvPs and PPs were found in a preverbal position, and 31 of these occurred in the immediately preverbal position. This can be compared with table 4 above, which shows that in total 41 out of 116 adverbs and PPs (35%) occurred in a preverbal position. The remaining 5 preverbal constituents from this figure of 41 are unfocused or in sentences with multiple focus. The 75 other unfocused and multiple focused AdvP/PPs are found in a postverbal position. Clearly, a focused AdvP or PP occurs preverbally in the Sandawe irrealis. At this point it is useful to consider the position of the object in relation to a focused AdvP, as shown in the next table:

*Table 6 The position of the object in sentences with AdvP focus*

<i>Word order</i>	SOAdvPV <sup>L</sup>	SAdvPV <sup>L</sup> O	SAdvPOV <sup>L</sup>	<i>Total</i>
<i>Frequency</i>	14	7	3	24

The particularly interesting point to be made here is that the position of the object seems unimportant. That is, it can be either preverbal or postverbal. Instead, what seems to be important is that the immediately preverbal position is made available for the focused AdvP.

As well as word order, the tone pattern of the verb was also investigated in the sentence anagram task:

*Table 7 Distribution of verb tone patterns*

<i>Tone pattern</i>	<i>Number of occurrences</i>
V <sup>L</sup>	149
V <sup>H</sup>	26
V <sup>H*</sup>	11
<i>Total</i>	186

V<sup>H\*</sup> stands for a verb which occurred with a high tone pattern because it followed an all low-toned noun. This is a phonological rule which is observed regardless of any focus considerations. All occurrences of a high-toned verb which cannot be explained by this rule are represented by V<sup>H</sup> in

the second row of the table.  $V^H$  occurs mainly with polarity focus and non-contrastive verb focus (21 examples), but also with multiple focus items including a focused verb (3 examples) and with single contrastive verb focus (2 examples). It is therefore best described as a verb focus pattern. The next table shows more clearly how these results break down and how they relate to word order:

*Table 8 A comparison of verb focus and polarity focus patterns*

<i>Focus type</i>	<i>Sentence pattern and frequency of occurrence</i>		
	S(O)V <sup>L</sup> (AP/PP)	S(O)V <sup>H</sup> (AP/PP)	Other
V(non-con)	1	9	SAPV <sup>H</sup> (1), SOV <sup>H*</sup> (1)
V(con)	9	2	SOV <sup>H*</sup> (1)
V+	23	3	SV <sup>L</sup> O (3), SOV <sup>H*</sup> (2), SOAPV <sup>L</sup> (1)
Polarity	0	5	SV <sup>H</sup> O (3), SV <sup>H</sup> OAP (2), SOV <sup>H*</sup> (1), SV <sup>H</sup> OPP (1)
<i>Total</i>	33	19	16

The basic word order is used for all but one of the 24 items with single verb focus. The main variation is instead in the tone pattern of the verb. There is a clear difference between non-contrastive and contrastive focus with respect to this. Excluding the  $V^{H*}$  items, 10 out of 11 sentences with a non-contrastively focused verb have a verb with a high tone pattern, whereas 9 out of 11 sentences with a contrastively focused verb have a verb with a low tone pattern. When the verb is focused along with other constituents, it tends to have a low tone pattern (27 out of 30 items, excluding  $V^{H*}$ ). It is interesting that multiple focus here patterns with *contrastive* verb focus, since all the multiple focus items in the task involve *non-contrastive* focus. On the other hand, polarity focus, which involves *contrastive* focus, follows the *non-contrastive* single verb focus pattern of having a verb with a high tone pattern (11 out of 11 items, excluding  $V^{H*}$ ). The word order of the polarity focus items is also interesting as SVO is the preferred order for 6 examples and SOV for 4, which goes against the overall trend in the data for SOV.

It was suggested in section 2.2 above that the word order SV<sup>L</sup>O is a pattern used for examples of subject focus. However, in the 56 examples in



the data set which contain a focused subject (excluding the 12 examples of sentence focus, but including other examples of multiple focus), only 4 follow this pattern. These examples are the only ones to follow the pattern  $SV^L O$ . 3 of the 4 examples are from the group of 12 sentences in which the subject and the verb are focused. Is the verb positioned between the subject and the object in these three examples because the subject is focused or because the verb is focused? The examples of verb focus discussed in the preceding paragraph do not suggest that the position of the verb is affected by its own focus status, although the examples of polarity focus provide evidence for the opposite view. The remaining example of the pattern  $SV^L O$  is of target sentence A with the subject as its only focused constituent. Putting the tone pattern of the verb to one side, 9 of the 17 occurrences of the order SVO involve a focused verb, 7 involve a focused AdvP and one a single focused subject.

In summary,  $S(O)V^L(AdvP/PP)$  is the basic sentence pattern and occurs with all sentence focus and object focus items and most examples of subject focus and multiple focus.  $AdvP/PP, V^L$  is the sentence pattern for AdvP or PP focus.  $V^H$  occurs to varying extents with different kinds of verb focus.  $SV^L O$  occurs with a few subject focus examples.

#### ***4. Further discussion***

The overall pattern for the Sandawe irrealis is that non-verbs are focused by preceding a low-toned verb, whereas verbs can be focused by occurring with a high tone pattern, but also occur with a low tone pattern. Parallels can be drawn here with the realis, where there is a distinction between verbs and non-verbs. In the realis, non-verbs are focused by being suffixed with an appropriate inflectional morpheme. A verb may also be focused by being inflected if it is the only inflected constituent of a sentence, but otherwise it is usually left uninflected. For both the irrealis and the realis, non-verbs follow a particular focusing mechanism more closely than verbs do.

A further similarity is that a focused non-verb tends to occur before a verb in both the irrealis and the realis, but for different reasons. In the irrealis, a focused constituent sometimes occurs immediately before the verb, in which case it is focused by the phonological contrast with the low tone of the verb. In the realis, a focused constituent must usually occur before the verb so that it may be suffixed with an inflectional morpheme. This is because the language has the restriction that an uninflected verb

may not precede the first inflected constituent of a sentence (Elderkin 1989:106, Kagaya 1990:5).

Whilst a low-toned verb in the irrealis and an uninflected verb in the realis can be seen as equivalent since they are often preceded by focused constituents, a high-toned verb in the irrealis and an inflected verb in the realis are equivalent in another sense. A sentence-initial verb in the irrealis must occur with a high tone pattern, regardless of the focus of the sentence. The realis parallel of this is that a sentence-initial verb in the realis must be inflected, irrespective of its focus status.

For the realis, word order is not a primary factor in sentence pattern choice, but one at least partly dependent on other factors. Focus primarily affects the choice of inflection pattern, which in turn may affect word order. In contrast, for the irrealis, word order is more important. The mechanism of focusing a constituent by it immediately preceding a low-toned verb is an equal combination of phonological and word order means. In addition, those focused constituents which are not immediately preverbal, do tend to occur preverbally, without any apparent mediating cause. That is, this word order effect is not dependent on any other factor but focus.

In the case of the irrealis, the tone pattern of the verb and the linear position of the focused constituent relative to the verb are the important factors. With the realis, the inflectional status of the focused constituent and its linear position relative to the verb are the important factors. For both the irrealis and the realis, sentence pattern choice is affected by focus.

The question was posed in the introduction as to whether Sandawe can be described as a discourse configurational language. By Kiss's (1995:6) definition, one of the properties of discourse configurational languages is that the '(discourse-)semantic function 'focus', expressing identification, is realized through a particular structural relation (that is, by movement into a particular structural position)'. Sandawe does exhibit this property to an extent as the data presented here has argued for the position that focus in the Sandawe irrealis is realised partly by word order. However, two important differences must be noted between Kiss's definition of the property and what can be observed in the Sandawe data. Firstly, the type of focus considered in this paper has been *information* focus (as defined by Lambrecht 1994:213) rather than identificational focus, as explained in the introduction. Secondly, the definition refers to a 'movement into a particular structural position' and it has not been claimed here that such a movement takes place in Sandawe. In fact, the data suggests that it is a particular linear position rather than a particular structural one which is important in realising focus. Recall that in the Sandawe irrealis focused

constituents are found preverbally, but not necessarily immediately preverbally. Thus both the orders SOAdvPV and SAdvPOV express AdvP focus.

It was mentioned in the introduction how Kiss states that information focus ‘does not have a designated structural position in the sentence’. The data presented here from Sandawe supports this position as it does not show that constituents with information focus are found in a particular structural position. However, the data does suggest that such constituents are found in a designated *linear* position, namely in a preverbal position. This phenomenon is not the same as the tendency for many languages to show a preference for new information to follow old. As was shown in *Table 5* above, all 80 single focused non-verbs in the data set were found preverbally. In Sandawe this linear effect is not a statistical preference, but a rule without exceptions.

56 out of these 80 single focused constituents in the data set (70%) occur immediately preverbally. They follow the often observed tendency that focus position is next to an inflected verb (Kiss 1995:20). Kiss (1997:720) also notes that there is a strong cross-linguistic correlation between SOV word order and preverbal structural focus. Sandawe follows this tendency with respect to its preverbal linear focus position, as it is an SOV language (Dalgish 1979:274, Elderkin 1989:27).

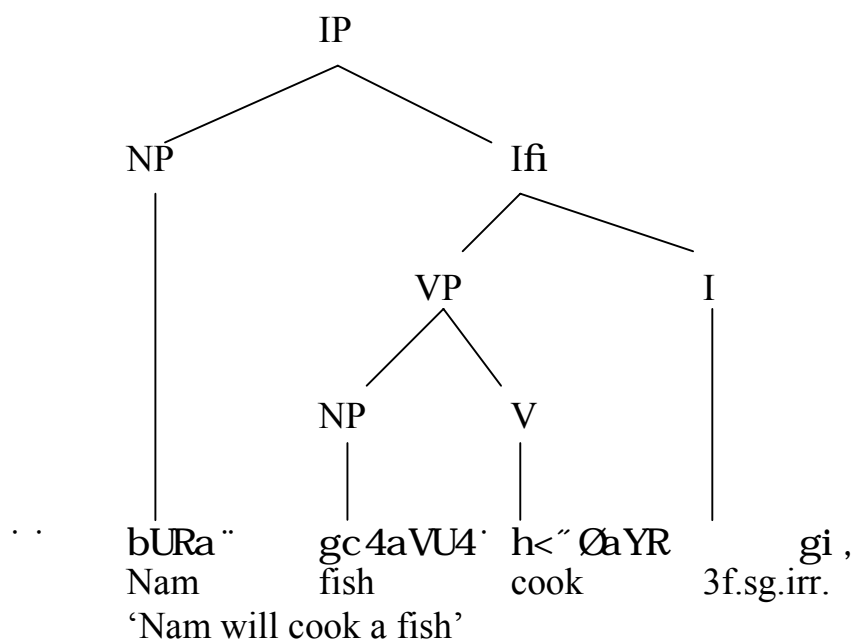
It was noted earlier that the position of the object in sentences with AdvP focus in the data set varied considerably, but what seemed to be important was that the object left the immediately preverbal slot free for the AdvP. In this sense, Sandawe could be analysed as employing an indirect focus marking strategy, in a similar way to Swahili, as Bearth (1999:129) explains:

Swahili, for instance, makes extensive use of a type of word order permutation whereby focus effects, particularly on the verb, are obtained by moving postverbal constituents to the front, in topic position, and thereby defocusing them. It is not the focus constituent which undergoes movement, but the special focus effect is obtained by moving a non-focalised constituent out of its (unmarked) focus position.

In Sandawe, it is sometimes the case that the normally immediately preverbal object is moved so that another constituent may occur in the focus position. The adjacent verb occurs with a low tone pattern in order to mark information focus on the immediately preverbal constituent. Thus in two senses, the focus marking strategy for the Sandawe irrealis is indirect

as it may involve the movement of an unfocused constituent and the lowering of the tone of another unfocused constituent.

Although this paper has concentrated on analysing the linear word order of the Sandawe irrealis rather than on proposing a structural analysis, a few comments should be made here with respect to structure. Assuming a standard government and binding framework, Sandawe is taken to be a consistently head-final language<sup>9</sup> with verb movement from V to I as the verb must be inflected. The basic clause structure is therefore the following:



*Fig. 1 Structure of an irrealis declarative*

According to such a structure, the order SOAdvP<sub>+FOC</sub>V involves the right adjunction of the AdvP to the VP and the order SAdvP<sub>+FOC</sub>OV involves the left adjunction of the AdvP to the VP. Unfocused AdvP is found postverbally, as a right adjunction to IP.

## 5. Conclusion

<sup>9</sup> Some support for this assumption is that possessive constructions such as 'Nam's fish' follow the order modifier and then head and that the language has postpositions.

The analysis proposed above must be seen as a provisional one. Further work is clearly required. In particular, the question of how such an analysis handles the realis needs to be addressed. With its many different inflectional possibilities and the interactions of these with word order and focus, the realis presents additional challenges for a structural analysis.

The data presented in this paper has shown that word order in the Sandawe irrealis is not free in the sense of allowing alternative sequences at random, but it is influenced by focus considerations and interacts with the tone pattern of the verb.

### *References*

- Bearth, T. (1987). La logique des focalisations - le cas du toura. *Mandenkan* **14-15**. 29-45.
- Bearth, T. (1992). Constituent structure, natural focus hierarchy and focus types in Toura. *Folia Linguistica* **26/1-2**. 75-94.
- Bearth, T. (1999). The contribution of African linguistics towards a general theory of focus. Update and critical review. *Journal of African Languages and Linguistics* **20**. 121-156.
- Dalgish, G. M. (1979). Subject identification strategies and free word order: the case of Sandawe. *Studies in African Linguistics* **10**. 273-310.
- Eaton, H. C. (forthcoming). *The grammar of focus in Sandawe*. Ph.D. dissertation, University of Reading.
- Elderkin, E. D. (1986). Diachronic inferences from basic sentence and noun structure in Central Khoisan and Sandawe. *Sprache und Geschichte in Afrika* **7**. 131-156.
- Elderkin, E. D. (1989). *The significance and origin of the use of pitch in Sandawe*. D. Phil. thesis, University of York.
- Elderkin, E. D. (1991). Clause structure and tone in Sandawe. *York Papers in Linguistics* **15**. 93-115.
- Elderkin, E. D. (1994). *Sandawe verbal sentences*. Unpublished manuscript.
- Elliott, J. R. (2000). Realis and irrealis: forms and concepts of the grammaticalisation of reality. *Linguistic Typology* **4**. 55-90.
- Hagman, R. S. (1977). *Nama Hottentot Grammar*. Bloomington, Indiana: Indiana University Press.
- Kagaya, R. (1990). Restriction on word order of free word order language – the case of the Sandawe language. *Journal of Asian and African Studies* **40**. 1-12. Trans. F. Ingham-Nagasawa.
- Kagaya, R. (1993). A classified vocabulary of the Sandawe language. *Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA)*. *Asian and African Lexicon* **26**. Itubashi, Tokyo: Tokyo Press Co. Ltd.
- Kiss, K. É. (1995). Discourse configurational languages. Introduction. In Kiss, K. É. (ed.) *Discourse configurational languages*. Oxford: Oxford University Press. 3-27.
- Kiss, K. É. (1997). Discourse-configurationality in the languages of Europe. In Siewierska, A. (ed.) *Constituent order in the languages of Europe*. Berlin: Mouton de Gruyter. 681-727.

- Kiss, K. É. (1998). Identificational focus versus information focus. *Language* 74. 245-273.
- Lambrecht, K. (1994). *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

### Appendix

The 12 target sentences used in the sentence anagram task were the following:

- A    ^ < c ~ d ~ 4 ~    .    .    \ i ~ a V i E ~ R ~    X c ~ a c R - ] ~  
       ^ < o p i                      c o w - d e f .    b u y - 3 m . s g . i r r .  
       ‘ ^ < o p i w i l l b u y t h e c o w ’
- B    X Y 4 [ Y 4 f U R ~    .    .    ¿ U ñ ö ~ R ~    .    X c ~ a c ~ \_ U ~ R - ] ~  
       D e g e r a                      g o a t - d e f .    s e l l - 3 m . s g . i r r .  
       ‘ D e g e r a w i l l s e l l t h e g o a t ’
- C    ^ c ~ \_ U 4 ~    .    .    h < ~ Ú a Y ~ R - g i , ~    ¿ Y 4 #  
       ^ o k a                      c o o k - 3 f . s g . i r r .    l a t e r  
       ‘ ^ o k a w i l l c o o k l a t e r ’
- D    \_ f i U ~ h g f i U ~ k U 4 ~    .    .    „ < Y ~ a Y ~ R - g i , ~    d < Y R ~  
       K ’ a t s ’ a w a                      s w e e p - 3 f . s g . i r r .    t o m o r r o w  
       ‘ K ’ a t s ’ a w a w i l l s w e e p t o m o r r o w ’
- E    h U 4 ö a V U R ~ ~    X U ~ b ~ ß ä ~ ~    .    ~ Ú h G U ~ k U ~ R - ] ~  
       .    ¿ Y 4 #  
       T a a m b a                      a r r o w - d e f .                      l o o k f o r - 3 m . s g . i r r .                      l a t e r  
       ‘ T a a m b a w i l l l o o k f o r t h e a r r o w l a t e r ’
- F    ` Y 4 ö V U R ~    .    .    ` U ~ # g Y Ñ ñ ~ R ~    .    .    .  
       .    O , k Y ~ R - ] ~    d < Y R ~  
       L e e b a                      t h r e s h i n g t r o u g h - d e f .                      m e n d - 3 m . s g . i r r .    t o m o r r o w  
       ‘ L e e b a w i l l m e n d t h e t h r e s h i n g t r o u g h t o m o r r o w ’
- G    b U R a ~    .    .    b ^ U 4 b ^ U ~ ^ U ~ ö ~ R - g i , ~  
       .    a ~ Ú b X U ~ - h U ~ R ~    .    .    .  
       N a m                      t o m a t o e s                      p l a n t - 3 f . s g . i r r .                      f i e l d - l o c .  
       ‘ N a m w i l l p l a n t t o m a t o e s i n t h e f i e l d ’

- H h<Ỹ B\_U4ˆ fiŨ ö^URˆˆ · hkŨ #kŨ öˆ R-gi ,  
 · h<cRöbhŨ ·  
 T<enka acb\_Ymi d` i aˆˆ pick-3f.sg. irr. steppe-loc.  
 ‘T<enka will pick acb\_Ymi d` i aˆˆ in the steppe’ ·
- I l c̃ #U4ˆ · [ i ˘ ` U4` U4öˆ · ^Ũ ^Ũ ·  
 · ˘ ÚhGŨ kŨ öˆ R-gi , ·  
 Xo’a Gulalaa brother look for-3f.sg. irr.  
 ‘X’oa will look for Gulalaa’s brother’
- J VURf ] c̃ · \_c/ öB[ c4öˆ · VŨ öVŨ ·  
 · XŨ f Ũ öˆ R-gi , ·  
 Bar Koongoo father wait for-3f.sg. irr.  
 ‘Bar will wait for Koongoo’s father’ ·
- K [ Y4` Y4ˆ · ^ <c̃ dˆ Øˆ · ac̃ \_c̃ ˘ ` Ũ öˆ R-] ·  
 Gele ^ <opi greet-3m.sg. irr.  
 ‘Gele will greet ^ <opi’
- L [ U4kURˆ · XY4[ Y4f URˆ h` fiŨ di ˘ aỸ · R-] ·  
 Gawa Degera beat up-3m.sg. irr.  
 ‘Gawa will beat up Degera’