Word order and information structure in Makhuwa-Enahara
Wal, G.J. van der

3.   Grammar and information structure

In this second part of the thesis, which consists of chapters 3, 4 and 5, I discuss the syntax and information structure in Makhuwa. The language exhibits variability in word order, and information structure seems to be an influential factor in the word order and the conjoint/disjoint alternation. The current chapter discusses the general notions of information structure and minimalist syntax and introduces two models combining syntax and information structure. These models are applied to the Makhuwa data in chapters 4 and 5.

3.1   Configurationality

Mostly on the basis of word order properties, some languages have been called "configurational" and others "non-configurational". In a configurational language, the grammatical functions of subject and object appear in a particular structural relationship to each other. English is the standard example of a configurational language, where the syntactic functions of subject and object can be deduced from their position in the sentence. Hale (1983) was the first to describe the Australian language Warlpiri as non-configurational. He proposed the Configurationality Parameter, according to which non-configurational languages have three characteristics: 1. free word order (i.e., subject, verb and object can occur in any order); 2. extensive use of null-anaphora (pro-drop); 3. use of discontinuous NP constituents. Case-marking has been added to these properties (Neeleman and Weerman 1999, among others), since it was observed that free word order and case-marking often co-occur. A number of other languages, which do not exhibit all these characteristics, have also been named non-configurational, under a broader definition of non-configurationality suggested by (Bresnan and Mchombo 1987): subject and object functions are not distinctively encoded by phrase structure. Baker (2003) provides a list of these languages.

An *analysis* may be referred to as non-configurational when it explains the variable word order without refering to structure or configuration. However, since all sentences in all languages have a certain configuration, the term "non-configurational" does not seem appropriate to refer to *languages*. There are striking differences between languages in terms of word order and constructions, so the question is: what determines the configuration of sentences in a language? For the "configurational" languages, the most influential factor is the syntactic functions and argument relations. For languages where phrase structure does not reflect only syntactic functions, it has been proposed that their word order is determined by principles of discourse.

Li and Thompson (1976) distinguish languages according to the prominence of subject and topic. They claim that some languages, such as Chinese, can be more insightfully described by taking the discourse notion of topic to be basic and analysing the basic structure as topic-comment (rather than subject-predicate). This implies that in

topic-prominent languages the structural encoding of the discourse function "topic" is more important than the encoding of the syntactic function "subject" in the word order. Languages where the words in a sentence seem to be ordered according to the discourse functions have been called "discourse-configurational". É. Kiss (1995: 6) defines discourse-configurationality as follows. A language is discourse-configurational if (in intuitive terms):

A. The (discourse-)semantic function "topic", serving to foreground a specific individual that something will be predicated about (not necessarily identical with the grammatical subject), is expressed through a particular structural relation (in other words, it is associated with a particular structural position).

or

B. The (discourse-)semantic function "focus", expressing identification, is realised through a particular structural relation (that is, by movement into a particular structural position).

Languages can also have both properties A and B. É. Kiss (1995:5) provides a list of languages that have been identified as discourse-configurational, some of which are also in Baker's (2003) list of non-configurational languages. These languages come from a range of language families. Probably the best-known example of a discourse-configurational language is Hungarian, where an identificationally focused element must occur in the position immediately preceding the verb. The object in Hungarian typically occurs after the verb, like *kalapot* 'hat' in (498b), but is preposed to precede the verb when interpreted as identificational focus (498a).

Hungarian (É. Kiss 1998:247)
(498)   a.   Mari egy kalapot nézett   ki   magának
                   Mary a   hat.ACC picked   out herself.DAT
                   'it was a hat that Mary picked for herself'

         b.   Mari ki   nézett   magának   egy kalapot
                   Mary out picked   herself.DAT a   hat.ACC
                   'Mary picked for herself a hat'

A similar phenomenon is observed in Aghem, a Grassfields Bantu language. Watters (1979) establishes the Immediate After Verb (IAV) position as the position for focus in Aghem. In the canonical sentence in (499a) the locative *án 'sóm* 'in the farm' is in its canonical sentence-final position. When it is the answer to a question, it is considered the focus of the sentence, and hence it occurs in IAV position (499c). Note

that the question word *ghέ* 'where' (499b) is also in IAV position, as *wh*-words are assumed to be inherently focused. Even the subject in this SVO language should appear in IAV when focused, as illustrated in (500), where the *wh*-word as well as the answer "Inah" must appear after the verb.

Aghem (Watters 1979:147)

(499)   a.   fíl   á   mɔ̀ zɨ́ kɨ́-bέ án   'sóm
                  friends SM P2   eat fufu   in   farm
                  'the friends ate fufu in the farm'

         b.   fíl   á   mɔ̀ zɨ́ ghέ   bέ-'kɔ́?
                  friends SM P2   eat where fufu
                  'where did the friends eat fufu?'

         c.   (fíl   á   mɔ̀ zɨ́) án   'sóm   (bέ-'kɔ́)
                  friends SM P2   eat   in   farm   fufu
                  '(the friends ate fufu) in the farm'

(Watters 1979:144)

(500)   a.   * ndúghɔ́   mɔ̀ ñɨ́ŋ   (nô)?
                     who   P2   run   foc

         b.   à   mɔ̀ ñɨ́ŋ ndúghɔ́?
                  DS   P2   run   who
                  'who ran?'

         c.   à   mɔ̀ ñɨ́ŋ énáʔ?
                  DS   P2   run   Inah
                  '*Inah* ran'

   Aghem could thus be considered a discourse-configurational language. Other Bantu languages that have been reported to display a free word order (and where word order seems to be related to discourse principles) are Xhosa (du Plessis and Visser 1992), Chichewa (Bresnan and Mchombo 1987), Northern Sotho (Zerbian 2006) and also Makhuwa (Stucky 1985). The sentences in (501) exemplify the variability in word order and the influence of the discourse in Northern Sotho. They show that the logical subject and object may precede or follow the verb. As is reflected in the English translation, there is a difference in interpretation between the SVO, VS and OV orders. These word orders, their interpretation, and the subject agreement are further discussed in chapter 4, section 4.3.2.

Northern Sotho (Zerbian 2006:171,58)
(501)   a.   monna o ngwala lengwalo
                   1.man   1 write   5.letter
                   'the man is writing a letter'

         b.   go fihla monna
                   17 arrive 1.man
                   'there arrives a man'

         c.   lengwalo, ke   le ngwad-ile
                   5.letter   1SG 5 write-PAST
                   'the letter, I wrote it'

   Although word order in these Bantu languages is not as strict as in English, for example, it is certainly not as free as the word order in the languages Mithun (1987) describes. She shows that languages like Cayuga (Iroquoian), Ngandi (Australian) and Coos (Oregon) display all possible combinations of subject, verb and object, and do not have any preference regarding word order out of context. For (502) Mithun reports that a Cayuga speaker found all three sentences grammatical, but that there was no preferred reading for the arguments ("it was unclear who beat whom"). Instead, the word order is fully pragmatically based, according to the "relative newsworthiness within the discourse at hand" (Mithun 1987:325). She concludes that these pragmatically based languages do not have a basic word order.

Cayuga (Mithun 1987:286)
(502)   a.   Khyotro:wę́:   Ohswe:ké̜'   ahǫwati:kwéni' (SOV/OSV)
                   Buffalo   Six Nations   they beat them
         b.   ahǫwati:kwéni'   Khyotro:wę́:   Ohswe:ké̜' (VSO/VOS)
                   they beat them   Buffalo   Six Nations
         c.   Ohswe:ké̜'   ahǫwati:kwéni'   Khyotro:wę́: (SVO/OSV)
                   Six Nations   they beat them   Buffalo

   Bantu languages differ from these pragmatically based languages in various aspects. First, the Bantu languages can still be said to have a basic or canonical word order out of context, namely SVO. Apart from the free word order property, which is not as prominent in Bantu (cf. Morimoto 2000), other characteristics of non-configurationality do not apply fully either to Bantu languages: there is no case-marking on nouns,[17] and while the full arguments of the verb can easily be left out, the subject

---

[17] Schadeberg (1986) describes tone cases in Umbundu, and Kavari and Marten (2005) describe a system in Herero which uses four different tone patterns on the object to indicate what they call default case, complement

and object agreement morphology on the verb can be argued to sometimes be pronominal (incorporated pronoun) and sometimes not (grammatical agreement) (Bresnan and Mchombo 1987).[18] As such, the Bantu languages mentioned earlier seem quite discourse-configurational when compared to a language like English, but the picture rapidly changes when comparing to the discourse-configurational languages Mithun (1987) describes and analyses.

Comparing various non-configurational or discourse-configurational languages, the conclusion must be drawn that there is a lot of variation within these languages and that "there is no single non-configurational type" (Pensalfini 2004:393). It seems that word order in languages can be partly determined by discourse, or be partly "configurational". For Makhuwa specifically, Stucky (1985:192) concludes that the language "seems to be about midway between the relatively fixed order of English and the very free order of Warlpiri. This relativity is suggestive of a continuum rather than a clear-cut distinction".

In this light, it is obvious that a division into configurational vs. non-configurational, or into a tripartite division with a third category discourse-configurational, is descriptively inadequate and very unlikely to be valid. Instead, word order could be viewed as a linguistic means used to express both syntactic functions and discourse functions, where it is seldom the case that languages have their word order determined purely by syntactic principles or solely by discourse principles. The continuum Stucky suggests would then not only involve syntax, but discourse as well. All languages are somewhere on the continuum between these factors determining word order, reaching from a high influence of discourse on one end of the continuum to a high influence of syntax on the other. Word order is thus never free, but is always to some extent determined by syntax and/or discourse.

Where a language is on this continuum may be related to the alternative means a language has to express syntactic relations or discourse functions, besides word order. All languages use word order to some extent, but since word order cannot encode all syntactic relations and discourse functions at the same time, languages must have some other means to encode (at least some of) these properties. Languages vary in the means they have available and in the functions that these means can encode. If a language has a broader inventory of means to encode syntactic relations, for example the morphological marking of case and agreement, the word order in that language is more easily used to encode discourse functions. If, on the other hand, a language lacks these alternatives, the word order is used to make clear what the syntactic subject or object is. In that case, that language can resort to other means, such as prosody, for the encoding of the discourse information.

---

case, presentative and vocative. This system seems to be related more to information structure and does not resemble either ergative or accusative case systems.

[18] Discontinuous constituents are also often found in non-configurational languages, but these have not been the focus of the research. See however Morimoto and Mchombo (2004) and Mchombo (2006).

The question posed in the second part of this thesis is: where is Makhuwa on this continuum? Or, more precisely, how do discourse and syntax interact in Makhuwa? In order to answer this question, this chapter first presents the basic notions of information structure (discourse functions in the sentence) and of the Minimalist Programme (a theoretical direction in syntax), which I assume and use in chapters 4 and 5. Then I show how information structure can be combined with minimalist syntax, discussing the cartographic model and an interface account. In the present chapter the discussion is more general and abstract, whereas in the following chapters (4 and 5) the models of syntax and information structure are applied to word order in Makhuwa and the conjoint/disjoint system, respectively. Chapter 6, the conclusion, tries to answer the question about the interaction of syntax and discourse for Makhuwa.

3.2   Information Structure

3.2.1   Information structure, accessibility and salience

The term "information structure" (IS) was first coined by Halliday (1967), to describe the fact that the linguistic and extralinguistic context of a sentence can have an influence on the structure of that sentence. As many linguists have noted (Chafe 1976, Prince 1981, Firbas 1992, Lambrecht 1994, among many others) it is necessary to take into account this context in order to fully understand the formal properties of a sentence. Since many scholars have developed ideas about the functional and formal theory and application of information structure, many different definitions and terms have been used. I give two definitions below, which illustrate various relevant notions in IS.

> de Swart and de Hoop (1995:3)
> … information structuring, that is, presentation of information as old and new. Successful communication requires a balanced presentation of old and new information: too much new information can make it hard to establish the connection with previous discourse and leads to incoherence. Every new sentence in a discourse connects to the previously established context, and, at the same time, adds a new piece of information. Depending on what is new in a given context, the same piece of information can be presented in different ways.

> Foley (1994)
> Information structure is the encoding of the relative salience of the constituents of a clause, especially nominals, and is realised as choices among alternative syntactic arrangements. The IS of a particular clause is determined by the larger sentence or discourse of which it is a part (i.e., its context). The communicative effect of the IS is to foreground certain aspects of the message of the clause, but to background others. The need

> to encode IS is a language universal, but the formal means to do so vary
> widely across the languages of the world.

First, these descriptions make clear that IS has to do with the context of a sentence, the discourse, as already noted above. However, IS is not concerned with the organisation of the discourse itself, but rather with the organisation of a sentence within the discourse (Lambrecht 1994:7). This means that the order and logic of paragraphs in a text, or of turn-taking in a conversation are not relevant for IS, except for that part where the context influences the structure of the sentence. Only the connections between the context and the elements in one sentence are relevant for IS. Broader principles such as the Gricean Maxims (Grice 1975) are thus only indirectly linked to the IS in a sentence.

Second, especially in the area of IS, confusion sometimes arises about what exactly is denoted by certain IS notions. IS is about the text-external world and takes the mental representations of the elements in this world as its primary objects. These concepts (referents and events) are referred to by linguistic expressions. This is an important difference to observe: IS uses *concepts* whereas linguistic structure uses *expressions*. Only the concepts can have a certain IS status, not the expressions. For example, when uttering the sentence "Ali has got malaria" it is not the *word* "Ali" which is familiar to us and apparently has malaria, but it is the *person* Ali. Saying that this sentence is about Ali, or that Ali is the topic of the sentence, means that the referent Ali (or actually its mental representation) is the one being ascribed a certain property. I refer to the things, people and circumstances as the discourse *referents* or *events* or together as *concepts*, and to their linguistic counterparts as *expressions* or (linguistic) *elements*. When discussing the status or value in IS I use "the referent corresponding to element X" most often, but for the sake of brevity I sometimes state that "an element/expression is interpreted as…", by which I still intend the referents the elements correspond to.

Third, IS concerns the *presentation* of a message rather than the *content* of the message. The meaning of a linguistic utterance in terms of lexical and/or propositional content remains constant.[19] However, depending on the speaker's hypothesis about the hearer's state of mind (assumptions, attention), that same meaning may be packaged in different ways. In other words: how a speaker chooses to express a certain meaning depends (partly) on what she thinks is new or old information for the hearer. Vallduví (1993:14) characterises "information packaging" as a "set of instructions with which a speaker directs a hearer to retrieve the information encoded in a sentence and enter it into her/his knowledge store". Only if the speaker adjusts the encoding of the message to the needs of the hearer can fruitful communication take place.

The distinction old vs. new, as put forward in the first definition (de Swart and de Hoop 1995), is one important property in the presentation of information. The second definition (Foley 1994) singles out another important property: relative salience, or the foregrounding and backgrounding of certain aspects. These properties turn out to be very

---

[19] Exhaustive focus, and focus particles like "only" can be claimed to have a truth-conditional semantic impact.

important in information structure in general, and specifically in the model I use in the next chapters.

The relative newness of a piece of content depends on what the hearer already knows. IS is thus based on the speaker's assumptions of the hearer's knowledge and should help the hearer understand what the speaker intends. Yet not all the information a hearer has in her head is taken into account, neither is it coded in the grammar. As Chafe (1976, 1987) notes, the conveying of information not only involves knowledge (long term memory), but also consciousness (short term memory). Since our minds can only focus on very few concepts at a time, only a limited number of concepts can be cognitively "active". Chafe (1987) suggests that a concept can then be in one of three possible activation states: active, semi-active or inactive. A concept is active only for a short while, when it is "lit up" as the centre of consciousness, and then becomes semi-active, which means that it is still in the awareness of the speaker, but more peripheral. After a while, it can get back to the inactive state: equal to most concepts that were unused in the previous discourse.

Concepts can be activated in three ways: by previous mention in the discourse (textually accessible), by the current situation or in general the text-external world (situationally accessible), or by a semantic frame (inferentially accessible) (Lambrecht 1994:99). As an example of the first two possibilities of activation, imagine we have a conversation in which the referent "sailing boat" becomes active in our minds. This could be the case, for example, when you have just told me you went sailing with your boat last weekend ("text"), or when we happen to be sitting at the harbour and a yacht passes by (situation). In both cases the referent is activated in our minds. The situationally accessible referents always include the referents who are present in the current discourse situation (me, you), but also the concepts that are always accessible (to a certain degree) by common knowledge, such as "the moon" and "the train" (Erteschik-Shir 2007). The third possibility, the activation by a semantic frame, happens through the semantic connection with a related concept that is activated. For example, when "pancakes" are mentioned, not only this referent gets activated in the mind of the hearer, but also the syrup and icing sugar she normally puts on her pancake become more activated, because they are in the same semantic frame as the pancakes.

Returning to the three activation states that Chafe suggests, Lambrecht (1994:100) observes that "from the psychological point of view, there is no theoretical upper limit to the number and kinds of cognitive states which [concepts] may have in the course of a conversation". Slioussar (2007) applies this insight in her activation network model. After a concept has been activated it does not immediately switch to be inactive, but the activation will *gradually* decrease, so that at different points in time, concepts have numerous different states of activation. This means that a concept is not "active" or "inactive", but rather that it has "more activation" or "less activation". The higher the amount of activation on a concept, the more accessible it is. All concepts thus move along an *accessibility* scale (cf. the Givenness Hierarchy of Gundel, Hedberg and Zacharski (1993)), where each concept has a value for accessibility. However, this does

not imply that this exact value on the scale is what is encoded in the language. Later in this chapter I explain that it is only the *relative* accessibility with respect to the other elements in the sentence which is reflected in the grammar.

So, the first relevant property is relative accessibility. The second relevant property is what Foley called the relative salience. This property represents the speaker's intentions for further discourse. Just as the current discourse representations reflect what has been mentioned before in the discourse (accessibility), they also keep track of the intended amount of attention for the next part of the discourse. What does the speaker want to highlight? What should be more backgrounded? Before uttering a sentence, new discourse representations have been constructed by the speaker, and the concepts of the next sentence have already lit up in her mind. Thus, whenever a concept is selected to be spoken about, it automatically has a value for accessibility and salience.

How accessible and salient a concept is in the discourse can, for example, be measured in texts. Ariel's accessibility theory (1985, 1990, 2001) accounts for the choice of referential expression used for a referent at a given point in the discourse or text. When a full noun is used, for instance, the referent is more likely to be low in accessibility than when only a pronoun is used, or just a prefix. She lists several factors that influence the accessibility, such as the inherent importance of the referent (e.g., being a participant), the number of times a referent has been mentioned before, the number of referents mentioned between two expressions for the same referent, the cohesive linking within a paragraph, the grammatical role, etc. These are all factors that influence the accessibility of the referent and thus its encoding in the language. Accessibility is probably not only influential in the choice of referential expression, but I assume also in the word order (and possibly other strategies marking information structure).

Whereas the accessibility of a concept can be determined by looking at the previous discourse or text, the salience of a concept is visible in the role the referent plays in the following discourse. Gernsbacher (1989) notes that the way a referent is encoded does not only reflect the current degree of accessibility, but also contribute to the future accessibility status of the referent. This "extra" function corresponds to what is here referred to as salience.

Following Slioussar (2007) I take accessibility and salience to be the aspects of IS that are relevant for grammar. The way in which accessibility and salience are encoded is explained in section 3.4. If these are the notions the grammar needs, then two other notions frequently used in IS can stay within the realm of pragmatics: topic and focus. The difference between accessibility and salience on the one hand, and topic and focus on the other hand, is that the former are properties or *states* of individual referents, and the latter are "pragmatic *relations* established between these [referents] and the propositions in which they play the role of predicates or arguments" (Lambrecht 1994:49). Referents thus have a certain IS status, and on the basis of that status they can have a topic or focus relation to the proposition. For example, a referent can be very accessible and may even be the most accessible of all concepts in the sentence. The

grammar could encode this accessibility by putting the expression corresponding to that referent in a sentence-initial position. The pragmatic relation of this referent to the proposition is then that of "topic". In order to better understand the terms "topic" and "focus" and the functional relations they indicate, I discuss them below.

3.2.2   Topic and focus

An abundance of terms have been proposed to indicate semantic, pragmatic and syntactic properties related to "topic" and "focus" in some way, and there are even more definitions that have been proposed for these terms (see the intricate map in Kruijff-Korbayová and Steedman 2003:254). I do not discuss all of these here, but indicate which terms I use and what I understand by them. For further information and explanation on different kinds of topic and focus and terms or definitions used for them, see Gundel (1999), Gussenhoven (2007) and Krifka and Féry (2008).

Topic
One distinction I would like to clarify is that between "discourse topic" and "sentence topic". The discourse topic can be the issue of debate for a longer stretch of time, or for a larger unit than the sentence (paragraph, text, whole conversation), and it can be more abstract (Reinhart 1981). Sentence topics, on the other hand, can vary for each sentence in the discourse and often correspond to an expression in the sentence (topic expression). For example, within a conversation the discourse topic may remain "making pancakes", but one sentence in the conversation may have the batter as its topic, while other sentences may concern the frying pan or the syrup, and have that as a topic. The study of IS only relates to sentence topics: as already mentioned above, IS is concerned with the organisation of a sentence within the discourse, not with the organisation of the discourse itself.

The topic of a sentence has in the literature been defined as a) that part which is old or presupposed information or b) as that what the sentence "is about" (leaving aside syntactic, prosodic and psychological definitions[20]). Although topic referents are usually associated with presupposed or old information, there is still a certain gradience in the "oldness", as mentioned above. Prince (1981) and Reinhart (1981) show that being discourse-old is neither necessary nor sufficient to function as a topic. Instead, taking pragmatic aboutness as a defining notion, the topic can be viewed as the referent to which the information in the proposition is applied, or the entry under which the information in the proposition should be stored. It is then used as a means in the language to express the ordering and categorising of the information in a discourse. This

---

[20] These could be, respectively, "first position in the sentence" (Halliday 1967), "non-stressed elements" (Chomsky 1971), or "center of speaker's attention" (Schachter 1973), as mentioned by Reinhart (1981:56)

view on the function of a topic is consistent with Chafe's (1976:50) description of topics:[21]

> What topics appear to do is limit the applicability of the main predication to a certain restricted domain. [...] The topic sets a spatial, temporal, or individual framework within which the main predication holds.

Under this definition, a sentence can have more than one topic. The frame within which the proposition should be evaluated and stored can be specified for both space and time, as in the typical phrase at the beginning of a story: "Once upon a time, in a country far, far away…". This time and location set the scene or frame within which all the information that follows in the sentence should be evaluated. Multiple topics are not restricted to adverbial expressions, but they may also be individual referents, as in "Scones, my mother really bakes them the best". The information about the quality of the scones and the baking is then stored under and assessed for both the "scones" and "my mother". More information about multiple topics can be found in chapter 4, sections 4.2.4 and 4.2.5.

Erteschik-Shir (2007), taking more or less the same definition of topic as do Reinhart (1981) and Strawson (1964), specifies topics as the "pivot for truth value assessment" (p.15). The proposition is evaluated within the frame that is set by the topic and it is only within the limits of this topic that a proposition can be judged true or false. Since every sentence is assigned a truth value, every sentence must have a topic, according to Erteschik-Schir (2007:15).[22] That is, every sentence has a *pragmatic* topic, but this is not necessarily overtly realised in every sentence (Gundel 1988). A sentence can thus have a pragmatic topic (a referent/event), but lack a topic expression (a word or phrase). When a sentence lacks a linguistic expression to refer to the topic (the topic expression is dropped), the pragmatic topic is the "here and now". This is referred to as a stage topic (Gundel 1974).

In summary, I take "topic" to be a pragmatic relation between a referent and a proposition. The proposition is evaluated with respect to the topic, or, in other words, the topic restricts the domain in which the proposition is judged true or false and indicates where the information in the proposition should be stored. An important difference is that between the topic referent in the real world (and its mental representations) and the topic expression, which is the linguistic element corresponding to the topic referent. All sentences have at least one topic *referent*, but may have more, while the topic *expression* can be absent.

The seemingly topic-less sentences, which have a stage-topic, are thetic sentences, alternatively called "all-new" utterances. In the literature concerning IS a

---

[21] Chafe (1976) makes a distinction between "what the sentence is about" and "the frame within which the sentence holds", but the way Reinhart (1981) explains "aboutness" unites these definitions.

[22] Deciding on the truth value and topic of direct speech, or a sentence with imperative mood, could be difficult, though.

distinction has often been made between categorical and thetic statements (Kuroda 1972; Sasse 1987, 1996). A categorical statement is a twofold judgement, stating the existence of an entity and then predicating something on it. A thetic statement, on the other hand, is an unstructured judgement expressing only the recognition (or rejection) of an event or a state. Sasse (1987) uses (504) and (503) as typical examples of these two types of judgements. The categorical judgement in (503) first names the entity John and then predicates of him that he is intelligent. On the other hand, the thetic judgement in (504) does not involve the independent recognition of some entity, but simply affirms the state or situation of "raining". It is easy to see how (504) predicates about the "here and now", the stage topic.

(503)   John   is intelligent
         entity   statement

(504)   it is raining
         statement

Thetic sentences are important in this thesis, since they are expressed by the non-canonical word order VS in Makhuwa (505), or a special presentational construction (see chapter 4, section 4.3.2). Thetic sentences are typically used "out of the blue" or at the beginning of stories, i.e., when there is no discourse context yet and when the referents are introduced which will be referred to in the following discourse. All the elements in the sentence can thus be expected to be presented as equally salient: there are no elements that have a specific function as topic or focus. Hence, the term "presentational focus", which has also been used to describe the pragmatic function of the thetic sentence, is controversial and confusing. A thetic sentence can indeed introduce a referent into the discourse, or mention an event taking place, but it does not contain an exclusive focus (as defined below). Instead, the elements in a thetic sentence form the comment to a stage topic.

(505)   e-náá-rúpá   epúla
         9-PRES.DJ-rain   9.rain
         'it is raining!'

Focus
A categorical judgement is two-fold, consisting of the recognition of an entity and a statement about that entity. Categorical sentences express this split linguistically as the topic expression and the comment. The comment is what is assessed relative to the topic. Within the comment a further distinction can be made between the concept that is interpreted as the focus of the sentence and the background. While every sentence has a comment (otherwise there would be no point in saying the sentence), not every sentence needs to have a focus.

Several authors distinguish different types of focus, which may be encoded differently in a language. É. Kiss (1998) shows for Hungarian that there is a difference in interpretation between postverbal objects and immediately preverbal ones. She claims that the postverbal element receives "information focus" and the preverbal element has "identificational" focus, as illustrated in (498) above and (506) below. In (506a) the indirect object Mary has identificational focus, whereas in (506b) it is the new information of the sentence and is not interpreted as identificational focus. Likewise, Hyman and Watters (1984) distinguish between "assertive" and "contrastive" focus. The first type (information or assertive) is the new information the speaker gives without a special background or reference set in mind, for example as the answer to a *wh*-question. The second type (identificational or contrastive) indicates that the concept is selected from a restricted set and that for the rest of the members of that set the proposition does not hold.

(506) a. tegnap este Marinak mutattam be Pétert
last night Mary.DAT introduced.1SG PERF Peter.ACC
'it was to Mary that I introduced Peter last night'

b. tegnap este be mutattam Pétert Marinak
last night PERF introduced.1SG Peter.ACC Mary.DAT
'last night I introduced Peter to Mary'

Makhuwa-Enahara does not seem to mark such a distinction. Instead, the notion of "exclusivity" seems the most relevant. What is marked in Makhuwa is the element that is selected to the exclusion of some alternative. This does not always entail exhaustive identification (although it may) or contrast. The term "contrast" I use to refer to a contrast made explicit in the context, not a contrast with alternatives, as in Rooth's (1996) theory of alternative semantics. My notion of exclusivity is consistent with the basic idea of alternative semantics, which proposes that the meaning of focus is that it evokes possible alternatives for the focused constituent. The referent of the element marked as exclusive is identified as the referent for which the proposition holds, and the proposition does not hold for (at least) some other referent. This is a weak version of exclusivity, and I cannot prove that a strong version (exhaustivity) always holds.

One way in which IS is marked in Makhuwa is in the difference between conjoint and disjoint conjugations. Clear evidence for the marking of an exclusive interpretation in these forms is found in sentences where an expression is modified by the particle "only", which induces an exclusive reading. Such an expression must follow a conjoint verb form (507a) and is ungrammatical with a disjoint verb form (507b). The concept of exclusivity and the conjoint/disjoint alternation are further elaborated in chapter 5.

(507)   a.   CJ   o-lomw-é   ehopa   paáhi
                          1-fish-PERF.CJ 10.fish   only
                          'he caught only fish'

        b.   DJ   * oo-lówá   ehópá   paáhi
                            1.PERF.DJ-fish 10.fish   only
                          int. 'he caught only fish'

Whereas some analyses make use of a contrastive focus type, I do not take contrastiveness to be a type of focus or a function an element can have. This holds for Makhuwa, but I suggest, along the lines of Lambrecht (1994) that an explicit contrast and a contrastive interpretation is in general dependent on context. Lambrecht (1994:290, 291) states that

> the impression of contrastiveness [...] arises from particular inferences
> which we draw on the basis of given conversational contexts. [...]
> Contrastiveness [...] is not a category of grammar but the result of the
> general cognitive processes referred to as 'conversational implicatures'.

This view on contrastiveness is unlike the one Beninca and Poletto (2004) adhere to; they assume a special projection for contrastive focus in the syntactic representation. Most of the cases in Makhuwa for which a contrastive focus reading could be claimed can actually be reanalysed as having an exclusive interpretation, in addition to which the context indicates an explicit contrast.

Furthermore, a contrastive interpretation is found not only with focus referents, but also with topics. The same applies here: there is no specific type of topic called "contrastive topic". Various preverbal elements in Makhuwa, which can be said to have a topic relation to the proposition, can also be contrasted. These different elements can have various syntactic functions. In (508) the subject *oóréera nrímá* 'good ones' is contrasted to *oótákhala nrímá* 'bad ones' in the next sentence; in (509) and (510) the contrasted elements are the left-dislocated objects *olávíláví* 'trick' and *ntsúwáki* 'toothbrush', respectively. Adjuncts such as the temporal adverbs *otháná* 'during the day' and *ohíyú* 'in the evening' in (511) can also be contrasted. These elements are analysed occupying different preverbal positions, and they are only interpreted as contrastive when a following or preceding phrase indicates the contrast.

(508)   oóréera   nrímá   a-n-khálá   warattá-ní
        2.good.REL   3.spirit 2-PRES.CJ-stay 16.lagoon-LOC

        oótákhala   nrímá   a-n-khálá   nshawóro
        2.ugly.REL   3.spirit 2-PRES.CJ-stay 18.bathroom
        'the good ones (frogs) live in the lagoon, the bad ones live in the bathroom'

(509)   <u>oláví láví</u>   woo-phwánya   só   apátthány' aáwé   khaá-vo (H7.14)
          14.trick   2SG.PERF.DJ-meet   only   2.friend   2.POSS.1 NEG.2-LOC
          'the trick you've found, just its friend is not here'

(510)   nyú   moo-thúm-átsa   tsootéene
          2PL.PRO   2PL.PERF.DJ-buy-PLUR   10.all

          masi <u>ntsúwáki</u>   khu-thum-ále
          but   3.toothbrush NEG.2SG-buy-PERF.DJ
          'you bought everything, but you didn't buy a toothbrush'

(511)   <u>otháná</u>   ni-m-váréla   ntsuwá
          17.daylight   5-PRES -burn.REL 5.sun.PL

          <u>ohíyú</u>   o-n-aáryá   mweerí
          17.night   3-PRES-shine.REL 3.moon.PL
          'during the day the sun shines, at night the moon shines'

      In summary, focus is a relation between a referent/event and a proposition, like topic. In Makhuwa it evokes alternatives for the focused concept, and exclusivity is the property encoded in the grammar. Conversely, contrastiveness is dependent on the context and is not as such grammatically expressed.
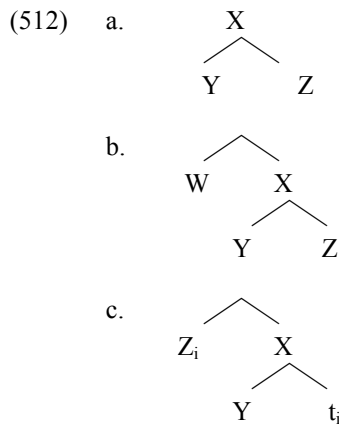
      In this section the most important concepts of information structure have been discussed. The pragmatic *relations* topic and focus and the various definitions and associations have been clarified. Accessibility, salience and exclusivity are identified as relevant *properties* of referents in the grammatical encoding of IS. The question posed earlier can now be narrowed down: how do *these* properties of IS interact with the syntax in Makhuwa? The next section introduces the basic ideas and operations in the minimalist model of syntax, without referring to IS. In section 4 the possible combination of the two (IS and syntax) in a model is examined.

3.3   **Minimalist syntax**

Generative syntax has always been concerned with the design of the human language faculty. Language, Chomsky (1966) argues, is a separate cognitive system that interacts with other cognitive systems (see also Jackendoff 1997). It allows us to formulate and thereby structure our thinking, as well as communicate with other human beings. The structure-building part of the language system (syntax) can be studied independently of the lexical meaning or context (Chomsky 1957). For example, one can still judge the grammaticality of a nonsensical sentence, as in Chomsky's now famous sentence "colourless green ideas sleep furiously": a perfectly grammatical sentence, that does not have a (logical) meaning. In the last decades the hypothesis has been examined that syntax is a perfect and economical system. The question posed under this hypothesis is the following: suppose that the syntax has minimal means to structure meaning: how far

can we get in explaining the properties of linguistic constructions? This is the line of research the Minimalist Programme follows.

The input for the structures to be built is the lexical items. These are first selected from the mental lexicon to form the exhaustive collection of elements the sentence will consist of, which is called the Numeration. What syntax does with these lexical elements is to combine them to form new constituents. This happens by applying the operation "Merge", which is the only operation postulated in current Minimalist syntax (Chomsky 1995, 2004, 2005). This operation takes two linguistic elements and combines them, thereby creating a new unit (like Y and Z are merged to form X in (512a)). Merging another element to that new unit extends the derivation and forms another unit. To this new unit another element can be merged and so on. However, only one unit is added at a time, and hence Merge creates binary branching structures. When extending the derivation by one element, this element can be either new from the lexicon, like W in (512b) or from the derivation itself, i.e., an element that has already been merged before, like Z in (512c). The first type of Merge is referred to as External Merge and the second type is called Internal Merge. Since in Internal Merge an element leaves its original position in the derivation and ends up in another position (leaving a trace t), this operation is also referred to as "Move".[23] I use the terms "move" and "movement" to refer to the operation Internal Merge.

(512)   a.   X
                      /\
                    Y   Z

         b.   /\
                    W   X
                        /\
                      Y   Z

         c.   /\
                  Z_i X
                        /\
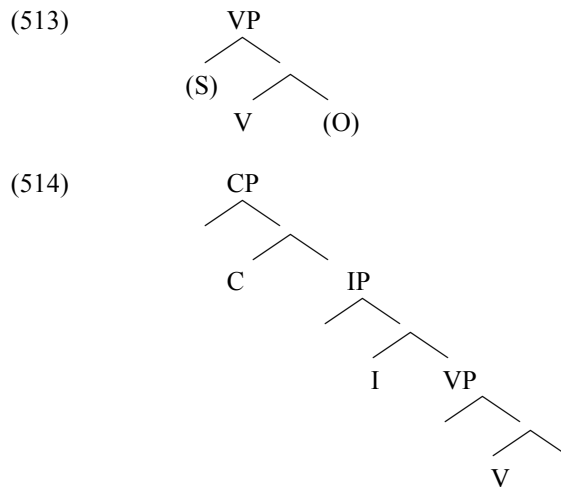   Y t_i

Properties of lexical items can be projected to a maximal projection, of which the lexical item is the head. In (513) the maximal projection is VP and the head is V. The element to which a head is first merged is its complement (O in (513)), and the position directly under the maximal projection is the specifier (S in (513)). On top of such a maximal projection another projection is built, etc. The derivation of a sentence proceeds

---

[23] One can also read here that an element is copied and merged in another position, but since for my purposes the effect is the same, I will use the term movement.

from the lexical/thematic domain (the verb phrase, VP) to the inflectional domain (the inflectional phrase, IP), and on top of that the complementiser domain is derived (complementiser phrase, CP), as in the tree structure in (514). The CP is typically analysed as the domain where sentence type (relative, embedded, question) and pragmatic interpretation are encoded. This is the most relevant domain for notions of IS.

The projections in IP and CP are functional projections, and their heads are active in establishing syntactic relations. Still, the inflectional domain is more related to the lexical domain, since in the inflectional domain lexical elements can be licensed. The postions in which arguments are merged and/or licensed are called A(rgument) positions, and VP and IP together are thus traditionally called the A domain. The complementiser domain on the other hand is the A-bar domain. When an element is moved, it can be moved to an A position, or and A-bar position. The latter type of movement is intended in this thesis when an element is said to be (left- or right-) dislocated.

(513)

```
          VP
         /\
      (S)   /\
          V   (O)
```

(514)

```
          CP
         /\
        C   \
           IP
          /\
         I   VP
            /\
           V
```

By combining linguistic elements to form larger units, the syntax creates relations and dependencies between these elements. One such relation is c-command: a node c-commands all other nodes under the first branching node up that it does not dominate. In (515), for example, going upwards from W the first branching node is V, and the other node under it is X, which contains Y and Z. W thus c-commands X, and with that also Y and Z. W does not c-command V, T or U, neither can it c-command itself. Node V only c-commands U in this structure and cannot c-command W or X, because it dominates them. In this thesis the c-command relation is used in the interface rule referring to the conjoint verb form, proposed in chapter 5, section 5.4.2.

(515)

```
        T
       / \
      U   V
         / \
        W   X
           / \
          Y   Z
```
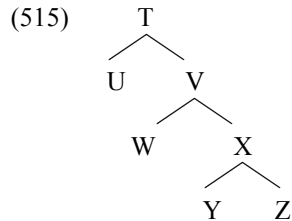
The relation between W and Y in (515) can also be called asymmetric c-command, because W c-commands Y, but Y does not c-command W. This is unlike W and X, for example, which c-command each other. Asymmetric c-command is used in defining the relation between the hierarchical syntactic derivation and the linear word order. The hierarchical structures correspond to linear word order as formulated in Kayne's (1994) Linear Correspondence Axiom. According to this axiom, if an element asymmetrically c-commands another, it will precede this other element. Thus, the spell-out of the structure in (515) would put U before W (and X), and W before Y and Z.

Another relation between syntactic elements, which is often marked morphologically, is Agree. When two elements agree, they share certain features. These can be present on either one of them or on both. Such features include phi-features such as person, number, gender and case features. The overt expression of an Agree relation can, for example, be a prefix on the verb, such as the subject marker in the Bantu languages. In minimalist syntax an Agree relation is initiated by a head –the probe– that searches in the derivation that has been built up so far (the c-command domain). When it encounters an element that has the feature specification that the probe is searching for – the goal–, an Agree relation is established between the probe and the goal. A distinction is often made between interpretable and uninterpretable features. Number and person, for example, are interpretable features of a noun phrase, because they play a role in the interpretation of the noun phrase, but the same features are uninterpretable on a grammatical agreement prefix, because it does not play a role in the semantics of the sentence. The checking of these uninterpretable features by matching with interpretable features is thus like fitting the pieces of a jigsaw puzzle. Feature sharing in an Agree relation is for example the case with object marking in Makhuwa. The head AgrO (for object agreement) is merged into the derivation, in which the verb, the subject and the object are already present. This head then searches its complement and finds as its goal the object. Probe and goal agree, and the AgrO head now displays the features of the object: in (516) it is specified as class 1 and spelled out as the object marker *-mʹ-*.

(516)   nthíyáná   o-ni-ḿ̲-mána   n̲l̲o̲p̲w̲á̲n̲a̲
         1.woman   1-PRES.CJ-1̲-hit 1̲.man
         'the woman hits the man'

For the Bantu languages, it has been argued by Carstens (2005) and Buell (2005) that when the head AgrS (for subject agreement) establishes an Agree relation, the goal should always be moved to the specifier of AgrSP.[24] For example, the subject marker on the verb in (517) is *tsi-*, which agrees in class with the subject *minépa* 'ghosts' (both class 4), and accordingly the subject moves up and precedes the verb. The position in which the moved phrase ends up is the specifier of the probing head. If in (517) the probing head is AgrS, the subject is moved to specAgrSP. The position from which an element is moved is indicated by "t" and an index.

(517) masi <u>minépa</u>$_i$ <u>tsi</u>-ńní-wá-aka t$_i$ vá (H12.52)
       but <u>4</u>.ghosts <u>4</u>-HAB-come-DUR 16.PRO
       'but the ghosts use to come here'

Although many generative syntacticians have tried to find the reason behind this movement to the subject position, so far it has only been described as a principle: the Extended Projection Principle (EPP): "every sentence needs to have a subject". Suggesting that Agree is linked to EPP or that AgrS "has an EPP-feature" comes down to saying that the element determining subject agreement on the verb must be moved to the position above the verb. While I would like to provide a more satisfactory explanation for the necessity of movement with subject agreement in Makhuwa, this issue is far too intricate, and I take it as a given.

     Agreement can thus be one circumstance under which elements undergo movement in the derivation. Otherwise, movement can only occur if it has interpretational effects, or, as Chomsky (2005:7) puts it: "Internal merge yields discourse-related properties such as old information and specificity, along with scopal effects." This is where the general idea of an independent syntactic module interacting with other cognitive modules becomes interesting. In essence, the computational system of the syntax is very simple: only Merge is used. Although there are two versions of Merge, external and internal (move), the system is still very limited. This makes its output in principle *un*limited, as the operation can basically combine any given linguistic object with another, creating all possible derivations. These derivations, as the output of the computational system, should be legible at the interface with the other cognitive modules, or at least the conceptual-intentional interface (C-I) and the sensori-motor interface (S-M). The C-I interface checks the interpretation of the sentence and the S-M interface instructs the speech organs to pronounce the sentence. The syntax must thus make sure that whatever structure it derives has the right form and interpretation at the interfaces. As such, these interfaces form restrictions on the derivations that the computational system derives by applying Merge. The way in which a certain

---

[24] Baker (2008) proposes that languages are parameterised with respect to the direction of agreement. In Bantu languages the subject agreement would then always be with an element higher in the tree and not in the c-command domain.

configuration with a certain interpretation is filtered or matched contitutes an interesting issue, which is treated in the next section.

3.4   Combining IS and syntax

There are many different ways to combine IS and syntax. In multi-level models of grammar, such as Lexical Functional Grammar, IS can easily be integrated as a separate level. This level is then matched with other levels like argument structure and phonological structure. Erteschik-Shir (2007) compactly sketches the various models of grammar and how they could incorporate IS (see also Stucky's (1985) analysis of Makhuwa in Phrase Structure Grammar). In this section I discuss a cartographic model and a configurational interface model and discuss their advantages and disadvantages. These models, adapted with rules specific for Makhuwa, are applied to the Makhuwa data in chapters 4 and 5.

3.4.1   *Cartographic model*

One of the difficulties in combining IS and syntax is the fact that IS uses abstract scalar notions (an increasing or decreasing amount of activation on a concept), whereas grammar has no gradient means, but uses discrete values, such as singular/plural. Slioussar (2007:11) describes two strategies to encode IS: as categorical labels, or as relational notions. The first is pursued in the cartographic model of grammar, as put forward by Rizzi (1997). In this model a certain interpretation is realised as a projection in the left periphery of the sentence, the extended CP domain. Rizzi proposed two topic projections (TopP) and a Focus projection (FocP), and later works have proposed even more fine-grained distinctions and projections related to pragmatic interpretation (cf. Beninca and Poletto 2004). In this way, an explicit map is formed of the projections in the left periphery of the sentence, hence the name. The idea is that an element only receives a certain interpretation when it is in the correct position, that is, when it has checked the features of the relevant head and moved to the specifier of that projection.

For example, a focused element can only receive this focused interpretation when the uninterpretable focus feature of the Foc head is checked, and the focused element has moved to the specifier of FocP. This implies that lexical items do not only have phi-features such as person and number, but can also receive an extra feature, such as [foc] for focus or [top] for topical elements. The head of the TopP or FocP has an uninterpretable feature [top] or [foc] and probes down to find an item with a matching feature, the probe. The features are checked and the goal is moved to the specifier. Movement is still dependent on features here, and the checking of uninterpretational features makes sure that the derivation passes at the interfaces. In (518) the head Foc is the probe, and the XP marked with a focus feature [foc] is the goal. The two agree, and the goal moves to the specifier of the probe, specFocP.

(518)

FocP

Foc[foc] IP

Agree I VP

V XP[foc]

FocP

XP[foc]

Foc[foc] IP

Move I VP

V
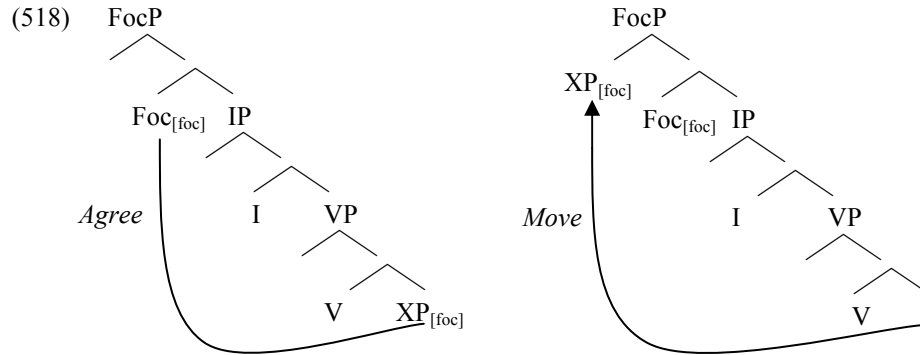
This analysis works for IS and syntax in many (European) languages and is especially well demonstrated for Italian (Rizzi 1997, Frascarelli 1997, 2000, 2004, Beninca and Poletto 2004) and Hungarian (Brody 1990, 1995, Horvath 1995, É. Kiss 2007). The cartographic approach has also been applied in the analysis of IS and syntax in several African languages and language families, such as Chadic (Tuller 1992), Kirundi (Ndayiragije 1999), Kîîtharaka (Muriungi and Abels 2006), Kikuyu (Schwarz 2007), and Kwa and Bantu (Aboh 2007b). However, there are three weaknesses that make the model less attractive in general and for Makhuwa specifically.

First, adding features to lexical elements after these have been retrieved from the lexicon violates Chomsky's (1995:228) Inclusiveness Principle, according to which all the features in a syntactic derivation must be derivable from its lexical units:

> a "perfect language" should meet the condition of inclusiveness: any structure formed by the computation […] is constituted of elements already present in the lexical items selected for N [numeration]; no new objects are added in the course of the computation apart from rearrangements of lexical properties.

Unlike phi-features, which are inherent properties of each lexical element, a focus or topic feature is not always a property of a lexical item. These features have to be added after a lexical element has been retrieved from the lexicon, and the information added by the features is thus not linked to a lexical entry. Focus and topic features therefore violate the Inclusiveness Principle. Erteschik-Shir and Strahov (2003) also use topic and focus features and admit that these (and only these) features violate the Inclusiveness Principle. In their model, the topic and focus features differ from other features in that they are assigned after narrow syntax, after which they are checked at P(honological) syntax. P-syntax does not use hierarchical structure, since the narrow syntax has been closed off. Further operations needed to check the features in P-syntax may include

changing the word order, as in scrambling. However, these word order changes "after syntax" are quite problematic, since they can destroy the syntactic relations built up in the narrow syntax, while at the same time changing the interpretation of the sentence: an unwanted result. Their model of syntax thus faces more problems than just Inclusiveness.

Aboh (2007a) treats the problem of Inclusiveness in his analysis of focus structures in Kwa languages. He argues that information structural features must be present in the Numeration (the set of elements selected for a sentence) and the lexicon and thus do not violate inclusiveness. Similarly to case and phi-features, core features of IS are introduced in the numeration, Aboh proposes. An important argument for his hypothesis comes from the language Gungbe, which uses focus particles. Since speakers must acquire these discourse items, they must be in the lexicon. Another argument is in the comparison between *wh*-features and focus features: if an interrogative feature is syntactic, why would its counterpart focus not be? Inclusiveness should thus not form a problem in this analysis, since IS features are in the Numeration. Nevertheless, there are other objections to the use of topic and focus features and of corresponding projections.

A second weakness is found in the answer to Aboh's question on the difference between a *wh*-feature and a focus feature. A fundamental interpretational problem for discourse features is that the notions "focus" and "topic" are relational (Jackendoff 1972, Lambrecht 1994), but a feature on a syntactic element is not. If a constituent is focused, then the rest of the comment is backgrounded, and in the same way a constituent is never a topic by itself but always the topic of a proposition. Topic and focus encode the information structure of two parts in a sentence relative to each other. It will thus always be problematic to label a syntactic element as topic depending on the checking of a feature but independent of the rest of the sentence or context. The relational nature of topic and focus is easier to implement in a linguistic theory if these notions are understood to be pragmatic relations, which are not directly encoded in the syntax.

Third, in the cartographic analysis an element always moves only to get a certain interpretation itself. However, there are cases where an element moves in order to *not* get the interpretation associated with the original position, or in order for another element to get a certain interpretation. One example from Makhuwa is the VS order, which is discussed more extensively in chapter 4, section 4.3.2. In the derivation of the VS order, the verb moves not to receive or check a certain interpretation for itself but so that the subject does *not* get a topical interpretation, which would be the case if the subject were preverbal in SV order. This movement for negative or altruistic reasons cannot be explained in a theory that makes use of interpretational features.

An additional problem for a cartographic model is the conjoint/disjoint (CJ/DJ) alternation (see chapter 2, section 2.6.5 for a description of the verb forms in Makhuwa and chapter 5 for an analysis). Various southern Bantu languages display this alternation, but in some languages the choice for the one or the other verb form seems to be largely dependent on the interpretation of the element immediately following the verb (exclusive focus or neutral, as in Makhuwa), whereas in others the form of the verb is more determined by constituency (phrase-final or not, as in Zulu or Sotho). In the latter, the

constituency-dependent type, the verb takes its CJ form when the verb is not phrase-final and some element still follows (Buell 2006). This is the case when, for instance, the object is in situ and has not been dislocated. However, if the object is not in situ and does not immediately follow the verb, the verb is phrase-final and takes its DJ form. Since it can easily be the case that the object is left-dislocated after the inflectional part of the verb is derived, the choice for a CJ or DJ verb form and the corresponding morphology of the verb (TAM affixes etc.) can only be determined if the whole derivation and surface representation is taken into consideration. Therefore a filter is needed anyway to determine the morphological form of the verb.

In summary, although the cartographic model can account for certain interpretational effects and word orders, the origin of the syntactic IS features is unclear, and the encoding of relational notions is problematic. Furthermore, movement in order to escape a certain interpretation cannot be accounted for in the cartographic model.

3.4.2 *Interface model*

Slioussar's (2007) configurational IS model
The whole configuration (or representation) of a sentence is found to be relevant in the combination of IS and grammar.[25] What receives an interpretation is not a particular element with an absolute feature, or a particular position, but rather the configuration that the syntax creates. This can be implemented in a configurational model of IS and word order. Earlier configurational models were mainly based on prosody, suggesting that the position of the sentence stress influences or determines the word order (e.g., Szendröi 2003). Unfortunately, these models are very hard to apply to languages that do not use stress as a primary indication of focus or that do not have stress at all. In a configurational model that does not assume a direct influence of stress or prosody on the derivation, the IS is encoded in the final hierarchical relations between the constituents in a sentence. These relations are interpreted at the interfaces according to universal and language-specific conditions, constraints and/or rules.

The most important advantage of such a configurational model is that it allows for the encoding of relative patterns. In a configurational model the grammar does not translate the continuum of possible accessibility values, for example, to a limited number of categories (labeled "accessible", "semi-accessible", or "inaccessible"), nor does it mark the absolute value for accessibility of a concept (say, 64% accessible). Instead, the grammar indicates whether a concept is to be interpreted as more or less accessible than another concept. For example in OV word order it is not the interpretation of the preverbal accessible object or the verb per se, but rather their combination and their status relative to each other which is encoded and interpreted. The encoding of such relative properties is the basis for Slioussar's (2007) configurational IS model, in which

---

[25] I use the term "configurational" here to refer to the *model* (not a language), in the sense that the model relates to the whole representation or configuration of the sentence, and not to the movements itself, to particular stages in the derivation, or to specific projections.

only two notions of IS are encoded in the grammar: the relative accessibility and the relative salience of each referent or event.

In this model, each concept has a value on the accessibility scale as well as the salience scale. These values are dependent on discourse representations and they change along with the development of the discourse. This can be seen as an activation network, as was briefly explained earlier in this chapter. The discourse representations determine the status of each concept in the sentence. Whenever a sentence is uttered, the discourse representations are updated, and these new representations form the input for the next sentence. In this way, the concepts corresponding to the linguistic elements in a sentence all have a specific value for accessibility and salience. The grammar can encode these values in the order of the linguistic elements, for example. The way syntax organises these elements with respect to each other (the derivation) should be in accordance with the interface rules, which make reference to both the hierarchical syntactic relations and the IS values. The interface rules thus restrict the grammatical derivations and interpretations, and function as a filter to derivations made in the syntax. The interface rule Slioussar (2007) proposes for Russian scrambling is given in (519).

(519)   If X is (re)merged above Y, the discourse entity corresponding to X is at least as accessible and at most as salient as the one corresponding to Y. If there are no independent reasons to remerge X above Y, the discourse entity corresponding to X is more accessible and less salient than the one corresponding to Y.

With this rule Slioussar can explain the relative order of two objects and their interpretations in double object constructions in Russian. In the non-scrambled word order S V IO DO, the IO is at least as accessible and at most as salient as the DO. If the DO precedes the IO, as in the scrambled word order (S V DO IO), the DO must be more accessible and/or less salient. Since the movement of the DO over the IO is not related to agreement, it must be motivated by the need to obey the interface rule and have an effect on the interpretation. In (520) the DO *medvežonka* 'bear cub' is given in the context (provided between brackets), and hence it is more accessible than the IO *cirku* 'the circus'. According to the rule, the element corresponding to the more accessible referent (DO) must precede the element corresponding to the less accessible and more salient referent (IO), which is indeed the case.

Russian (Slioussar 2007:183, adapted)
(520)   (And Umka (bear cub) ended up here by accident.)

        Sergej.Šojgu   podaril   medvežonka   cirku
        Sergej.Shoygu.NOM   gave   bear.cub.ACC   circus.DAT
        'Sergey Shoygu presented the bear cub to the circus'

Because Slioussar's model is configurational, the referents referred to in a sentence are interpreted with respect to each other independent of whether they have moved or remain in their original position. Configurations rather than movements are assessed. For the double object constructions in Russian this implies that the higher element is interpreted as more accessible not only if that is a moved element, such as the DO in (520), but also if no movement has taken place and both objects are *in situ* (IO DO). The advantage of interpreting any order, whether scrambled or not, and the possibility of encoding relative notions are the main reasons for Slioussar to develop her configurational interface model for Russian word order.

On the technical side of such an approach, Slioussar takes Chomsky's (2001, 2004, 2005) Phase theory as a basis, but departs from it in several ways. One modification is in the "right position" and interpretation of each element. According to Chomsky, the correct interpretation of each element at the interfaces is determined by the final position it reaches. In the cartographic approach this is a fixed position in the hierarchy, but Slioussar stresses that in her model the correct position for a certain interpretation is the final position *relative* to other elements. Another modification is that Slioussar assumes that movement is not separated from Agree *per se*. There are two different motivations for movement in her model: movement may occur if there is an agreement relation where features are checked (as in *wh*-movement) or if the resulting word order has interpretational effects (differences in scope or IS),[26] as also explained in the previous section.

Although her model does not specifically depend on Phase theory, Slioussar uses one of its mechanisms for IS-related movement. In Phase theory, the maximal projections vP and CP are assumed to be phases in the derivation, and the complements of the phase heads are sent off to be spelled out directly after the phase is completed. Only the elements at the edge of a phase remain visible after closure of the phase, but the other elements are no longer accessible. In order for elements to be moved to the edge of a phase, all lexical items that enter the computation have an edge feature (EF). Futhermore, the phase-heads v and C also have an EF (Chomsky 2005). The EFs on the phase-heads are somehow special, since they can attract constituents in the clause to their specifiers. Thus it seems that there are two different kinds of EFs: those that can attract and those that cannot, but the distinction is not discussed further by Chomsky or Slioussar. The most important aspect of EFs is that they do not involve feature-matching, which is why Slioussar's model uses EFs for the "free" reordering. Any element can thus move to the specifier of the attracting head with an EF, as long as the interpretation at the interfaces is correct.

---

[26] Interpretation-related movement has also been referred to as "free movement" or "free internal merge". However, since there is a clear motivation behind this movement (namely, a difference in interpretation), it does not seem appropriate to call it "free" (just as "free word order" is not really free).

Unlike Chomsky, Slioussar assumes that a number of other heads also carry an attracting EF, namely adverbs, V and T.[27] In the summary she states that all projecting heads have an EF. I briefly repeat her arguments for this point of view here. The projecting heads with EFs are heads that can attract constituents for non-IS-related reasons. Slioussar finds empirical evidence from Russian that these heads also participate in IS-related movement, for example in the reordering of objects within the VP. In this example, the head V has the DO as its complement and the IO in its specifier. V can attract the DO and move it to a second specifier if the DO is more accessible and less salient than the IO.[28] Another example is the EF of T in VS word order. T always needs to have a nominative subject or expletive in its specifier in Russian (EPP), but when T attracts its complement to a second specifier, the subject is sentence-final and is interpreted as the most salient and least accessible.

The adverbs must also have an attracting EF, since they can also be involved in scrambling. In this model, adverbs are not adjuncts, but they form their own projections. Because Slioussar assumes a strict hierarchy of adverbs (Cinque 1999), reordering of a higher adverb and the verb or an object only happen by moving the lower verb and/or object over the higher adverb. The sentence in (521a) represents the neutral word order, with the verb and object following the adverbs. In (521b) the complement of the adverb *medlenno* 'slowly', the predicate *est kašu* 'eats porridge', is moved for interpretational reasons: the adverb is now interpreted as the most salient element in the sentence. This IS-related movement around the adverb suggests that the adverb also has an attracting EF, according to Slioussar.

(521)   a.   ėtot.mal'čik   vsegda   medlenno   est   kašu
               this.boy.NOM   always   slowly   eats   porridge.ACC

        b.   ėtot.mal'čik   vsegda   est   kašu   medlenno
               this.boy.NOM   always   eats   porridge.ACC   slowly

If these heads, which are not at the edge of a phase, are said to carry an attracting EF, the feature is no longer a real "edge feature" and has become a technical way to state that anything around these heads can reorder without feature checking. The only real limitation to this movement is still at the interface, because reordering is only licensed if the resulting hierarchy is interpreted accordingly. If that is the case, the only mechanism needed for IS-related movement is Merge, plus the filter at the interface. This filter is needed to assess derivations in any variant of the model, whether mediated by EFs or simply by performing the operation Internal Merge.

---

[27] With the assumption of attracting EFs on these heads, it is clear that Slioussar does not assume feature inheritance of T from C, or parallel attraction and movement to these phrases, which violates binary Merge.
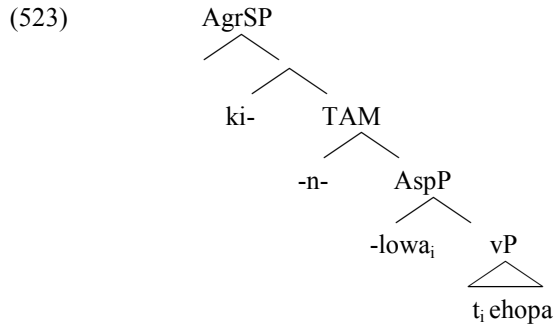
[28] Indeed, in Chomsky (2005) and Slioussar (2007) there is in principle no limit to the number of specifiers a head can have.

Concluding, reordering happens by means of movement, which could be brought about by EFs. There are two types of movement in Slioussar's model. The first is agreement-related movement, which is restricted by feature matching. The interpretational effects are a result of the Agree relation and the sharing of a categorical feature (not the movement). The second is IS-related movement, which is restricted by the interface rules. In this type the movement and the resulting difference in word order causes interpretational effects, according to the interface rules. Unlike the features in the cartographic model, the interface rule can encode relative notions. In either type of movement is the interpretation dependent on the position of the elements in the sentence relative to each other.

Interface model in Makhuwa

Slioussar's study demonstrates that in Russian scrambling, the objects and adverbs are ordered relative to *each other*, keeping in line with the interface rule. In Makhuwa the elements in a sentence are also ordered according to their relative accessibility and salience, but here the IS values are evaluated relative to the *verb*. What appears to be more important in Makhuwa is the placement in the pre- or postverbal domain rather than the mutual ordering of any two elements. Gundel (1988) notes that the verb seems to have a topic-demarcating function in SVO languages: only the subject in an SVO sentence has the topic function. This is often harder to determine in an SOV language, where topic markers are used more often to limit the topic function to the subject. Apart from this crosslinguistic tendency to mark the topic domain, I can think of two possible reasons why the verb would play such a central role, and function more or less as the pivot of the sentence. The first is the syntactic structure of the verb. As I explain in chapter 4, section 4.1, I take the verb to be a composition of in-situ inflectional heads and a verb stem, rather than one complex head. The verb moves to a position just above the vP (say, AspP), and the prefixes for negation, subject agreement, tense, aspect and mood are all morphological heads spelled out in their base position. The whole composition undergoes morphological or phonological merger to form one word. This is exemplified in (523), the tree structure of (522): the verb stem -*lowa* 'to fish' only moves to Asp, not to TAM and AgrS. These heads are filled by the subject marker *ki-* and the present tense marker -*n*-, respectively.

(522)   ki-n-lówá   ehopá
         1SG-PRES.CJ-fish 9.fish
         'I am catching fish'

(523)   AgrSP
                  ⟋⟍
              ki-   TAM
                       ⟋⟍
                   -n-   AspP
                           ⟋⟍
                    -lowa$_i$ vP
                                ⟋⟍
                             t$_i$ ehopa

If the verbal word is actually a phonologically joined sequence of morphemes, its position in the derivation is fixed, and moving the verb would require movement of the whole chunk, a phrasal remnant; unlike moving one head when the verb has head-moved to AgrS or T, as in French, for example. When the position of the verb is less flexible, the sentence is naturally divided into a domain preceding it and a domain following it.

A second hint for why the elements are ordered around the verb is found in diachronic processes. Givón (1976) proposes that the subject and object markers on the verb historically started out as pronouns, used for reference to topics. These then became cliticised to the verb, were reanalysed as anaphoric/pronominal subject and object markers, and then grammaticalised to grammatical agreement markers. In the stage where the prefixes were always used pronominally, the whole argument structure of the verb was expressed on that verb by means of the prefixes; other elements in the sentence were in a sense optional. Even synchronically in a language such as Makhuwa, where the object markers are grammaticalised and function as purely grammatical agreement markers, the verb is still the necessary and central part of the sentence, and other phrases find their positions around it.

In Makhuwa, words in a sentence appear to be ordered topic>comment in pragmatic terms, or, in other words, elements preceding the verb are more accessible and less salient than the verb, and the elements following it are less accessible and more salient, or equal to the verb in accessibility and salience. Because word order is in such a way dependent on the relative notions of IS, I chose Slioussar's model to account for the relation between word order and IS in Makhuwa. However, the tendency in Makhuwa to take the verb as the nucleus of the sentence and to encode the accessibility and salience relative to the verb necessitates adaptation of the interface rule in (519). The interface rule for Makhuwa concerning interpretation of pre- and postverbal elements is introduced and discussed in chapter 4, after presenting more elaborate data on the properties of the elements in the pre- and postverbal domains.

Another characteristic of the grammar of Makhuwa-Enahara is the IAV position, which is connected to the conjoint verb form. This is also a position relative to the verb, but the interpretation of the element in that position is not relative to the other elements

in the sentence; it is interpreted as exclusive. As such, its position and interpretation can be captured in a model using a feature, but in chapter 5 I show that it is also possible to design an interface rule that accounts for the interpretation of the element directly following the conjoint verb form. Before discussing this second interface rule, chapter 5 first provides more information on the conjoint/disjoint alternation.

There is one other predication that the interface model makes. As mentioned, in this model movement can take place for reasons of agreement or for interpretational motivations. The interpretational motivations can of course be related to IS, as formulated by the need to comply with interface rules such as the one in (519). Apart from the IS-related movement, movement can also take place for scope effects. These are not so much IS related, neither are they induced by agreement. Slioussar shows that Russian indeed makes use of non-agreement related movement to obtain a certain scope (2007:130), and the same can be demonstrated for Makhuwa. In an affirmative VS sentence the postverbal subject may not be modified by "only" (524a). Such a subject would have to appear in a cleft or copular sentence, as in (524b). See also chapter 5, section 5.2.5. However, in a negative conjugation the VS order is the only way to derive the scope not>only, and then the VS order is allowed: the verb is moved to a position above the subject in order to obtain the desired scope effect (525). While the cleft is a good alternative for the affirmative conjugation (524b), in a negative cleft the subject c-commands the negative relative verb and still gets the other scope: only>not (526). In order to achieve the right scope, the negative verb must move over the subject, creating a VS order. Movement in Makhuwa can thus be motivated by agreement, IS interpretation or scopal interpretation, as Slioussar also shows for Russian.

(524) a. * oo-vár-íya   latáráw' uúlé   paáhi
           1.PERF.DJ-grab-PASS 1.thief 1.DEM.III only
           int. 'only that thief was caught'

       b. o-var-iy-alé   lataraw' uúlé   paáhi
           1-grab-PASS-PERF.REL 1.thief.PL 1.DEM.III only
           'only that thief was caught'
           lit: '(the one) who was caught was only that thief'

(525) kha-ń-sómá   anámwáne   paáhi
       NEG.2-PRES-read.DJ   2.children only
       'not only childern study' (parents study as well)

(526) Cicica   paáhí o-hi-´m-wéha   efíílíme
       1.Cicica.PL only 1-NEG-PRES-look.REL 9.film
       'only Cicica doesn't watch the film'
       lit: 'it is only Cicica who doesn't watch the film'

Implementation of a configurational IS model
The universal and language specific interface rules can be implemented and elaborated in roughly two different models of the computational system and its interactions: an evaluational model and a derivational model. In the evaluational model the syntax freely creates derivations, which are then checked for their interpretation by the interface rule. At this interface, only the optimal combination of form and interpretation comes through the filter. In order to let the interface rule work as a filter, more than one derivation is formed by the narrow syntax. Although the number of syntactic derivations can in some way be limited by syntactic conditions on the generation of such derivations, there will always be overgeneration: several sentences are generated in order for one to be selected.

This generation-and-selection process is of course best known from Optimality Theory (OT). In OT, various universal constraints are ordered differently in each language, in such a way that the filter of constraints selects the sentence that is optimal in that language. In recent attempts to move the interpretational component from the narrow syntax to the interfaces, combinations have been made of Minimalism and OT (see Samek-Lodovici 2005, 2006 and other papers in Broekhuis and Vogel 2006). In such a model, as for example Broekhuis's (2008) Derivations and Evaluations model, the IS-related interface rules could fit in easily. Zerbian (2006) already shows for Northern Sotho how IS, prosody and word order can be modeled in OT. I refer the reader to her work for more information and an overview of the issues that arise in general in the combination of a syntactic framework with OT. Morimoto (2000) combines OT and LFG to account for inversion constructions and their agreement in Bantu languages like Kirundi.

Implementing IS interface rules as OT constraints entails that these constraints are ranked with respect to other constraints regarding word order. As a simple example, one constraint may require that elements with a referent more salient than the verb occur under/after the verb, and another constraint may require (agreeing) subjects to be moved to a position higher than the verb (EPP). The first constraint concerns word order and IS, the second word order and syntactic functions. In the case that the subject is more salient than the verb, the constraints are in conflict and have a different optimal output: the "IS constraint" prefers VS order, whereas the "syntax constraint" prefers SV order. Depending on the ranking of these constraints, the one word order or the other comes out as optimal. In a language where the constraints on IS in word order are ranked very high, these constraints should not be violated. This implies that the word order encodes the discourse functions, rather than the syntactic functions. In the example of the salient subject, the VS word order would come out as optimal in this language, as for example in Sesotho (Demuth 1990). In a language that ranks the constraints on syntactic functions higher than those concerning IS, the optimal word order encodes the syntactic functions rather than the discourse functions. In the example, this language would have SV as the optimal output, as for example English. Ranking the "IS constraints" and the "syntax constraints" in an intermingled way results in a word order that is determined partly by the need to encode discourse functions and partly by the need to encode

syntactic functions. This is of course a simplified picture, and there are far more constraints, which result in a far more complex interaction. The "configurationality" of a language could in an evaluational (OT) model be said to be dependent on the relative ranking of constraints, or, in other words, different rankings would correspond to different positions on the continuum from "(syntax-) configurational" to "discourse-configurational".

Another way to implement the interface rules is to view them as rules that ensure the mapping of syntax and IS during the derivational process. This derivational approach, more in the line of thought of Epstein et al. (1998), creates only one derivation. This seems to be more economical than the overgeneration in the OT model, but in order to ensure that this one derivation is indeed a correct one, the sentence has to be evaluated at every step of the derivation. The derivation and evaluation continues until at some representation of the sentence the interface rules are met and the sentence is ready.

When I say in this thesis that an element moves for interpretational reasons, or because of the interface rule, I do not mean that the syntax can look ahead and anticipate the interface rules. Rather, an element can be moved in syntax and be found to occupy the right position at the interface in either of the implementations described above, whether this happens only once after the derivation (OT) or several times during the derivation. When the reason that an element is in the correct position after movement is that the configuration complies with the interpretation rule, one can say that the element moved there for interpretational reasons.

Further research will have to show how exactly the combination of IS and syntax can be fruitfully implemented in a model of syntax and its interactions with other cognitive modules. This would have to include not only research on the prosody-focus interaction or scrambling, but also on how focus in various languages is adjacent to the verb, as well as the differences in interpretation found in languages like Makhuwa between elements in the preverbal and postverbal domain. In this thesis only the interface rules are discussed, which can be implemented in either system.

3.5   Conclusion

Information structure is concerned with the linguistic packaging of information, reflecting the discourse representations. Not all discourse information is encoded in the sentence, however. The relative accessibility and salience are relevant, as is exclusivity in Makhuwa. These are the properties the grammar encodes, and notions like topic and focus refer to the pragmatic relation between a referent and the proposition.

Word order is one of the ways in which IS can be expressed, but word order is also used to encode syntactic functions. This tension between syntax and IS is resolved differently in every language. Possibly depending on the alternative means available for encoding the syntax (or IS), word order can be used to encode more of the syntax or more of the IS. Hence, "configurationality" could be viewed as a continuum between the extremes of syntax-configurational and discourse-configurational.

Applying Slioussar's (2007) configurational model of grammar and information structure, I assume that referents have a certain value for accessibility and salience and sometimes also for exclusivity in Makhuwa. In every sentence the syntax derives, these values are encoded in their position relative to the verb. The sentences that the syntax derives are checked, or filtered, at the interface, according to universal and language-specific interface rules. These ensure the right interpretation of the referents and the right position of the linguistic expressions with respect to each other.

The cartographic model, another way to combine minimalist syntax and information structure, cannot account for movement for negative or altruistic reasons and is problematic in the encoding of relative notions. These problems are exemplified in the chapters 4 and 5, where the configurational model is applied to Makhuwa.

Chapter 4 discusses the elements occurring in the pre- and postverbal domains in terms of their syntactic properties and IS status. Specific attention is paid to elements corresponding to concepts that are necessarily low in accessibility, such as *wh*-words, and elements that are high in salience, such as answers to *wh*-questions. After exploring the possibilities and impossibilities of elements with various IS values in various positions, it is demonstrated how an interface rule adapted for Makhuwa can account for most of the different word orders and interpretations in Makhuwa. Chapter 5 takes a closer look at the conjoint/disjoint system, examining the exact interpretation of the elements following the two different verb forms and the contexts in which the CJ or DJ verb form is obligatorily or preferrably used. Although the cartographic model can account for the CJ/DJ facts in Makhuwa, it is shown that the interpretation and use of the CJ/DJ alternation can also be formulated in the configurational interface model: an additional interface rule accounts for the exclusive interpretation of the element immediately after the verb.