A Role and Reference Account of Information Structure in Somali

Jennifer Girard

Follow this and additional works at: https://commons.und.edu/theses

Part of the Psychology Commons

Recommended Citation
https://commons.und.edu/theses/957
This thesis, submitted by Jennifer F. Girard in partial fulfillment of the requirements for the Degree of Master of Arts from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

[Signatures]

This thesis meets the standards for appearance, conforms to the style and format requirements of the Graduate School of the University of North Dakota, and is hereby approved.

[Signatures]

Date

December 11, 2001

Date
PERMISSION

Title A Role and Reference Account of Information Structure in Somali
Department Linguistics
Degree Master of Arts

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work, or in his absence, by the chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Signature  

Date  December 1, 2001
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>ix</td>
</tr>
<tr>
<td><strong>CHAPTER 1: INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Purpose and Methodology</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Role and Reference Grammar</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Information structure in Role and Reference Grammar (RRG)</td>
<td>9</td>
</tr>
<tr>
<td><strong>CHAPTER 2: ANALYSIS OF SOMALI FOCUS STRUCTURE</strong></td>
<td>18</td>
</tr>
<tr>
<td>2.1 Overview of Somali Sentence Structure</td>
<td>18</td>
</tr>
<tr>
<td>2.1.1 Pronoun Paradigms</td>
<td>21</td>
</tr>
<tr>
<td>2.1.2 Verb Agreement</td>
<td>25</td>
</tr>
<tr>
<td>2.1.3 Case Marking</td>
<td>27</td>
</tr>
<tr>
<td>2.2 Analysis of Somali focus structure</td>
<td>28</td>
</tr>
<tr>
<td>2.2.1 The baa construction</td>
<td>28</td>
</tr>
<tr>
<td>2.2.2 The waxaa construction</td>
<td>29</td>
</tr>
<tr>
<td>2.2.3 The waa construction</td>
<td>30</td>
</tr>
<tr>
<td>2.2.4 Narrow focus: subjects</td>
<td>31</td>
</tr>
<tr>
<td>2.2.5 Narrow focus: objects</td>
<td>40</td>
</tr>
<tr>
<td>2.2.6 Predicate focus</td>
<td>45</td>
</tr>
<tr>
<td>2.2.7 Sentence focus</td>
<td>54</td>
</tr>
<tr>
<td>2.2.8 Yes/No Questions</td>
<td>55</td>
</tr>
<tr>
<td>2.2.9 Narrow focus: Predicative and Non-Predicative Adpositional Phrases</td>
<td>56</td>
</tr>
<tr>
<td><strong>CHAPTER 3: ALTERNATIVE ANALYSES OF SOMALI FOCUS STRUCTURE</strong></td>
<td>62</td>
</tr>
<tr>
<td>3.1 Overview</td>
<td>62</td>
</tr>
<tr>
<td>3.1.1 Saeed (1984)</td>
<td>62</td>
</tr>
<tr>
<td>3.1.2 Livnat (1983, 1984)</td>
<td>70</td>
</tr>
<tr>
<td>3.1.3 Aspects of Discourse Configurationality</td>
<td>72</td>
</tr>
<tr>
<td>3.2 Conclusion</td>
<td>74</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Organization of RRG (Van Valin and LaPolla 1997:21) ..................................4
Figure 2. Abstract LSC (Van Valin and LaPolla 1997:38) ................................................5
Figure 3. Japanese LSC .......................................................................................................7
Figure 4. English WH-question (adapted from Van Valin and LaPolla 1997:235) ..........9
Figure 5. Focus structure in English (adapted from Van Valin and LaPolla 1997:235) .. 16
Figure 6. Narrow Focus (Intransitives) ..............................................................................34
Figure 7. Narrow focus (subject), for 31b ........................................................................37
Figure 8. Narrow focus (subject), for 31c ........................................................................37
Figure 9. Narrow Focus (subject [for 31d] ) .................................................................. 38
Figure 10. Narrow focus (subject [for 31e]) in waxaa constructions. ............................... 39
Figure 11. Narrow focus with waxaa (object) .................................................................... 44
Figure 12. Narrow focus (object), for 37c ........................................................................44
Figure 13. Question eliciting predicate focus ................................................................... 50
Figure 14. Predicate Focus ............................................................................................... 52
Figure 15. Argument-Adjunct Focus ............................................................................... 58
Figure 16. Direct Core Argument Focus .......................................................................... 59
Figure 17. Peripheral Focus ............................................................................................. 61
Figure 18. Predicate focus ............................................................................................... 68
LIST OF TABLES

Table 1. Operators and the LSC................................................................. 7
Table 2. Nominative Pronouns (Livnat 1984:11)................................. 22
Table 3. Accusative Pronouns (Livnat 1984:11)................................. 22
Table 4. Verb Paradigms (Suffix Only) (based on Saeed 1984:55-56)...... 25
Table 5. Verb Paradigms (Prefix/Suffix) (based on Svolacchia et al. 1995:94) 25
ACKNOWLEDGEMENTS

Special thanks to my thesis committee, Dr. Steven Marlett, Chairperson, Dr. Albert Bickford, and Dr. Cheryl Black, for their willingness to advise me and serve in this capacity. Their suggestions and comments were most helpful in bringing this thesis into its final form. Thanks also to Dr. Daniel Everett, whose helpful insights contributed to this thesis.

My deepest gratitude and appreciation goes to Salim Hadad, his wife Munira, and their children: Bilqis, Mohamed, Isam, Yasir and Oshine. Salim gave many hours of his free time in the collection and verification of the data in this thesis, and his family welcomed into their home as one of their own, with warm hospitality and friendship. Special thanks also goes to Abdirahman Jama and his brother Ali Farah, who generously assisted me in the collection of Somali data, provided me with Somali literature, and patiently gave their insights. Without their faithful assistance and patience, this project would not have been completed.

Many thanks to my friends and family in Charlotte, NC and at SIL-UND, who supported me and prayed for me during the writing of this thesis. Above all, my deepest thanks goes to God, Creator of the beautiful Somali language and people. He has sustained me throughout this project.
ABSTRACT

This thesis is a study of focus structure in Somali within the framework of Role and Reference Grammar (RRG). Focus is an important phenomenon in Somali, and this thesis investigates the different types of focus which occur in Somali, the effects of the focus markers on other features in the language, and the interaction between syntax and focus structure. Data on Somali focus structure was collected in consultation with three native speakers of Somali, and also gathered from previous research on the Somali language. Although word order in Somali is relatively free, focus structure is correspondingly rigid, a finding which supports the RRG hypothesis that syntax and focus structure adapt to each other as necessary.
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>ADV</td>
<td>adverb</td>
</tr>
<tr>
<td>AP</td>
<td>adpositional phrase</td>
</tr>
<tr>
<td>ARG</td>
<td>argument</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
</tr>
<tr>
<td>EXC</td>
<td>exclusive</td>
</tr>
<tr>
<td>EXT</td>
<td>extensive</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>FOC</td>
<td>focus marker</td>
</tr>
<tr>
<td>INC</td>
<td>inclusive</td>
</tr>
<tr>
<td>LDP</td>
<td>left-detached position</td>
</tr>
<tr>
<td>LSC</td>
<td>layered structure of the clause</td>
</tr>
<tr>
<td>M</td>
<td>masculine</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>NUC</td>
<td>nucleus</td>
</tr>
<tr>
<td>O</td>
<td>object</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>PoCS</td>
<td>postcore slot</td>
</tr>
<tr>
<td>PP</td>
<td>prepositional phrase</td>
</tr>
<tr>
<td>PrCS</td>
<td>precore slot</td>
</tr>
<tr>
<td>PRED</td>
<td>predicate</td>
</tr>
<tr>
<td>PRO</td>
<td>pronoun</td>
</tr>
<tr>
<td>PSA</td>
<td>privileged syntactic argument</td>
</tr>
<tr>
<td>RDP</td>
<td>right-detached position</td>
</tr>
<tr>
<td>REST</td>
<td>restrictive</td>
</tr>
<tr>
<td>RRG</td>
<td>Role and Reference Grammar</td>
</tr>
<tr>
<td>S</td>
<td>subject</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>V</td>
<td>verb</td>
</tr>
</tbody>
</table>

CHAPTER 1: INTRODUCTION

1.1 Purpose and Methodology

The goal of this study is to provide a description and analysis of affirmative declarative main clause focus structure in Somali within the framework of Role and Reference Grammar (RRG) (Van Valin and LaPolla 1997). The relationship of declarative clause focus structure to interrogatives will also be explored. Previously, focus structure in Somali has been analyzed solely within the framework of generative grammar (Livnat 1983, 1984; Saeed, 1984, Svolacchia et al. 1995) and, with a few exceptions, these analyses have overlooked important semantic and pragmatic components of focus structure. Chapter 1 includes a brief outline of the relevant points from RRG, and the Lambrechtian (1994) paradigms used to elicit various types of focus structure. Chapter 2 describes important features of Somali syntax which are relevant to information structure and also analyzes Somali information structure within the framework of RRG. Chapter 3 considers previous analyses of Somali focus structure and some advantages of the RRG description. The principal purpose of this thesis is to show how Somali exemplifies the typological model presented in RRG which posits that syntax and focus structure adapt to each other, and also to show how focus structure affects various morphosyntactic phenomena in the language, namely verb agreement and the occurrence of clitic pronouns.
The data presented in this thesis was collected and verified with the assistance of three native speakers of Somali: Abdirahman Hasan Jama, Ali Farah, and Salim Hadad. The acceptability of all examples was verified with at least one of these native speakers. Although these speakers have lived outside of Somalia since a civil war began there in 1990, they continue to speak Somali on a daily basis with friends and relatives. All three have at least the equivalent of two years education in a Somali university, and two are experienced in educating children and adults in both Somali and English.

The methodology I used was to consult the literature, in order to orient myself to the basic sentence structure of Somali. Then I designed elicitation sessions on the basis of relevant data which I found in the literature (Mathews 1984, Saeed 1984, Livnat 1984, Ajello 1984, Hayward and Saeed 1983, Svolacchia et al. 1995). (Any data that was directly taken from another source will be documented as such in the text.) Initially, I tested some of the examples in the literature with the consultants, to verify their abilities, and then devised question/answer paradigms by adapting Lambrecht’s (1994) paradigms into Somali. Based on previous research on Somali focus structure, elicitation sessions were conducted using the question/answer pairs listed in Chapter 2.

In some instances, the consultant was given a question in Somali and a list of potential answers in Somali. He was asked to verify which answers were appropriate to the question, which answers were inappropriate, and which answers were ungrammatical. Appropriate answers were also ranked in order of preference. Other elicitation sessions consisted of questions given in English, and a list of desired answers given in English; translations were provided by the consultant. Another format for elicitation involved presenting the question in Somali, and requesting the best possible answers in Somali.
Somali belongs to the East Cushitic subgroup, the Cushitic subfamily, and the Afroasiatic family. It is spoken by over ten million people, mainly in Somalia, Kenya and Ethiopia. In 1972, the Somali orthography was standardized; this standard orthography will be used throughout the thesis (Andrzejewski 1975:7).

1.2 Role and Reference Grammar

The analysis in this thesis is based on the theoretical framework of Role and Reference Grammar. RRG is a syntactic theory based upon the proposal that "grammatical structure can only be understood with reference to its semantic and communicative functions" (Van Valin and LaPolla 1997:13). RRG is also concerned with Typological Adequacy; one of its goals is to develop a theory of syntax based on an extensive list of languages (Van Valin and LaPolla 1997:14-15). It is my hope that this study of Somali will assist in furthering this goal. An additional goal of RRG is to present a descriptive framework which can be used by field linguists for writing grammars (Van Valin and LaPolla 1997:15).

There are several key points in RRG which distinguish it from more formally-oriented theories:

1) There are no multiple levels of syntactic representation underlying syntactic forms. In RRG there is only a surface syntactic form, which is mapped directly onto the semantic representation (and vice versa) through a linking algorithm (Van Valin and LaPolla 1997:21). The relationship between the syntactic and semantic levels is not derived, but it simply involves an association between different types of information, semantic and syntactic, which describe the same structure.
There are two general requirements that RRG considers necessary for a solid theory of clause structure:

a) The theory should capture all of the universal features of language without imposing features on languages in which there is no evidence for them.

b) A theory should represent comparable structures in different languages in comparable ways. (Van Valin and LaPolla 1997:22).

The following two categories play a role in the syntax of every language, and the contrasts defined by these categories must be represented in the clause structure proposed by RRG:

a) There is a contrast between predicating elements (i.e. verbs, copulas or nominal predicates), and non-predicating elements (which include all other constituents).

b) There is also a contrast between (semantic) arguments and non-arguments of the predicate, i.e. between those NPs and adpositional phrases which
are required by the semantic representation of the verb, and those which are not. (Van Valin and LaPolla 1997:25,26).

In RRG, a tree diagram represents the various universal aspects of the layered structure of the clause (LSC). Parentheses indicate elements which are not required in every sentence.

Figure 2. Abstract LSC (Van Valin and LaPolla 1997:38)

As shown in Figure 2, there are five layers which comprise the LSC. (The predicate and the nucleus are identical; therefore, the predicate does not represent a separate layer.) The bottom layer is comprised of CONSTITUENTS, which are the syntactic categories represented in the sentence. The next level contains the NUCLEUS, consisting solely of the predicate (the verb plus any of its operators¹), and also the ARGUMENTS of the verb. The third level is the CORE, which is composed of the nucleus plus any syntactic ARGUMENTS of the verb. Next is the CLAUSE, which contains the core plus any constituents which are not syntactic arguments of the verb. The PRECORE SLOT (PrCS) and

¹Operators include any elements (words, clitics or affixes) such as tense, aspect, negation and illocutionary force.
POSTCORE SLOT (POCS) may contain either semantic arguments of the verb or non-arguments. They generally include PPs or ADVs which are not set off by a pause or intonation break. The PERIPHERY may contain only adjuncts of the verb. In English, adjuncts generally include locative and temporal PPs (e.g. *after the party*) and bare NP adverbials (e.g. *yesterday*) (VanValin and LaPolla 1997:27). Finally, there is the SENTENCE, which consists of the clause plus any extra-CLAUSAAL participants (which are in the LEFT DETACHED (LDP) or RIGHT DETACHED POSITION (RDP)). Constituents in the LDP or RDP; these constituents are set off by a pause or intonation break. Although linear order is not completely constrained (i.e. lines may cross in tree diagrams), there are several universal linear precedence rules which apply in RRG:

1) The LDP must precede the clause.
2) The clause must precede the RDP.
3) The PrCS must precede the core.
4) The core must precede the PoCS. (Van Valin and LaPolla 1997:71)+

The layered structure of the clause (LSC) may represent any language, because of the flexibility in constituent order. According to Van Valin and LaPolla (1997:32), the linear order of the core arguments and the predicate is irrelevant to the determination of whether an element is in the nucleus, core or periphery.’ In other words, the nucleus, core arguments and periphery may appear in any order, as shown in Figure 3.
In addition, some grammatical features which modify the constituents in the LSC are referred to as operators. Each type of operator is a member of a particular layer of the clause structure, as shown in Table 1.

Table 1. Operators and the LSC

<table>
<thead>
<tr>
<th>RRG layer</th>
<th>RRG operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleus</td>
<td>Aspect</td>
</tr>
<tr>
<td></td>
<td>Negation</td>
</tr>
<tr>
<td>Core</td>
<td>Directionals</td>
</tr>
<tr>
<td></td>
<td>Modality (Root)</td>
</tr>
<tr>
<td></td>
<td>Negation</td>
</tr>
<tr>
<td>Clause</td>
<td>Status$^2$, negation</td>
</tr>
<tr>
<td></td>
<td>Tense</td>
</tr>
<tr>
<td></td>
<td>Evidentials</td>
</tr>
<tr>
<td></td>
<td>Illocutionary force</td>
</tr>
<tr>
<td>Sentence</td>
<td>None</td>
</tr>
</tbody>
</table>

(Adapted from Van Valin and LaPolla 1997:47).

RRG does not posit grammatical relations such as subject and object. Instead, RRG distributes the properties of grammatical relations among several different syntactic, semantic and pragmatic components of the grammar. The first set of

$^2$Status refers to modality and categories such as realis and irrealis (Van Valin 1997:41)
components distinguishes between the semantic roles of NPs in the sentence, which are essentially two in number: **ACTOR** or **UNDERGOER**. These **MACROROLES** (the term coined by Van Valin to describe these semantic roles) are "generalizations across the argument-types found with particular verbs." Grammatical rules in the languages generally refer to these macroroles, rather than "specific arguments in the logical structure" of the verb (Van Valin and LaPolla 1997:139). NPs are classified according to the semantic role they have in the sentence: the actor is generally agentive, while the undergoer is patient-like. For example, in a standard active transitive sentence in English, the actor would correspond to the subject while the undergoer would correspond to the object.

The second set of components distinguishes between the **PRIVILEGED SYNTACTIC ARGUMENT** (PSA) and the macrorole which is not the PSA. The PSA corresponds roughly to the notion of syntactic subject, while the non-PSA corresponds roughly to the notion of syntactic object. The PSA is construction-specific, and in a syntactically accusative language, the highest-ranked macrorole in the sentence will always be chosen as the PSA. In a syntactically ergative language, the lowest-ranked macrorole will be chosen as the PSA. Actors will always outrank undergoers in choosing a PSA in a syntactically accusative language. If a sentence contains an actor and an undergoer, the actor will be the PSA, but if the sentence only contains an undergoer as argument, the undergoer will be selected. In English, the undergoer functions as the PSA in a passive sentence.

Core (syntactic) arguments are represented by the ARG node which occurs above the constituent, while semantic arguments are represented by the ARG node occurring below the constituent (Van Valin and LaPolla 1997:215). Both core and semantic
arguments are part of the semantic representation of the verb (Van Valin and LaPolla 1997:26).

Figure 4. English WH-question (adapted from Van Valin and LaPolla 1997:235)
centered on syntactic form, and functional theories, whose approach is centered on the communicative aspects of language. His purpose is to relate the formal structure of sentences to the communicative situations in which they are used (Lambrecht 1994:1).

One of the central questions of Lambrecht's theory of information structure is one which has been curiously ignored in generative grammar: "Why should grammars provide the means of generating so many different syntactic and prosodic structures for expressing one and the same propositional content?" (Lambrecht 1994:8).

Prior to Lambrecht, Prague School linguists such as Daneš considered the "level of the utterance" (i.e. information structure) as one of the three components of grammar (along with syntax and semantics). Firbas states that the purpose of the utterance is to "make it possible [for speakers] to understand how the semantic and grammatical structures function in the very act of communication" (Firbas, quoted in Daneš 1966:27, in Lambrecht 1994:6). Unlike Lambrecht, however, the Prague School stated that the purpose of the level of the utterance was to "convey some extra-linguistic reality reflected by thought" (Firbas, quoted in Danes 1966:27, quoted in Lambrecht 1994:6).

Lambrecht's paradigm of information structure strives to integrate form and function more closely than the Prague School did. As a result, he proposes that any information structure analysis must only refer to phenomena which are represented linguistically in the form of the sentence. Any interpretation of a sentence must be based on the actual form, not just inference. Lambrecht (1994:7) considers the clause, not the entire discourse, to be the domain of information structure, although his theory of information structure does consider the interaction of clauses with their surrounding discourse.
Only one version of generative grammar, the Extended Standard Theory (EST), made even minimal mention of information structural factors as relevant to syntax (see Chomsky 1970, Jackendoff 1972, Akmajian 1973, quoted in Lambrecht 1994). But EST never developed any significant proposals along these lines. Previous research (Livnat 1984) on Somali focus structure is based on a version of EST and briefly considers the semantic and pragmatic implications of its syntactic analysis.

Information structure distinctions are manifested cross-linguistically in several different forms in sentences: morphologically (usually through affixes or clitics), syntactically (through word order variations), prosodically (through stress, tone and/or intonation) and lexically (through choices of specialized lexical items). In Somali, information structure distinctions are manifested primarily morphologically, through clitics.

An information structure analysis of a particular language must be based on a comparison between semantically equivalent, but formally and pragmatically divergent pairs of sentences, known as allosentences (Lambrecht 1994:6). Allosentences are utterances which differ in only one feature. An English example of allosentences follows:

(1) a. Mary saw JOHN.
   b. MARY saw John.

\[ ^3 \] The concept of allosentences is mentioned by Zellig Harris (1951:46). He uses different terminology and refers only to phonological differences between sentences.
It is only through these contrasts that information structure may be analyzed.

There are two important assumptions which guide my analysis of the information structure of Somali sentences:

1) There is a distinction between the propositional (basic, non-contextual) meaning and the utterance-specific meaning of a sentence within its context. In this paper, I will use yes-no and WH-questions to determine the utterance-specific structures and their meanings in Somali, and based on these structures, present an analysis of Somali focus structure within the framework of information structure presented in RRG.

2) The unmarked constituent order of a language is best determined using full lexical, not pronominal, arguments because pronominal arguments often obey different syntactic and prosodic constraints (Lambrecht 1994:15). This can be challenging to do when using questions and replies, because the most natural replies are often in pronominal form. In Somali, using pronominal arguments in elicitations is the only way to determine the verb paradigms and the occurrence of coreferential pronoun clitics, because the third person clitics are often null. Independent pronouns, however, receive focus and trigger verb agreement in the same way as lexical nouns. It is difficult to determine a basic, unmarked word order in Somali, because there are no declarative affirmative sentences which remain morphologically unmarked for focus structure.

Lambrecht (1994:213) defines FOCUS as the semantic component of a pragmatically structured proposition whereby the assertion differs from the
presupposition. The **PRESUPPOSITION** is the portion of the semantic structure of an utterance that contains the information that is shared between the speaker and the hearer, and the **ASSERTION** is the portion which contains information not shared by the speaker and the hearer. For all practical purposes, the assertion and the focus coincide. Every utterance must contain at least an assertion, i.e. every utterance must give the hearer new information. A focused constituent is one which fills in the missing variable in an open proposition (Lambrecht 1994:122, 223). The distinction between assertion and focus is not important for the purpose of this thesis, since it is mostly concerned with the focused constituent in an utterance; i.e. the lexicogrammatical representation of the focus.

In contrast, a **TOPIC** is the "entity within the pragmatic presupposition that has the function of naming the referent that the assertion is about" (Van Valin and LaPolla 1997:203). It is not obligatory in every utterance, though it is usually present. The topic must be part of the pragmatic presupposition.

Lambrecht (1994:223) presents a taxonomy which distinguishes the following three types of focus structure: **PREDICATE FOCUS**, **NARROW (constituent) FOCUS** and **SENTENCE FOCUS**. They are distinguished as follows (focused constituents are indicated in bold caps, and prosodically prominent constituents are underlined):

1) An utterance with predicate focus corresponds closely to the notion of a topic-comment structure, in which the predicate is focused and comments on the topic, which is the subject. This is by far the most common type of focus cross-linguistically. For example:

(2) Q. What happened to your car?

   A. My car/It **BROKE DOWN**. (Lambrecht 1994:223).
In example 2, the focused constituent is the predicate. In English, the focused constituent is indicated by prosodic prominence. The prosodically marked constituent does not necessarily correspond one to one with the focused constituent in every instance (i.e. in example 2, the predicate is ‘broke down’, while the prosodic prominence only falls on ‘down’).

2) In an utterance with narrow focus, the focused constituent identifies the “missing argument in a presupposed open proposition” (Lambrecht 1994:223). 4 For example:

(3) Q. Did your **MOTORCYCLE** break down?

A. No, my **CAR** broke down.

Sentences with narrow focus are also referred to as identificational, which means that their function is to provide the referent solicited by the WH-word in a WH-question (Lambrecht 1994:122). In narrow focus constructions, the focus is limited to a single constituent (Van Valin and LaPolla 1997:208).

3) An utterance with sentence focus is characterized by the lack of a presupposition, i.e. the assertion coincides with the focus and extends over the entire proposition (Lambrecht 1994:233). The purpose of sentence focus is to introduce a referent into discourse or to announce an event which introduces a referent into the discourse (Lambrecht 1994:14). For example:

4 Lambrecht’s terminology to refer to this phenomenon is ‘argument focus’, while Van Valin (1997:208) uses the term ‘narrow focus’. There is no difference in meaning between the two terms. I am adopting the term ‘narrow focus’ to avoid confusion with the term ‘argument’, defined as ‘a constituent required by the verb’s semantic structure’.
(4) Q. What happened?

A. **MY CAR BROKE DOWN.**

In example 4, the utterance with sentence focus is homophonous to the utterance with narrow focus, though this is not necessarily true cross-linguistically. The only distinction between the two sentences is the presupposition evoked in the question: in example 4, there is no presupposition, while in example 3, the presupposition is that something broke down.

RRG groups the three categories of predicate, narrow and sentence focus into two larger categories. The first is **BROAD FOCUS**, in which more than one constituent is focused (Van Valin and LaPolla 1997:206). Predicate focus and sentence focus are subsumed under this category. The second category is **NARROW FOCUS**, which subsumes the category of narrow (constituent) focus.

RRG also distinguishes between **MARKED** and **UNMARKED FOCUS STRUCTURE**. Predicate focus is universally considered the unmarked structure, while narrow and sentence focus are marked (Lambrecht 1994:296). Because of the unmarked status of predicate focus sentences and because predicates contain more than one constituent, they may have more than one interpretation. The possibility of multiple interpretations means that the intended reading must be multiply marked, e.g. prosodically and syntactically, or prosodically and morphologically, etc. Lambrecht concludes that predicate focus structure is 'inherently ambiguous', due to the presence of an unaccented topical element (i.e. the object) within the predicate focus domain (1994:303).

RRG helps to clarify the potential ambiguity of predicate focus sentences by positing a distinction between the actual and potential focus domains (Van Valin and...
Both the actual and potential focus domains are construction-specific: they differ among utterances within a language based on the type of focus structure in a particular sentence. The *actual focus domain* is the syntactic domain in which the focus actually occurs in a particular utterance. The *potential focus domain* is the syntactic domain in which the focus may occur in a particular utterance. The actual focus domain must occur within the potential focus domain. The only universal constraint on the potential focus domain is that it must occur within the clause. The focus structure is represented graphically by a separate projection. In Figure 5, for example, the potential focus domain is the entire sentence, while the actual focus domain is the WH-phrase in the PrCS.

![Figure 5. Focus structure in English (adapted from Van Valin and LaPolla 1997:235)](image)

RRG adopts a central assumption of Lambrecht’s theory of information structure, which states that all sentences contain focus; i.e. there are no sentences that are neutral with regard to focus structure. One of the basic purposes of communication is to convey new information, and focus is closely (though not solely) associated with the introduction
of new information. There are various ways in which focus structure is represented (i.e. syntactically, morphologically, or prosodically), even within a language. Lambrecht rejects proposals previously made by generative linguists stating that the various focus structures within a language are transformationally related (1994:319). Previous research on Somali focus structure (Saeed 1984:81) proposes that two of the constructions which indicate focus in the language are related by transformations. In Chapter 3, I will consider how the information structure principles in RRG make possible an alternative analysis which renders this transformational explanation unnecessary.

In the following chapters, I demonstrate how Lambrecht’s (1994) principles of information structure, which are geared toward languages that express focus prosodically and syntactically, can be adapted to a language such as Somali, in which focus structure is primarily expressed through morphology.

There are two morphological indicators of focus in Somali and three different constructions in which these indicators appear. Exactly one of these two indicators must occur in each main declarative affirmative clause in Somali. Using Lambrecht’s paradigms, I analyze the types of focus indicated by each of these indicators, investigating whether they are mutually exclusive in the types of focus they indicated. I present the types of constituents which can be focused and analyze the potential and actual focus domains for each type of focus structure. In addition, I show how information structure in Somali interacts with verb agreement and the occurrence of coreferential pronoun clitics, and explain how the relatively free word order in Somali is accounted for by the relative rigidity of its focus structure.
2.1 Overview of Somali Sentence Structure

Each declarative affirmative main clause in Somali contains a focus marker; therefore, any basic description of Somali syntax should begin with a description of focus structure. Basic word order in Somali is generally defined as SOV, although word order is relatively free and SVO and OSV are also common word orders, as shown in example 5:

   Ali FOC Maryan defeated-3SG.M
   ‘ALI defeated Maryan.’

b. Cali baa ka raay-ay Maryan.

c. Maryan Cali baa ka raayay.7
   ‘ALI defeated Maryan.’

All the examples in 5 have the same propositional meaning, and their usage is governed by pragmatic factors. Examples 5a-c are considered standard declarative Somali sentences.

5 The meaning of the verb ‘ka raayay’ is to defeat someone in a game.

6 Cali is transliterated into English as Ali; the grapheme <c> represents a voiced pharyngeal fricative.

7 The presence of a pause following Maryan in 5c is possible, but not preferred.
One of the most important phenomena of Somali syntax is the use of a group of particles that have special semantic functions in the sentence. Andrzejewski (1975:3) uses the term **INDICATOR** to refer to these particles. According to him, the main function of these indicator particles is to signal completeness of a sentence; that is, without one of these indicators, a sentence is merely a dependent clause. This is consistent with the fact that no main declarative sentence in Somali may occur without a focus marker, because focus markers belong to the group of indicator particles. In 5a-c, *Cali* is the focused constituent and *baa* is the indicator particle, which denotes focus. There are nine particles in Andrzejewski’s list of indicators\(^8\), which include *baa* and *waa*, the two focus markers described below. Every main declarative affirmative Somali sentence must contain either *baa* or *waa*, and every sentence must contain one of the indicator particles.

Based on Andrzejewski’s (1975) description, I have classified the six remaining indicator particles in his list as **ILLOCUTIONARY FORCE** operators, whose function is to determine whether a sentence is an "assertion, question, command or wish" (Van Valin and LaPolla 1997:41). Illocutionary force operators are universally present in language. The particles are listed below:

- *ha* – negative imperative
- *ma* – interrogative
- *miyaa* – Andrzejewski analyzes this particle as a separate marker, but other sources present it as a combination of *ma* and *baa*.

\(^8\) Andrzejewski classifies *waxa* as an indicator particle; its classification as such conflicts with my analysis and I have not listed it here.
- **soo** – negative interrogative
- **yaa** – negative optative
- **weeye** – copula

Every main declarative affirmative clause in Somali contains a morphological indicator of focus, according to Livnat (1983:89), and the following examples show how *baa* (including its optional variant *ayaa*)\(^9\) and *waa* indicate focus in Somali:

1) *baa*- When *baa* or *ayaa* follows any NP or oblique, that constituent is the focused constituent, as shown in examples 6 and 7 (recall that focus is defined as the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition):

   (6) *Idink-a ayaa y-imid.*
   
   you(PL)-ACC FOC REST-came.
   
   'YOU (PL) came.'

---

\(^9\) Previous research on Somali focus structure (Saeed 1984, Livnat 1984, et al) has indicated that *baa* and *ayaa* are optional variants, though it is certainly possible that they play different roles in the discourse. I do not have enough information at this point to determine the distinction, if any, in their roles. Saeed (1984:78-79) states that *baa* may undergo phonological rules (see Footnote 8), while *ayaa* does not. The choice of *ayaa* may be motivated by a desire to preserve a 'more transparent morphological structure' (Saeed 1984). I did not investigate this phenomenon and could not find any detailed differentiation between them (beyond the above explanation) in the literature. I will use *baa* to refer to both *baa* and *ayaa* in the text.
Baa is an enclitic with respect to the constituent that it focuses, and may occur only once in each clause (Saeed 1984:23). No elements may occur between baa and the constituent it focuses, except for any suffixes on the focused constituent. In addition, various types of constituents such as ADV, NP, and WH-words, can be focused by baa. This fact lends evidence to the claim that baa is a clitic instead of an affix.

2) The proclitic waa focuses the predicate in clauses where it occurs, as shown in the following example:

(8) Cali waa qosl-ay.
    Ali  FOC  laughed-3SG.M.PST
    ‘Ali LAUGHED.’  (Livnat 1983:93)

Like baa, waa may only occur once in each clause. Certain operators may occur between waa and the verb. The predicate may consist of a verb, a noun (in an equative sentence), or an adjective plus a copula verb.

2.1.1 Pronoun Paradigms

In Somali, the form of a pronoun is determined by its syntactic role in the sentence: either subject or object. The nominative form is generally used for subjects, while the accusative form is generally used for non-subjects. The choice of which form to use is also affected by focus structure, as discussed in Section 2.1.3. The nominative and accusative pronouns each have two forms: 1) an independent form, which functions similarly to an NP. The independent form is comprised of a base plus the appropriate case marker. 2) a clitic form, which must be phonologically attached to another word or particle. The nominative clitics are enclitics and attach to a preceding word or clitic,
while the accusative clitics are proclitics and attach to the word or clitic that follows. The forms of the pronoun paradigms are given below:

Table 2. Nominative Pronouns (Livnat 1984:11)

<table>
<thead>
<tr>
<th></th>
<th>Independent Form</th>
<th>Enclitic Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>anigu</td>
<td>=aan</td>
</tr>
<tr>
<td>2SG</td>
<td>adigu</td>
<td>=aad</td>
</tr>
<tr>
<td>3SG.M</td>
<td>isagu</td>
<td>=uu</td>
</tr>
<tr>
<td>3SG.F</td>
<td>iyadu</td>
<td>=ay</td>
</tr>
<tr>
<td>1PL (INC)</td>
<td>innagu</td>
<td>=aynu</td>
</tr>
<tr>
<td>1PL (EXC)</td>
<td>annagu</td>
<td>=aanu</td>
</tr>
<tr>
<td>2PL</td>
<td>idinku</td>
<td>=aydu (aad)</td>
</tr>
<tr>
<td>3PL</td>
<td>iyagu</td>
<td>=ay</td>
</tr>
</tbody>
</table>

Table 3. Accusative Pronouns (Livnat 1984:11)

<table>
<thead>
<tr>
<th></th>
<th>Independent Form</th>
<th>Proclitic Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>aniga</td>
<td>i=</td>
</tr>
<tr>
<td>2SG</td>
<td>adiga</td>
<td>ku=</td>
</tr>
<tr>
<td>3SG.M</td>
<td>isaga</td>
<td>Ø</td>
</tr>
<tr>
<td>3SG.F</td>
<td>iyada</td>
<td>Ø</td>
</tr>
<tr>
<td>1PL (INC)</td>
<td>innaga</td>
<td>inna=</td>
</tr>
<tr>
<td>1PL (EXC)</td>
<td>annaga</td>
<td>na=</td>
</tr>
<tr>
<td>2PL</td>
<td>idinka</td>
<td>idin=</td>
</tr>
<tr>
<td>3PL</td>
<td>iyaga</td>
<td>Ø</td>
</tr>
</tbody>
</table>

Example 9 shows object focus, while example 10 shows subject focus:

(9) Anig-a ayay Maryan i= ga raay-say.
    me-ACC FOC 3SG.F(NOM) Maryan 1SG(ACC) defeated-3SG.F
    ‘Maryan defeated ME.’

(10) Niman ayaa i= ka raay-ay anig-a.
     men FOC 1SG(ACC) defeated-3PL.REST me-ACC
     ‘The MEN defeated me.’

In example 9, the accusative clitic *i* is coreferential with the object *aniga*, while the clitic *ay* is coreferential with the subject *Maryan*. In example 10, the accusative clitic *i* is again coreferential with the object *aniga*, while there is no clitic pronoun coreferential
with *niman*, the focused subject. Although it may initially appear asymmetrical that a nominative clitic follows *baa* when the object is focused, and not when the subject is focused, the pattern of clitic pronoun occurrence is actually symmetrical. There are two separate processes occurring simultaneously in Somali:

a) The object, whether focused or non-focused, always co-occurs with an accusative pronoun clitic (except in third person, where the clitics are null) (Svolacchia et al. 1995:87). This pronoun clitic (*i* in example 11, and *ku* in example 12) immediately precedes the verb. Its obligatory presence suggests that the accusative clitic pronoun actually functions as an agreement marker.

\[\text{(11) } \text{Anig-a ayay Maryan } i= \text{ ga raay-say.} \]

\[\text{ayaa}=\text{ay me-ACC FOC 3SG.F(NOM) Maryan 1SG(ACC) defeated-3SG.F} \]

'Maryan defeated ME.'

\[\text{(12) Niman ayaa kaa raay-ay adig-a.}^{10} \]

\[\text{ku= ka men FOC 2SG(ACC) defeated-3PL.REST you(SG)-ACC} \]

'The MEN defeated you.'

b) In addition, the topic (as defined in Chapter 1) always triggers the occurrence of a coreferential clitic pronoun following the focus marker (Svolacchia et al. 1995:87, Saeed 1984:31). This phenomenon is exemplified by the enclitic *ay* following *ayaa* in Example 11. If the topic is the subject, a nominative enclitic pronoun will occur (as in example 11). Likewise, if the topic is the object, the accusative proclitic pronoun will occur (as in example 12). The occurrence of the topic

---

\[^{10} \text{Kaa}=\text{ku + ka} \]
pronoun provides the identity of the topic in sentences where the topic NP is not overt, as shown below in example 13b:

(13) a. Cali miyuu arkay Maryan?
   Ali did see Maryan
   ‘Did Cali see Maryan?’

   b. Maya, (asaga) Amina buu ark-ay.
      baa=uu
      No, (he) Amina FOC=3SG.M(NOM) saw-3SG.M
      ‘No, he saw AMINA.’

In a sentence where the topic NP is optional, such as asaga in example 13b, the coreferential pronoun uu will always occur and will function as the subject of the verb. If the object is also the topic, as in example 12, the two coreferential pronouns which would normally occur are identical, and only one appears in the sentence.

The above explanation for the distribution of clitic pronouns accounts for the fact that a nominative clitic pronoun (when the subject is also the topic) may occur in an intransitive sentence with predicate focus (as in example 14), but not in an intransitive sentence where the subject is focused, and therefore not a topic (example 15).

(14) Ninkii wuu dhac-ay.
    waa=uu
    man-DEF FOC 3SG.M(NOM) fell.down-3SG.M
    ‘The man FELL DOWN.’

(15) *Nink-ii buu dhac-ay.
     baa=uu
     man-DEF FOC 3SG.M(NOM) fell.down-3SG.M
     ‘The MAN fell down.’

In each sentence, therefore, a coreferential pronoun will co-occur with both the topic and with the object of the sentence. The topic of a sentence plays a role in determining whether or not a coreferential clitic pronoun will be present after the focus marker. The following syntactic properties are therefore characteristic of the topic in
Somali: 1) the topic is not focused by *baa* and 2) a coreferential clitic pronoun occurs with it following the focus marker.

### 2.1.2 Verb Agreement

In Somali, there are two verb paradigms (Saeed 1984:55-56). In the **EXTENSIVE** paradigm there is a different form for each person/number combination, whereas the **RESTRICTIVE** paradigm only distinguishes three different forms. The selection of which paradigm to use depends on the focus structure of the sentence. Some verbs use only suffixes to indicate agreement with the subject (Table 4), while others use both a prefix and a suffix (Table 5).

#### Table 4. Verb Paradigms (Suffix Only) (based on Saeed 1984:55-56)

<table>
<thead>
<tr>
<th>EXTENSIVE</th>
<th>Translation</th>
<th>RESTRICTIVE</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG keenay</td>
<td>(I) brought</td>
<td>1SG keenay</td>
<td>(I) brought</td>
</tr>
<tr>
<td>2SG keentay</td>
<td>(You(SG)) brought</td>
<td>2SG keenay</td>
<td>(You(SG)) brought</td>
</tr>
<tr>
<td>3SG.M keenay</td>
<td>(He) brought</td>
<td>3SG.M keenay</td>
<td>(He) brought</td>
</tr>
<tr>
<td>3SG.F keentay</td>
<td>(She) brought</td>
<td>3SG.F keentay</td>
<td>(She) brought</td>
</tr>
<tr>
<td>1PL keennay</td>
<td>(We) brought</td>
<td>1PL keennay</td>
<td>(We) brought</td>
</tr>
<tr>
<td>2PL keenteen</td>
<td>(You (PL)) brought</td>
<td>2PL keenay</td>
<td>(You (PL)) brought</td>
</tr>
<tr>
<td>3PL keeneen</td>
<td>(They) brought</td>
<td>3PL keenay</td>
<td>(They) brought</td>
</tr>
</tbody>
</table>

#### Table 5. Verb Paradigms (Prefix/Suffix) (based on Svolacchia et al. 1995:94)

<table>
<thead>
<tr>
<th>EXTENSIVE</th>
<th>Translation</th>
<th>RESTRICTIVE</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG imid</td>
<td>(I) came</td>
<td>1SG imid</td>
<td>(I) came</td>
</tr>
<tr>
<td>2SG timid</td>
<td>(You(SG)) came</td>
<td>2SG yimid</td>
<td>(You(SG)) came</td>
</tr>
<tr>
<td>3SG.M yimid</td>
<td>(He) came</td>
<td>3SG.M yimid</td>
<td>(He) came</td>
</tr>
<tr>
<td>3SG.F timid</td>
<td>(She) came</td>
<td>3SG.F timid</td>
<td>(She) came</td>
</tr>
<tr>
<td>1PL nimid</td>
<td>(We) came</td>
<td>1PL nimid</td>
<td>(We) came</td>
</tr>
<tr>
<td>2PL timaadeen</td>
<td>(You(PL)) came</td>
<td>2PL yimid</td>
<td>(You(PL)) came</td>
</tr>
<tr>
<td>3PL yimaadeen</td>
<td>(They) came</td>
<td>3PL yimid</td>
<td>(They) came</td>
</tr>
</tbody>
</table>

---

11 In my data, the 2SG restrictive form is *timid*, the same as the 3SG.F form. It is probable that the choice between the two forms is dependent on the gender of the hearer.
When the subject of a sentence is focused, the restrictive paradigm is used (as in example 16a and 16c). Otherwise, i.e. when the subject is not focused, the extensive paradigm is used (see examples 16b and 16d) (Livnat 1984:61).

The pattern of verb selection is exemplified in example 16:

(16) a. *Baabuurra*-d-*ii* ayya *i= dhaaf-*ay.*
    cars-F-DEF  FOC 1SG(ACC) passed-REST.
    ‘The CARS passed me.’

    b. *Baabuurra*-d-*ii* anig-*a* ayya *i= dhaaf-*een.*
    cars-F-DEF  me-ACC FOC 3PL(NOM) 1SG(ACC) passed-3PL.EXT.
    ‘The cars passed ME.’  (Saeed 1984:85)

    c. *Idink*-a ayya y-imid.12
       you(PL)-ACC FOC REST-came.
       ‘YOU (PL) came.’

    d. *Iyag*-u waay istaag-*een.*
       waa=ay
       they(NOM) FOC 3PL(NOM) stood.up-3PL.EXT.
       ‘They STOOD UP.’

From this distribution, it is apparent that the pragmatic status of the subject (i.e. whether or not it is focused) determines the conjugation of the verb. The subject/actor of 16a and 16b are identical, so the change in the verb paradigm cannot be attributed to a change in either the syntactic role of subject or the semantic role of actor. The choice of which verb paradigm to use is determined by the pragmatic status of the subject, i.e. whether it is focused or not.

Livnat (1984:iv) proposes that the use of the restrictive paradigm in sentences with focused subjects means that the subject is no longer the ‘grammatical subject’

---

12 The use of the accusative pronoun in this position will be explained in Section 2.1.3.
(similar to the RRG concept of syntactic subject) of the sentence, because in her view it can no longer trigger verb agreement. This analysis would be plausible if the restrictive form only had one verb form, but there are three forms, and the choice among them does depend on the person and number of the subject. So, for example, in 17 the first person plural pronoun selects the 1PL in the restrictive paradigm.

(17)  
\begin{verbatim}
Annaga  ayaa n-imid.
We(ACC) FOC 1PL-came.REST
'WE came.'
\end{verbatim}

In summary, verb agreement IS governed by the subject, and there is no reason not to consider it the grammatical subject of the sentence. When the subject is focused, fewer distinctions are made in the system of verbal agreement.

2.1.3 Case Marking

In Somali, there are two main cases: NOMINATIVE and ACCUSATIVE. These cases are indicated by the nominal suffixes \(-u\) (nominative) and \(-a\) (accusative). In general, the nominative case is used for subjects and the accusative case for non-subjects, including direct and indirect objects and other obliques. However, case marking is also influenced by focus structure, similar to the effects of focus structure on verb agreement and clitic pronoun occurrence.

When the subject of a sentence is not focused, it receives the nominative case, as expected.

\footnote{A third case, genitive, bears no relevance to this discussion. In addition, the definite affix \(-ii\) appears on nouns which are remote from the speaker (Saeed 1987:134), and is used for both nominative and accusative cases.}
(18) *Adig-u waad istaag-tay.*

You(SG)-NOM FOC-2SG(NOM) stood.up-2SG

‘You(SG) STOOD UP.’

However, when the subject is focused, it receives the accusative case. The object NP always receives the accusative case, regardless of whether or not it is focused. When neither the subject nor the object is focused, the subject receives the nominative case and the object receives the accusative case.

2.2 Analysis of Somali focus structure

As exemplified in Section 2.1, there are two focus markers which occur in Somali, *baa* and *waa*. These focus markers are used in three types of constructions, which are exemplified below:

2.2.1 The *baa* construction

The first type of construction which is used to indicate focus in Somali consists of an NP or oblique followed by *baa*. This construction is exemplified by examples 11 and 12 above, which are repeated as 19 and 20 below:

(19) *Anig-a ayay Maryan i= ga raay-say.*

ayaa=ay me-ACC FOC 3SG.F(NOM) Maryan 1SG(ACC) defeated-3SG.F

‘Maryan hit ME.’

(20) *Niman ayaa kaa raay-ay adig-a.*

ku= ka men FOC 2SG(ACC) defeated-3PL.REST you-ACC

‘The MEN defeated you.’

In example 19, the focused constituent is *aniga*, and focus is indicated by the marker *ayaa* which follows it. Example 21 shows that no constituent may occur between the coreferential proclitic pronoun *i* and the verb.
2.2.2 The waxaa construction

The second construction which indicates focus in Somali is the waxaa construction.

Waxaa is a fused form, consisting of waxa, a pronoun that can refer to a person or an object, followed by baa.\(^\text{15}\) Waxa is coreferential with a lexical NP that occurs postverbally (in the PoCS or in the periphery) shown in the following examples:

\begin{align*}
(22) \text{Cali} & \text{ wuxuu} & \text{ ark-ay} & \text{ Maryan.} \\
& \text{ waxa+ baa=uu } & & \\
& \text{ Ali one FOC 3SG.M(NOM) saw-3SG.M Maryan } & & \text{‘It was MARYAN whom Ali saw.’} \\
(23) \text{Maryan} & \text{ waxaa} & \text{ ark-ay} & \text{ Cali.} \\
& \text{ waxa+ baa } & & \\
& \text{ Maryan one FOC saw-3SG.M Ali } & & \text{‘It was ALI who saw Maryan.’}
\end{align*}

---

\(^{14}\) In the examples, a pound sign (#) indicates that a response is inappropriate for the given question, and a single star (*) indicates that the response is ungrammatical in the language.

\(^{15}\) When a word ending in /a/ or /o/ is followed by the focus marker baa, which ends in a long high vowel, the initial consonant of baa is deleted along with the final vowel of the first word (/x/ is a voiceless pharyngeal fricative):

\[
\text{waxa + baa} \rightarrow \text{waxaa}
\]

\text{(Livnat 1984:30)}
As seen in examples 22 and 23, the NP which is in the PoCS may be either the subject or object of the sentence. Waxa may also be coreferential with an oblique or adpositional phrase, occurring in the periphery. These are analyzed in 2.2.6.

2.2.3 The waa construction

The particle waa is a marker of predicate focus. The predicate may consist of either a verb (example 24), a nominal predicate (in the absence of a verb, as in example 25), or an adjectival predicate (plus copula) (example 26). Waa may focus the predicate in transitive sentences (examples 27 and 28), intransitive sentences (example 24), or verbless equatives (example 25). A clitic pronoun, coreferential with the topic of the sentence, may occur in the waa construction, as shown in 24 and 26-27.\(^{16}\) Note that the object cannot be the topic of the sentence, since it is part of the predicate which is in focus (Lambrecht 1994:20). The enclitic pronoun may not occur in a sentence such as 25, for reasons which will be explored later.

(24) \textit{Isag-u wuu istaag-ay.} \\
\hspace{1cm} \text{he-NOM FOC 3SGM(NOM) stood.up-3SG.M} \\
\hspace{1cm} \text{‘He stood up.’}

(25) \textit{Cali waa macallin.} \\
\hspace{1cm} \text{Ali FOC teacher} \\
\hspace{1cm} \text{‘Ali is a TEACHER.’ (Livnat 1984:97)}

(26) \textit{Cali wuu weyn yahay.} \\
\hspace{1cm} \text{Ali FOC=3SG.M(NOM) big be.3SG.M} \\
\hspace{1cm} \text{‘Ali is BIG.’} \\
\hspace{1cm} \text{(Ajello 1984:85)}

\(^{16}\) Recall that the third person accusative pronoun is null; thus, even if the topic is the object, no pronoun occurs (as in 28).
A sentence such as example 28 occurs in response to a question in which there is no constituent which could function as topic; e.g. What happened to the woman? The man DEFEATED the woman. ‘Man’ could not function as the topic in the above response since it is not presupposed in the question.

The following section presents an analysis of Somali focus structure using Lambrecht’s paradigms of narrow focus, predicate focus, and sentence focus. Furthermore, I differentiate between two types of narrow focus: subject and object. I also consider the differences in focus between transitive and intransitive sentences.

2.2.4 Narrow focus: subjects

Intransitives

To determine the manifestation of narrow focus in an intransitive sentence, I asked an intransitive question in which the single NP argument (i.e. the subject) was the missing argument. In intransitive sentences where the single NP argument is focused by baa, no coreferential pronoun clitic may occur, because no topic or object occurs in the sentence.
The intransitive question and its preferred responses (in descending order of preference) are given in example 29:  

(29) a. Yaa y-imid?
yaa+baa
who FOC 3SG.M-came
‘Who came?’ (Example 29 adapted from Saeed 1984:102)

b. Nink-ii baa y-imid.
man-DEF FOC 3SG.M-came
‘The MAN came.’

c. Waxaa y-imid nink-ii.
waxa-baa=uu
one FOC 3SG.M(NOM) 3SG.M-came man-DEF
‘The MAN came.’

d. *Nink-ii buu y-imid.
baa=uu
man-DEF FOC 3SGM(NOM) 3SG.M-came
‘The MAN came.’

Example 29 shows that an intransitive WH-question which contains a subject as the missing argument triggers a response in which the subject is focused, a predictable response according to Lambrecht’s theory of information structure. Both the baa and the waxaa constructions are acceptable responses to 29a. 29d, however, is an unacceptable response because the nominative clitic pronoun uu cannot be coreferential with ninkii, the focused NP. The coreferential nominative pronoun can only occur when the subject is not focused.

17 Every WH-word (yaa in example 29a) is followed by the focus marker baa, and adheres to the phonological coalescence rule in footnote 15.
The question presented in 30a is also intransitive with the subject as the missing argument, but its purpose is to determine whether or not there are differences in focus structure based on the semantic role of the missing argument, i.e. actor or undergoer. The missing argument of the question in 29a is an actor, while the missing argument in 30a is an undergoer. In examples 30b-d, where the focused constituent is an undergoer, the focus marking is identical to that in 29b-d, where the focused constituent is an actor. Therefore, the pattern of focus marking in \textit{baar} constructions does not differ according to the semantic macrorole (i.e. actor or undergoer) of the focused NP.

(30) a. \textit{Yaa dhac-ay?}
   \textit{yaa + baa}
   who fell.down-3SG.M
   ‘Who fell down?’

b. \textit{Nink-ii baa dhac-ay.}
   \textit{man-DEF FOC fell.down-3SG.M}
   ‘The \textbf{MAN} fell down.’

c. \textit{Waxaa kuf-ay nink-ii.}
   \textit{waxa + baa}
   one FOC fell.down-3SG.M man-DEF
   ‘The \textbf{MAN} fell down.’

d. *\textit{Nink-ii buu dhac-ay.}
   \textit{baa = uu}
   man-DEF FOC 3SG.M(S) fell.down-3SG.M.
   ‘The \textbf{MAN} fell down.’

Figure 6 shows the actual and potential focus domains for example 29b. Recall from Chapter 1 that the actual focus domain (represented by the gray shaded triangle shown in Figure 6) is the syntactic domain which is actually in focus, while the potential focus domain (represented by the dotted lines) is the syntactic domain in which focused constituents may occur. In Somali, the actual and potential focus domain for intransitive sentences in which the subject is focused by \textit{baar} is the first constituent in the core; this is
the only possible focus position. Therefore, in a *baa* construction, the actual and potential focus domains always coincide. In Figure 6, *ninkii* is a core argument of the verb.

Figure 6. Narrow Focus (Intransitives)

Transitives

Narrow focus on the subject, indicated by *baa*, also occurs in response to transitive questions in which the missing argument in the WH-question is the subject. The purpose of example 31 is to show that narrow focus on the subject is manifested identically for both intransitive and transitive sentences.  

---

18 Asking native speakers to rank the preferred responses to WH-questions seemed to be most accurate for the top one or two choices. The preferred order among the lower-ranked choices was harder to pinpoint. Judgments concerning the inappropriateness of
responses in order of their preference by a native speaker (recall that baa and ayaa are optional variants):

(31) a.  \[
\text{Yaa} \quad \text{ark-ay} \quad \text{Maryan?} \\
\begin{array}{c}
yaa+baa \\
\text{who FOC saw-3SG.M Maryan} \\
\text{'WHO saw Maryan?'}
\end{array}
\]

(Example 31 adapted from Livnat 1983:93)

b.  \[
\text{Cali ayaa} \quad \text{ark-ay} \quad \text{Maryan.} \\
\text{Ali FOC saw-3SG.M Maryan} \\
\text{'ALI saw Maryan.'}
\]

c.  \[
\text{Cali ayaa Maryan ark-ay.} \\
\text{Ali FOC Maryan saw-3SG.M} \\
\text{'ALI saw Maryan.'}
\]

d.  \[
\text{Maryan Cali baa ark-ay.} \\
\text{Maryan Ali FOC saw-3SG.M} \\
\text{'ALI saw Maryan.'}
\]

e.  \[
\text{Maryan waxaa ark-ay Cali.} \\
\begin{array}{c}
waxa+baa \\
\text{Maryan one FOC saw-3SG.M Ali} \\
\text{'It was ALI who saw Maryan.'}
\end{array}
\]

f.  \[
\#Cali wuxuu ark-ay Maryan. \\
\begin{array}{c}
waxa+baa=uu \\
\text{Ali one FOC 3SG.M(NOM) saw-3SG.M Maryan} \\
\text{'The one that Ali saw (was) MARYAN.'}
\end{array}
\]

g.  \[
\#Cali Maryan wuu ark-ay. \\
\begin{array}{c}
waa=uu \\
\text{Ali Maryan FOC 3SG.M(NOM) saw-3SG.M} \\
\text{'Ali SAW Maryan.'}
\end{array}
\]

sentences seemed more difficult than judgments concerning the ungrammaticality of sentences.
Example 31 shows that a transitive WH-question in which the subject is the missing argument triggers a transitive response in which the subject is focused, similar to the intransitive question/answer pair in example 29. 31b is preferred over 31c because it retains the same word order as the WH-question to which it responds. In Somali, the preferred response is always the one which retains the same word order as the WH-question to which it responds. However, the additional word orders presented in 31c and 31d are also acceptable, because the focused constituent still occurs in the unmarked focus position, i.e. the first position in the core. In 31b-d, the focused constituent remains in the core-initial position. This is the only focus position possible for focused NPs or obliques in the baa construction. Example 31d provides evidence to posit that Maryan is in the PrCS, because Maryan is not set off by a pause. 31e is possible because it focuses the appropriate constituent, which is the subject.

31f and 31g are unacceptable responses to 31a because the focus is not on the appropriate constituent. In 31f, the focused constituent is the object, which is presupposed in the question. In 31g, the predicate is in focus, which is not appropriate because it is presupposed in the question as well. Figure 7-9 represent 31b-d, showing that even though word order is variable in these constituents, the focused constituent remains in the same position. In Figure 7, Cali is the focused constituent and occurs as the first constituent in the core, which is the unmarked focus position.
In Figure 8, *Maryan* occurs in the core and is therefore both a syntactic and a semantic argument of the verb.
The focused constituent *Cali* remains in the same position in Figure 7-9, despite the variation in word order. The actual and potential focus domains for Figure 8 and Figure 9 are identical to those in Figure 7.

The second method of indicating narrow focus is with the *waxaa* construction, as shown above in example 31e. *Waxa* is a pronoun which can refer to a person or an object. In a *waxaa* construction, *waxa* is focused by *baa*. *Waxa* and *baa* coalesce to form *waxaa* (see footnote 9). Any NP which is coreferential with *waxa* always occurs postverbally, in the PoCS. Adpositional phrases coreferential with *waxa* may occur in the periphery.

Word order is also variable in a *waxaa* construction; any non-focused NP may occur in the PrCS, LDP, or core-internally. I propose that *waxa* always occurs core-initially, and there are three facts which support this analysis: 1) *Waxa* cannot be analyzed as occurring in the PrCS, because any NP that precedes it is not set off by a pause. 2)
When *waxa* appears in a WH-question, it may be preceded by two NPs, one in the PrCS, and one in the LDP. 3) Other NPs focused by *baa* are also core-initial, and *waxa* is focused by *baa*.

Figure 10 represents a *waxaa* construction in which the subject is focused. Again, the potential and actual focus domains coincide. The unmarked focus position for a *waxaa* construction is the same as that for a *baa* construction: the first position in the core.

Examples 31b-d give evidence to show that the focused constituent is in the core-initial position in *baa* sentences, and since *waxaa* constructions are a type of *baa* construction where *waxa* is focused by *baa, waxa* is also core-initial. WH-words are assigned to the core-initial position because they are focused with *baa* as well. Therefore, the unmarked focus position in Somali, for the *baa* and *waxaa* constructions, is the core-initial position.
2.2.5 Narrow focus: objects

Intrasentential pronominalization

In a sentence in which a pronominal, non-third person object is focused, both the subject, which is the topic, and the object appear with coreferential clitic pronouns, as shown in example 32:

(32) Anig-a ayay Maryan i= ga raay-say.
me-ACC FOC 3SG.F(NOM) Maryan 1SG(ACC) defeated-3SG.F
'Maryan defeated ME.'

However, in sentences where the object is a focused, lexical (i.e. third person) NP, there is no object clitic.

(33) Niman baan ka raay-ay anig-u.
men FOC 1SG(NOM) defeated-3-SG.M I-NOM
'I defeated the MEN.'

The absence of the accusative clitic in example 33 can be explained by the RRG principle that any lexical NP which occurs with a coreferential clitic pronoun must be outside the actual focus domain. This follows from the principle governing intrasentential pronominalization presented in RRG (VanValin and LaPolla 1997:224). Intrasentential pronominalization refers to the relationship between a lexical NP and a pronoun within the same sentence. One of the problems of intrasentential pronominalization is determining whether or not there is coreferentiality between the lexical NP and the pronoun in the sentence. In the following English sentence, for example, a coreferential interpretation between the possessive pronoun and the lexical NP is possible, but not necessary.

(34) Her mother loves Mary.
In order for coreferentiality to exist, either the lexical NP or the pronoun must occur outside the core (i.e. both constituents cannot occur in the core), and the lexical NP must be outside the actual focus domain. The principle governing intrasentential pronominalization is stated as follows:

*Coreference is possible between a lexical NP and a pronoun within the same sentence if and only if the lexical NP is outside of the actual focus domain.* (Van Valin and LaPolla 1997:224).

This principle can be directly applied to Somali. As shown in example 35, the focused subject, *niman*, does not occur with a coreferential pronoun. In example 36, however, with focus on the object, the nominative pronoun, *ay* can occur coreferentially with the subject *niman*. Clearly, Somali adheres to the above principle of intrasentential pronominalization.

(35)  
\[ \text{Niman ayaa ku=} \text{garaac-ay adig-a.} \]
\[ \text{men FOC 2SG(NOM) defeated-3PL.REST you-ACC} \]
\[ \text{‘The MEN defeated you.’} \]

(36)  
\[ \text{Nimanku waxay ka raay-een adig-a.} \]
\[ \text{waxa+baa=} \text{ay} \]
\[ \text{men.DEF one FOC 3SG.M(NOM) defeated-3PL you-ACC} \]
\[ \text{‘The one who the men defeated (is) YOU.’} \]

In the typical sentence involving intrasentential pronominalization, the pronoun is within the actual focus domain, while the lexical NP is outside of it. This appears to be true in Somali. As seen in example 36, the coreferential pronoun clitic *ay* immediately follows the focus marker *ayaa*, and is cliticized to it. Constraints on coreference, therefore, are not merely governed by syntax, but are also affected by pragmatic considerations as well.
Object focus

In 37a, the missing argument of the transitive verb is the object. Examples 37b-e show the possible responses (in descending order of preference) to the WH-question in 37a. These examples show how narrow focus on objects in Somali is indicated, and how the representation of narrow focus in Somali is consistent with Lambrecht’s principles, as outlined in Chapter 1 of this thesis.

(37) a. Yuu ark-ay Cali?
yaa+baa=uu  
who FOC 3SG.M(NOM) saw-3SG.M Ali  
‘Who did Ali see?’ (Example 37 adapted from Livnat 1983:93)

   b. (Cali) wuxuu ark-ay Maryan.
waxa+baa=uu  
(Ali) one- FOC -3SG.M(NOM) saw-3SG.M Maryan.  
‘The one whom he (Ali) saw was MARYAN.’

   c. Maryan buu ark-ay Cali.
   baa=uu  
Maryan FOC-3SG.M(NOM) saw-3SG.M Ali.  
‘Ali saw MARYAN.’

   d. Cali Maryan buu ark-ay.
   baa=uu  
Ali Maryan FOC-3SG.M(NOM) saw-3SG.M Maryan.  
‘Ali saw MARYAN.’

   e. ?Maryan buu Cali ark-ay.
   baa=uu  
Maryan FOC-3SG.M(NOM) Ali saw-3SG.M Maryan.  
‘Ali saw MARYAN.’

The responses in 37b-d are similar to those in 31b,d, and e, except that the focused constituent in 37b-e is the object instead of the subject. Unlike its counterpart in 31c, however, 37e appeared to lead to confusion about the syntactic roles of the argument NPs, although it was accepted as the preferred answer to 37a in another elicitation session. It seems that 37e may be problematic because the lexical NP Cali is located in
between the coreferential pronoun \textit{uu} and the verb. This dispreference follows from the fact that in 37e, both \textit{Cali} and \textit{uu} are eligible to function as core arguments of the verb. Apparently, it is dispreferred that two coreferential arguments of the verb should both occur in the core. Previous analyses (Livnat 1984, Saeed 1984, Svolacchia 1995 et al) have stated that the pronoun clitic \textit{uu} is optional in 37e, but not in 37b-d. The dispreference for two coreferential arguments in the core partially explains that optionality.

There are some differences among the responses to 31a and 37a. \textit{Waxaa} constructions are much more commonly used for object focus than they are for subject focus. Thus, 37b is the preferred answer to 37a, while its counterpart in 31e is acceptable but not preferred, possibly because the focused subject in 31e does not, and cannot, have a coreferential pronoun in the core, so its role in the sentence is less easily determined.

In \textit{waxaa} constructions, \textit{waxa} is always the focused constituent. The potential focus domain for \textit{waxaa} constructions is the initial position in the core, and the actual focus domain is the same, providing evidence for the rigidity of focus structure in Somali. In fact, the rigidity of focus structure in Somali corresponds with its relative flexibility in word order. The existence of languages with flexible word order and rigid focus structure is typologically attested to by Van Valin (1999:4).

Figure 11 shows the tree diagram for 37b, a \textit{waxaa} construction with object focus.
Figure 11. Narrow focus with *waxaa* (object)

The tree diagram in Figure 12 represents 37c. The same variation in word order is possible for in 37c-e as in 31b-d (represented in Figure 7-9) though I am only showing 37c here.

Figure 12. Narrow focus (object), for 37c
2.2.6 Predicate focus

In Section 2.1, I indicated that the focus marker *waa* indicates predicate focus. I used the following WH-question/answer pairs, adapted from previous Somali research (Livnat 1984:100) and from Lambrecht’s (1994) paradigms, to determine the scope of *waa* focus in Somali. This particular paradigm was used by Livnat (1984:100) to elicit ‘verbal’ focus (equivalent to Lambrecht’s term *predicate focus*) in Somali:

(38) a.  
\[\text{Muxuu} \quad \text{Cali samey-ay?} \]
\[\text{ma waxa+baa} \quad =uu^{19} \]
\[\text{INT what} \quad \text{FOC} \quad 3\text{SG.M} \quad \text{Ali do-3SG.M} \]
\[\text{‘What did Ali do?’} \]

The preferred response to 38a is given in 38b.

b.  
\[\text{Cali wuxuu} \quad \text{ka raay-ay} \quad \text{Maryan.} \]
\[\text{waxa+baa=uu} \]
\[\text{Ali one} \quad \text{FOC} \quad 3\text{SG.M(NOM)} \quad \text{defeated-3SG.M} \quad \text{Maryan.} \]
\[\text{‘Ali defeated MARYAN.’} \]

c.  
\[\text{Cali Maryan buu} \quad \text{ka raay-ay.} \]
\[\text{baa=uu} \]
\[\text{Ali Maryan FOC} \quad 3\text{SG.M(NOM)} \quad \text{defeated-3SG.M} \]
\[\text{‘Ali defeated MARYAN.’} \]

d.  
\[\text{Maryan buu} \quad \text{ka raay-ay} \quad \text{Cali.} \]
\[\text{baa=uu} \]
\[\text{Maryan FOC} \quad 3\text{SG.M(NOM)} \quad \text{defeated-3SG.M} \quad \text{Ali.} \]
\[\text{‘Ali defeated MARYAN.’} \]

e.  
\[\#\text{Cali bua} \quad \text{Maryan ka raay-ay.} \]
\[\text{Ali FOC Maryan defeated-3SG.M} \]
\[\text{‘ALI defeated Maryan.’} \]

f.  
\[\#\text{Cali bua} \quad \text{ka raay-ay} \quad \text{Maryan.} \]
\[\text{Ali FOC defeated -3SG.M Maryan} \]
\[\text{‘ALI defeated Maryan.’} \]

---

19 This analysis of muxuu is taken from Svolacchia et al. (1995:95).
g.  #Cali Maryan wuu ka raay-ay.
              waa=uu
    Ali Maryan FOC-3SG.M(NOM) defeated-3SG.M
        ‘Ali DEFEATED Maryan.’

h.  #Cali Maryan waa ka raay-ay.
    Ali Maryan FOC defeated-3SG.M
        ‘Ali DEFEATED Maryan.’

Question 38a, which Livnat intended to elicit a response with predicate focus, actually elicits a response with narrow focus on the object. Focus on the subject is not permitted in response to 38a, because the subject is presupposed in the question. Unexpectedly, predicate focus is not allowed either, even though the predicate is not presupposed in the question.

However, an intransitive response to 38a, as shown in 38i-k, has predicate focus, not narrow focus, since there is no object to receive narrow focus, and the subject, the single NP argument, is presupposed in 38a and cannot receive focus. The predicate is the only constituent which is not presupposed, and therefore the only candidate eligible for focus. In 38i-k, the single NP argument of each sentence is the actor and cannot receive focus, while 38i-n represent sentences in which the single argument is an undergoer and cannot receive focus (38i, j, l and m are equally preferred as responses to 38a):

(38)  i. Cali waa qosl-ay.
             Ali FOC laughed-3SG.M
         ‘Ali LAUGHED.’

      j. Cali wuu qosl-ay.
             waa=uu
           Ali FOC 3SG.M(NOM) laughed-3SG.M
       ‘Ali LAUGHED.’

      k. #Cali baa qosl-ay.
           Ali FOC laughed-3SG.M
      ‘ALI laughed.’
1. **Cali waa kuf-ay.**
   Ali FOC fell.down-3SG.M
   ‘Ali **FELL DOWN.**’

m. **Cali wuu kuf-ay.**
   waaw=uu
   Ali FOC 3SG.M(NOM) fell.down-3SG.M
   ‘Ali **FELL DOWN.**’

n. **#Cali baa kuf-ay.**
   Ali FOC fell.down-3SG.M
   ‘**ALI** fell down.’

The only difference between 38i and j is that a topic enclitic occurs in 38j but not in 38i. Likewise, one occurs in 38m but not 38l. The choice to consider *Cali* the topic lies with the speaker. When *Cali* is the topic, its coreferential enclitic pronoun occurs; when it is not, the pronoun does not occur. Thus, the topicality of any particular NP is determined by the presence of an enclitic pronoun, and provides an explanation for the acceptability of 38i, j, l and m in response to 38a.

Question 39 is identical to the one used in Lambrecht (1994:223) to elicit predicate focus. 39b-e show intransitive responses to 39a.

(39) a. **Maxaa Maryan ku dhac-ay?**
   ma+waxa+baa
   INT-thing-FOC Maryan to happened-3SG.M
   ‘What happened to Maryan?’

b. **Maryan way kuf-tay.**
   waaw=ay
   Maryan FOC 3SG.F(NOM) fell.down-3SG.F.
   ‘Maryan **FELL DOWN.**’

c. **Maryan waa kuf-tay.**
   Maryan FOC fell.down-3SG.F
   ‘Maryan **FELL DOWN.**’

d. **#Maryan baa kuf-tay.**
   Maryan FOC fell.down-3SG.F
   ‘**MARYAN** fell down.’
e. #Waxaa kuf-tay Maryan.
   waxa+baa
one FOC fell.down-3SG.F Maryan
‘The one who fell down (was) MARYAN.’

The only possible response to 39a is the waa construction, shown in 39b and c, with or without a topic enclitic. This construction indicates predicate focus. 39b and c are preferred because the predicate is focused; it is the only constituent eligible for focus in these sentences. Both 39d and 39e are inappropriate responses; Maryan cannot be focused because it is presupposed in question 39a.

When the response to 39a is a transitive sentence, the preferred response is narrow focus on the subject, although predicate focus is also permitted, as shown in examples 39f-j:

(39)  f. Maryan waxaa ka raay-ay Cali.
   waxa+baa
Maryan one FOC defeated-3SG.M Ali.
‘The one that defeated Maryan was ALI.’

 g. Cali baa Maryan ka raay-ay.
   Ali FOC Maryan defeated-3SG.M
‘ALI defeated Maryan.’

 h. Cali baa ka raay-ay Maryan.
   Ali FOC defeated-3SG.M Maryan
‘ALI defeated Maryan.’

 i. Maryan Cali baa ka raay-ay.
   Maryan Ali FOC defeated-3SG.M
‘ALI defeated Maryan.’

 j. Cali, Maryan wuu ka raay-ay.
   waa=uu
Ali, Maryan FOC 3SG.M(NOM) defeated-3SG.M
‘Ali DEFEATED Maryan.’
39f-j are listed in descending order of preference. Therefore, 39a evokes predicate focus for intransitives, and narrow focus on the subject or predicate focus for transitives. In 39b-c, Maryan is the subject, and cannot be focused because it is presupposed in question 39a. In 39f-j, however, Maryan is the object and the subject Cali is new information. Cali is therefore eligible for focus; narrow focus is apparently preferred over predicate focus in this instance. Although it is used by Lambrecht, question 39a is not an ideal question to elicit predicate focus in transitive responses, because there are two constituents which are eligible for focus: Cali and the predicate. Apparently, when there are two constituents eligible for focus, narrow focus is preferred over predicate focus.

The only type of question in Somali which consistently elicits predicate focus is one in which all NP arguments that occur in the response are presupposed in the WH-question, as shown in example 40. In example 40a, waxa functions as the WH-word, and ma provides interrogative illocutionary force. This analysis is consistent with the fact that waxa always occurs in the core-initial position in a declarative sentence, and that all WH-words occur in the core-initial position as well. Waxa must be analyzed as occurring in the core-initial position in 40a, because it is preceded by two NPs. Cali is in the LDP, and Maryan is in the PrCS.

Question 40a isolates the predicate as the only non-presupposed constituent in the WH-question, and therefore the only constituent eligible for focus.
Figure 13. Question eliciting predicate focus.

(40) a. *Cali Maryan muxuu ku samey-ay?*

\[
muxuu = \text{Cali Maryan}
\]

\[
\text{ma+waxa+baa= uu}
\]

Ali Maryan INT one FOC 3SG.M(NOM) to do-3SG.M

'WHAT did Ali do to Maryan?'

b. *Wuu ka raay-ay.*

\[
\text{waa= uu}
\]

FOC 3SG.M(NOM) defeated-3SG.M

'He **DEFEATED** (her).'

c. *Waa ka raay-ay.*

FOC defeated-3SG.M

'He **DEFEATED** (her).'

---

20 Native speaker judgment varied concerning the presence of the pause following the left-detached element; at best, the presence of the pause is marginal. I retain the analysis of the focused NP in the core-initial position, and *Cali* in the LDP, since it is most consistent with the grammar of the language.
40b is the preferred response because it does not repeat the NPs (Cali and Maryan) given in 40a. This follows from the concept of topic acceptability in RRG (Van Valin and LaPolla 1997:204-205), which states that an active referent (a description applying to both NP arguments in 40a) is most likely to occur as a topic. The topic is most likely to be coded as a zero morpheme, and secondly as a clitic/bound pronoun. The topic Cali is coded as a clitic pronoun in 40b. The subject is the default choice for topic when more than one possibility exists.

Any NP which occurs in the WH-question is optional in the response to that question, although it is still unusual that the object Maryan has no referent in responses 40b-c. They are only appropriate in response to a WH-question such as 40a, in which the argument Maryan is a presupposed, active referent and therefore does not have to be repeated in the response. (Although presupposed, Maryan cannot be considered the topic in 40b or 40c, because no coreferential topic pronoun occurs.) In general, the most preferred responses to WH-questions are those which repeat the least number of arguments.

Examples 40d-f are also acceptable responses to 40a, though they are less preferred than 40b-c.

(40)  

d. Cali, Maryan wuu ka raay-ay.  
   waa=uu
   Ali Maryan FOC-3SG.M(NOM) defeated-3SG.M
   ‘Ali DEFEATED Maryan.’

e. Cali, Maryan waa ka raay-ay.  
   Ali Maryan FOC defeated-3SG.M
   ‘Ali DEFEATED Maryan.’

f. Maryan, Cali waa ka raay-ay.  
   Maryan Ali FOC defeated-3SG.M
   ‘Ali DEFEATED Maryan.’
In 40d, the potential and actual focus domain is the predicate, as shown in Figure 14. The clitic pronoun \textit{uu} is a syntactic argument of the verb, because its referent \textit{Cali} is located in the LDP.

In an intransitive sentence with predicate focus, the coreferential topic pronoun may or may not occur. The sole NP in an intransitive sentence may function as the topic of the sentence, because it is presupposed in the question, as in example 41, or for discourse reasons requiring further exploration, it may not be considered the topic by the speaker, and the coreferential clitic pronoun will be omitted, as in example 42.

\begin{enumerate}
\item \textit{Cali wuu qosl-ay.}
\item \textit{Ali FOC 3SG.M(NOM) laughed-3SG.M ‘Ali LAUGHED.’}
\end{enumerate}

\textsuperscript{21} The focus marker \textit{waa} is not a constituent in the layered structure of the clause (LSC); it belongs only to the focus structure projection [indicated by the triangles below the sentence].
The following points summarize the responses given in 31-40:

1) For *waxaa* constructions, the actual focus domain is the core-initial position, because *waxa* is considered the focused constituent, since it is followed by the focus marker *baa*\(^ {22} \). It is coreferential with a lexical NP that occurs in the PoCS (or as shown in section 2.2.6, with an adpositional phrase in the periphery). The PoCS in Somali is reserved for the coreferential lexical NP in a *waxaa* construction; no other constituents will occur there in any construction. Any constituent occurring in the PoCS must be a semantic argument of the verb. In the *waxaa* construction, *waxa* receives primary focus (because it is followed by *baa*) and the constituent in the PoCS receives secondary focus. However, section 2.2.6 demonstrates that the PoCS is not the only possible position for constituents receiving secondary focus.

2) For *baa* constructions, the potential and actual focus domains are the core-initial position. If present, the enclitic pronoun that is coreferential with the topic immediately follows *baa*.

3) For *waa* constructions, the potential and actual focus domains are the predicate. If present, the enclitic pronoun coreferential with the topic immediately follows *waa*. The predicate may include a verb, or may be entirely nominal.

There are only two possible focus positions in Somali: the core-initial position and the predicate. In conclusion, although word order in Somali is relatively flexible, focus

---

\(^ {22} \) I thank Dan Everett for suggesting this possible analysis.
structure is rigid. This concurs with typological investigations on the interaction of syntax and focus structure (Van Valin and LaPolla 1997:213). In many languages, focus structure and syntax adapt to each other – if one is rigid, the other is relatively free. This is true for Somali.

2.2.7 Sentence focus

The example given in Lambrecht (1994) to elicit sentence focus is ‘What happened?’ I tested the equivalent of that question in Somali:

(43) a. *Maxaa dhac-ay?*
   What happened-3SG.M
   ‘What happened?’

   As shown in examples 43b-k, the above question elicits two possible responses in Somali: narrow focus on the subject or narrow focus on the object. It seems that the speaker was able to choose either NP to receive narrow focus, and there is no clear manifestation of sentence focus.

   b. *Nink-ii baa kuf-ay.*
      man-DEF FOC fell.down-3SG.M
      ‘The MAN fell down.’

   c. *Cali baa Maryan ka raayay.*
   d. *Cali baa ka raayay Maryan.*
   e. *Maryan, Cali baa ka raayay.*
      ‘ALI defeated Maryan.’ (translation for 43c-e)

   f. *Maryan baa Cali ka raayay.*
      ‘Ali defeated MARYAN.’ (translation for 43f)

   g. *Ninkii baa naagtii ka raayay.*
   h. *Ninkii baa ka raayay naagtii.*
   i. *Naagtii, ninkii baa ka raayay.*
      ‘The MAN defeated the woman.’ (translation for 43g-i)

      ‘The man defeated the WOMAN.’ (translation for 43j)
k. *Ninkii wuxuu qaadey naagtii.*
   ‘The man lifted the **WOMAN.**’ (translation for 43k)

In each of the grammatical responses, the focused NP is found core-initially, which is the universally unmarked focus position in Somali. In response to 43, the most frequently focused constituent was the subject. (Narrow focus on the subject is the most commonly used method, cross-linguistically, to indicate sentence focus.) However, the object could also be focused in response to 43a, and there is no consistent response to 43a. Therefore, the paradigm adapted from Lambrecht does not elicit sentence focus in Somali, and another strategy is needed.

2.2.8 Yes/No Questions

Narrow focus on a constituent in Somali is indicated by the focus marker *baa* following that constituent. Yes/no questions provide further confirmation for this analysis of narrow focus in Somali. In 44a-e, I asked a yes/no question with two NP arguments, and then replaced each of the arguments and the verb in separate responses, hoping to receive identical answers to those that I received for WH-questions with missing arguments. By replacing one constituent in the question by its counterpart in the response, my goal was to elicit narrow focus on the subject, object, and the verb. I hoped that these responses would compare with the corresponding examples above that give narrow focus on each of these constituents. For each response, I am listing two possible answers:

(44) a. *Cali miyuu*  
   *ark-ay*  
   *Maryan*

   *ma+ayaa=uu*  
   *Ali INT FOC 3SG.M(NOM) saw-3SG.M. Maryan*

   ‘Did Ali see Maryan?’
b.  (1) Haa, wuu arkay.
    (2) Haa, (asaga) wuu ark-ay.
        waa=uu
        yes (3SG.M) FOC-3SG.M(NOM) saw-3SG.M
        ‘Yes, (he) SAW (her).’

The confirmation response in 44b displays predicate focus. This shows that
predicate focus is the unmarked choice for focus, which corresponds to the assumption
made by Van Valin and LaPolla (1997:206) that predicate focus is the universally
unmarked type of focus structure.

Replacement of the object predictably elicits object focus.

(44) c.  (1) Maya, Amina buu arkay.
        (2) Maya, (asaga) Amina buu ark-ay.
            baa=uu
            no (3SG.M) Amina FOC-3SG.M(NOM) saw-3SG.M.
            ‘No, he saw AMINA.’

Replacement of the verb elicits both object focus and predicate focus in 44d.

(44) d.  (1) Maya, isagu wuxuu qaadey Maryan.
        ‘No, the one that he lifted was MARYAN.’
        (2) Maya, isagu wuu qaadey Maryan.
            ‘No, he LIFTED Maryan.’

44e shows replacement of the subject, which not surprisingly receives narrow
focus:

(44) e.  Maya, Axmed ayaa ark-ay Maryan.
        no Axmed FOC saw-3SG.M Maryan
        ‘No, AXMED saw Maryan.’

2.2.9 Narrow focus: Predicative and Non-Predicative Adpositional Phrases

In addition to focusing arguments, Somali allows focus on non-argument, oblique
constituents, including NPs and adpositional phrases. RRG distinguishes among three
types of constituents which are neither subjects nor objects.
1) Predicative adpositional phrases (argument-adjuncts (AAJ)): The adposition in argument-adjunct phrases is a predicate; it introduces an argument into the clause and assigns a semantic role to it. Argument-adjuncts represent the intermediate stage between the other two types of constituents, argument-marking and adjuncts. The following example shows a focused argument-adjunct phrase in Somali:

\[(45)\] a. Xagee \textit{buu} \[baa=uu\] Cali \textit{u carar-ay?}
   "Where FOC-3SG.M(NOM) Ali to run-3SG.M" \\
   "Where did Ali run?"

b. Wuxuu \[waxa+baa=uu\] u \textit{carar-ay dukank-ii.}
   place-FOC-3SG.M(NOM) to run-3SG.M store-DEF
   ‘He ran to the \textit{STORE}.’

As Figure 15 shows, the enclitic pronoun \textit{uu} functions as a core argument of the verb, in the absence of the lexical NP \textit{Cali}. \textit{Waxa} is the focused constituent and occurs in the unmarked focus position. The \textit{waxaa} construction is preferred when focusing longer constituents, such as adpositional phrases, because longer constituents are preferred clause-finally.
2) Non-predicative (argument-marking) phrases: The adposition in an argument-marking adpositional phrase introduces an argument into the verb’s logical structure, and they only mark semantic arguments of the verb. In argument-marking adpositional phrases, the adposition basically functions as a case marker (e.g. Robin gave the book TO Maryan). ‘Maryan’ is therefore an indirect core argument of the verb. If an adpositional phrase is non-predicative, its NP must also be able to occur without an adposition. This is the case in example 46:

(46) a. *Yuu siiy-ay Cali buug-ii?*

\[
\text{who FOC 3SG.M(NOM) gave-3SG.M Ali book-DEF}
\]

‘To whom did Ali give the book?’
b. (Cali) wuxuu waxa baa=uu siiy-ay Maryan.

(Ali) one FOC 3SG.M(NOM) gave-3SG.M Maryan 'He gave (it) to MARYAN.'

When the lexical NP is optional or absent (as *Cali* is in 46b), the coreferential enclitic pronoun *uu* functions as a core argument of the verb. This analysis is consistent with Van Valin and LaPolla’s (1997:332) claim that when a lexical NP is omissible in a sentence, any clitic or affix which is coreferential with that lexical NP will function as a syntactic argument of the verb. If present, the lexical N P is a semantic argument of the verb.

Figure 16 exemplifies focus marking of an NP which is a direct core argument of the verb.

Figure 16. Direct Core Argument Focus

As shown in Figure 16 and in previous examples, *waxa* is coreferential with the lexical NP *Maryan*, which is located in the PoCS. *Waxa* must be the syntactic argument
of the verb, because it occurs in the core, while Maryan cannot be the syntactic argument, because it occurs in the PoCS (syntactic arguments, by definition, must occur in the core). This is consistent with RRG’s analysis of bound pronominal affixes as syntactic arguments of the verb (in languages like Lakhota), based solely on their position in the core (Van Valin and LaPolla 1997:331). When independent lexical NPs occur outside the core, their coreferential pronoun clitics are the syntactic arguments of the verb, because syntactic arguments must be core-internal. Evidence for argument status of the clitic pronouns is based on the optionality of the independent NPs. Because the independent NPs are optional and the coreferential pronominal affixes are not, it is more consistent to analyze the clitic pronouns as the syntactic arguments of the verb (Van Valin and LaPolla 1997:331). In summary, the independent lexical NPs are arguments when they occur; in their absence, the clitic pronouns function as arguments of the verb. This phenomenon is commonly referred to as ‘pro-drop’ (Van Valin and LaPolla 1997:331).

Because of the occurrence of waxa in the core, I have analyzed it as a syntactic argument of the verb. Unlike other syntactic arguments, it does not possess any features which could trigger verb agreement, such as person, gender, or number, and its independent coreferential lexical NP is not optional because it provides features for waxa. Though RRG does not directly address the issue of non-lexical NPs such as waxa occurring as syntactic arguments, it does allow non-lexical (e.g. pronominal) NPs or clitics to function as syntactic arguments, and it seems consistent to analyze waxa as a syntactic argument as well. In addition, both Maryan and waxa must fill the same position as semantic arguments of the verb, per RRG’s analysis (Van Valin and LaPolla
1997:332), because they refer to the same entity. Only waxa, however, can function as the syntactic argument, because Maryan is not in the core.

3) Predicative (Adjunct) Adpositional Phrases occur in the periphery and only mark adjuncts (constituents that are not semantic or syntactic arguments of the verb).

\[(47)\]

\[(a)\] Goormuu \(\text{Cali arkay Maryan?}\)
\(\text{goormaa}+\text{baa}=\text{uu}\)
when FOC 3SG.M(NOM) Ali saw-3SG.M Maryan
‘When did Ali see Maryan?’

\[(b)\] Wuxuu \(\text{ark-ay (ayada) aroosk-a kadib.}\)
\(\text{waxa}+\text{baa}=\text{uu}\)
one FOC 3SG.M(NOM) saw-3SG.M her wedding.DEF after
‘He saw (her) AFTER THE WEDDING.’

Figure 17. Peripheral Focus
CHAPTER 3: ALTERNATIVE ANALYSES OF SOMALI FOCUS STRUCTURE

3.1 Overview

Because of the important role of focus structure in relation to Somali syntax, the phenomenon has been analyzed often in the literature. Some of the most influential and insightful analyses are summarized here, for the purpose of comparing their conclusions to those reached through studying Somali focus structure within the framework of RRG.

Two of the earliest analyses of Somali focus structures were those of Andrzejewski (1975) and Hetzron (1965). Along with Hetzron (1974:364), Andrzejewski (1975:23) defines the purpose of *baa*, *waa* and *waxa* in Somali as instances of

"focusing...elevat[ing] the communicational importance of an element above the level of the rest of the sentence."

3.1.1 Saeed (1984)

Saeed’s (1984) work expands previous analyses and presents a comprehensive analysis of focus constructions in Somali within the framework of generative grammar. Unlike other generative analyses, he does not presuppose the existence of a VP. He does, however, adhere to the notion that transformational rules are necessary in order to explain the similarities between *waxaa* and *baa* constructions (Saeed 1984:15). Livnat’s (1984:89) analysis of *waxaa* as a combination of *waxa* and *baa* (see footnote 9) renders any transformational explanation unnecessary. The two constructions show similar characteristics because they both exhibit narrow focus as indicated by the particle *baa*. 

62
Saeed (1984:20) defines focus as the introduction of new information or the marking of ‘old’ constituents as prominent. In RRG, what is new is the relationship between the focused constituent and the proposition in which it occurs, not just the information itself. Because Saeed does not include this relationship in his definition, it is not as precise as the distinctions given in RRG, and therefore difficult for him to present a cohesive analysis of the varying types of focus, as shown in my summary of his analyses below.

In addition, Saeed’s (1984:20) study is concerned mainly with the syntactic description of focus and he clearly states that any references to pragmatic roles and terms in his analysis are incidental. As a result, his study of Somali focus is mainly syntactic, and he does not concentrate on the semantic and pragmatic implications of his analyses. Because the model presented in RRG integrates syntax, semantics, and pragmatics, it provides a framework which allows further insights into the semantic and pragmatic functions of Somali focus structure.

One inconsistency with Saeed’s analysis surfaces with regard to baa-focused constituents. According to him, any NP focused by baa will be in the leftmost position of the ‘sentence proper’ (Saeed’s terminology for a constituent which appears to be equivalent to the clause in RRG). He must posit a ‘sentence proper’ in order to account for left-detached topics, as in example 48, although the term ‘sentence proper’ is not a term allowed within the framework of generative grammar. In order for Saeed’s analysis to be consistent with that of RRG, his definition of the ‘sentence proper’ must be equivalent to the core instead of the clause. The topic of the sentence is frequently in the ‘extra-sentential’ position, which immediately precedes the ‘sentence proper’. This
analysis does not explain how two constituents can precede the focused constituent in sentences such as example 48a, nor does it allow for the ‘sentence proper’ to be equivalent to the core in RRG.

(48) Maryan, Cali baa ka raay-ay.
Maryan, Ali FOC defeated-3SG.M ‘ALI defeated Maryan.’

According to Saeed, a ‘true’ topic is always in the ‘extra-sentential’ (i.e. LDP) or ‘aftertopic’ (i.e. RDP) position. However, his statement that a ‘true’ topic must occur in the LDP indicates that his definition of topic (which is clearly restricted syntactically) differs from that of Lambrecht, whose definition is more pragmatically-oriented. The position of the NP in the sentence does not prevent the occurrence of the enclitic pronoun, which is coreferential with the topic. The topic may occur anywhere in the sentence, as in example 49, where Maryan is the topic.

(49) Aniga ayay Maryan i= ka raay-say.
ayaa=ay me-ACC FOC 3SG.F(NOM) Maryan 1SG(ACC) defeated-3SG.F ‘Maryan hit ME.’

Saeed (1984:80) proposes that any constituent focused by baa is fronted and placed outside the ‘sentence proper’. When the subject of the clause is focused, the clause is therefore syntactically ‘subjectless’ and triggers the use of the restrictive paradigm.

(50) a. Baabuurra-d-ii ayaa i= dhaaf-ay
Cars-F-DEF FOC 1SG(ACC) passed-3PL.REST. ‘The CARS passed me.’

b. Baabuurra-d-ii anig-a ayay i= dhaaf-een
Cars-F-DEF me-ACC FOC 3PL(NOM) 1SG(ACC) passed-3PLEXT. ‘The cars passed ME.’ (Saeed 1984:85)
Focused subjects of *baa* sentences do not receive the nominative case, but the accusative instead (termed ‘absolutive’ by Saeed 1987:128). Therefore, Saeed concludes that the subject, when focused by *baa* in these types of sentences is really a complement, instead of a true subject because it does not have all the characteristics of a ‘true’ subject (i.e. nominative case marking). I conclude, however, that the seemingly unusual case marking actually parallels the system of verb agreement (when the subject is focused, the restrictive paradigm occurs; when the object is focused, the extensive paradigm occurs) outlined in Section 2.1.2. It appears that case marking is at least partially determined by the pragmatic status of the NP, i.e. whether that NP is the focus or the topic of the sentence, just as the choice of verb paradigm is determined by the pragmatic status of the subject. For example, when the subject is the topic of the sentence, it receives the ‘default’ nominative case (the suffix –*u*); when focused, however, it receives a ‘marked’, accusative case (the suffix –*a*). The object always receives the accusative case. For example:

(51) *Anig-a ayaa imid.*
  me-ACC FOC came.1SG.
  ‘I came.’

(52) *Anig-u waan istaag-ay.*
  waa=aan
  me-NOM FOC 1SG(NOM) stood.up-1SG.
  ‘I STOOD UP.’

(53) *Anig-a ayay Maryan i= di-shay.*
  ayaa=ay
  1SG-ACC FOC 3SG(NOM) Maryan 1SG(ACC) beat-3SG.F
  ‘Maryan beat ME.’

Saeed disagrees with Andrzejewski’s (1975) analysis of *waa* as a verbal focus marker. As a result, he also rejects the notion that every sentence must contain a
morphological representation of focus. Instead, he classifies *waa* as an indicator of declarative ‘clause marker’ (apparently equivalent to the RRG notion of illocutionary force and similar to the function of the ‘indicator particles’ described by Andrzejewski (1975)) and rejects the proposal that Somali has verbal focus in any form. He classifies *waa* as illocutionary force, not as a focus marker, mainly because it does not occur in interrogative or imperative sentences, only declaratives.

(54)  \[ \text{Ma Cali } \left[ \begin{array}{c} \text{*waa} \\ \text{baa} \end{array} \right] \text{ weyn ama yar?} \]

\begin{tabular}{lll}
INT & Ali & FOC big or small \\
& & \text{'Is ALI big or small?'}
\end{tabular}

Saeed’s argument that *waa* is only found in declarative sentences does not necessarily exclude it from classification as a focus marker. One would not expect it in WH-interrogative sentences, for example, because they only allow narrow focus on the WH word, and *waa* only indicates predicate focus, which is broad. This does not explain why it is not used in yes-no interrogatives, which permit broad focus but still do not allow *waa*. However, its exclusion from interrogatives does not prevent it from functioning as a focus marker in declaratives.

A basic assumption of RRG is that every sentence contains focus. In Somali, the primary means of indicating focus is morphological and it follows that every declarative sentence should contain either a morphological or prosodic indicator of focus. Saeed rejects this basic principle of information structure when he analyses *waa* as a ‘classifier’ (i.e. a marker of declarative illocutionary force). By defining *waa* only as a marker of illocutionary force, Saeed misses the fact that *waa* is parallel in discourse with *baa*, and that *waa* therefore plays a role in the information structure of Somali.
In addition, Saeed questions the analysis of \textit{waa} as a syntactic device reflecting verb focus, mainly because \textit{waa} occurs in verbless equative sentences such as (56). Saeed does not question the assumption that \textit{waa} is lexically empty, nor that it is pragmatically compatible with verbs being introduced as new information (1984:179). In one sense, Saeed is correct that \textit{waa} does not indicate verb focus. Narrow focus on the verb in Somali is impossible, as indicated by the following example:

(55)  
\begin{align*}
\text{a.} & \quad \textit{Anig-a ma ordaa?} \\
& \quad \text{Me-NOM INT run.PST} \\
& \quad \text{‘Did I run?’}
\end{align*}

\begin{align*}
\text{b.} & \quad \textit{Maya, adig-u (*waa) soco.} \\
& \quad \text{No, you-NOM walk.PST} \\
& \quad \text{‘No, you walked.’}
\end{align*}

The use of \textit{waa} in 55b was completely rejected by a native speaker. Saeed, however, only considers the possibility of verb focus, instead of nuclear or predicate focus, and misses a possibility which RRG allows – that even when no verb exists, \textit{waa} may focus the nominal or adjectival predicate of the utterance. For example:

(56) \quad \textit{Cali waa macallin.} \\
\quad \text{Ali FOC teacher} \\
\quad \text{‘Ali is a \textbf{TEACHER.’}}

In 56, \textit{macallin} is part of the nucleus, not an argument of the verb.
The copula can only be missing if the second argument is nominal, but not if it is adjectival.

(57) a. *Cali waa weyn.  
   Ali FOC big be.3SG.M  
   ‘Ali is BIG.’

b. *Cali waa weyn.

In fact, one of Saeed’s reasons for rejecting *waa as a marker of focus actually provides stronger evidence for its classification as such. *Waa never occurs in sentences with indefinite subject NPs, as the following examples show.

(58) a. Baabuur \(\left( \begin{array}{c} baa \\  \end{array} \right) i= dhaafay \\
   \) truck.INDEF FOC 1SG(ACC) passed.REST  
   ‘A TRUCK passed me.’ (Saeed 1984:170)

Because indefinite NPs are highly likely to be focused (Van Valin 1997:205), they would also be most unlikely to occur as the subject of a sentence in which the predicate is in focus.
Saeed’s analysis of *waa* as simply a ‘declarative clause marker’ fails to account for the fact that *waa* cannot co-occur with *baa*, whereas other illocutionary force markers (and WH words) can, as shown in examples 59 and 60:

(59) *Ma baabuur baa ku= dhaaf-ay?*  
 INT truck.INDEF FOC 2SG(ACC) passed-3SG.M  
 ‘Did a TRUCK pass you?’ (Saeed 1984:171)

(60) *Xagee buu ku siiy-ay Maryan buug-ii?*  
 Where FOC 3SG.M(NOM) to give-3SG.M Maryan book-DEF?  
 WHERE did Ali give the book to Maryan?’

He also claims that *waa*, if it is to be analyzed as a focus marker, should operate in parallel ways in discourse (1984:177) to *baa*. According to him, the following question-answer pairs demonstrate that *waa* does not function as a focus marker in the same way as *baa*:

(61) *Cali lacagtii buu keenay, sow ma aha?*  
 Ali money.DEF FOC 3SG.M(NOM) brought, IF NEG be.3SG  
 ‘Ali brought the MONEY, didn’t he?’

62 and 63 are both appropriate replies to 61, while 64 is inappropriate:

(62) *Haa, Cali lacag-tii buu keen-ay.*  
 Yes, Ali money-DEF FOC 3SG.M(NOM) brought-3SG.M  
 ‘Yes, Ali brought the MONEY.’

(63) *Haa, Cali lacagtii wuu keen-ay.*  
 Yes, Ali money.DEF FOC 3SG.M(NOM) brought-3SG.M  
 ‘Yes, Ali BROUGHT the money.’

(64) *#Haa, Cali baa lacagtii keen-ay.*  
 Yes, Ali FOC money.DEF brought-3SG.M  
 ‘Yes, ALI brought the money.’ (Saeed 1984:177-78)
Saeed (1984:28) claims that 63 should not be an appropriate response to 61, because 61 should only elicit a confirmation response, which he predicts to be narrow focus on the object. However, it seems that there are other discourse factors involved which allow 63 to be an appropriate response to 61, such as the fact that \textit{waa} constructions seem to be preferred as confirmation responses to Y/N questions, based on my data (shown in example 44b). The response in 63 may be a different type of focus, e.g. emphatic vs. the introduction of new information. This distinction is mentioned by Saeed (1984:27), but not used by him here to predict the acceptability of this response.

3.1.2 Livnat (1983,1984)

Additional research on Somali focus structure was completed by Livnat (1983, 1984). She analyzes Somali focus structure through the framework of Extended Standard Theory (EST), one of the early generative theories which acknowledged the influence of focus structure in grammar. Unlike Saeed, she concludes that every main declarative clause in Somali MUST have a focus marker (Livnat 1983:89).

One of Livnat’s goals is to explain the distribution of the clitic pronoun in conjunction with \textit{baa}. According to Livnat (1983:89,112), only a cleft construction analysis can explain the lack of a nominative clitic pronoun when the subject is in focus. In her analysis, she proposes that a \textit{baa} focused constituent is extracted from the clause, and that the clause is left syntactically subjectless. Because the coreferential clitic pronoun is actually an agreement marker, it has no subject to agree with and therefore does not occur. Yet when the subject is not focused and precedes the focused NP, the clitic pronoun which is coreferential with the subject/topic still occurs (as in 65). Even though the subject is outside the clause in this example, the pronoun may still occur.
Livnat’s explanation for the non-initial focused constituents is to propose a rule which scrambles the constituent order AFTER the focused constituent has been preposed and the coreferential subject clitic has been barred from occurrence. (1983, 126).

I also found that Livnat (1983:90) and Saeed’s descriptions of Somali as a free word order language to be somewhat overstated. They present examples of two verb-initial word orders, shown in examples 66-67, which were judged to be ungrammatical by my consultants (Livnat 1983:90). 

\textit{Waa} is a second-position clitic, and cannot appear in sentence or core-initial position.

\begin{equation}
\begin{aligned}
(66) \quad *Wuu & \quad \text{ark-ay} \quad \text{nink-ii} \quad \text{naagt-ii}. \\
& \text{waa}=uu \\
& \text{FOC 3SG.M(NOM)} \quad \text{saw-3SG.M man-DEF woman-DEF} \\
& \text{‘The man SAW the woman.’}
\end{aligned}
\end{equation}

\begin{equation}
\begin{aligned}
(67) \quad *Wuu & \quad \text{ark-ay} \quad \text{naagt-ii} \quad \text{nink-ii} \\
& \text{waa}=uu \\
& \text{FOC-3SG.M(NOM)} \quad \text{saw-3SG.M woman-DEF man-DEF} \\
& \text{‘The man SAW the woman.’}
\end{aligned}
\end{equation}

The following examples lead Livnat to conclude that a focus marker analysis of 
\textit{waa} is problematic.

\begin{equation}
\begin{aligned}
(68) \quad \text{Maxay} & \quad \text{samay-say} \quad \text{Caashi?} \\
& \text{ma+waxa+baa}=ay \\
& \text{INT what FOC 3SG.F(NOM) do-3SG.F Asha} \\
& \text{‘What did Asha do?’}
\end{aligned}
\end{equation}

She assumes that 68 should elicit verbal focus, though in fact it elicits focus on an argument of the verb (the location).
(69) a. *Gurig-ii bay tag-tay baa=ay*  
House.DEF FOC 3SG.F(NOM) went-3SG.F  
‘She went HOME.’

b. *#Way tag-tay gurig-ii waa=ay*  
FOC 3SG.F(NOM) went-3SG.F house-DEF  
‘She WENT home.’  
(Livnat 1984:100)

In 69, there is competition between the entire predicate and the location for focus, because neither is presupposed in the question. As stated before, there seems to be a preference for narrow focus over broad focus. However, if the response to 68 is an intransitive sentence, the *waa* construction does occur. That is, when the predicate is the only constituent eligible for focus (as in 70b), the *waa* construction is used to indicate predicate focus:

(70) a. *Muxuu Cali samey-ay? ma+waxa+baa=uu*  
INT what FOC 3SG.M(NOM) Ali do-3SG.M  
‘What did Ali do?’

b. *Cali wuu qosl-ay. waa=uu*  
Ali FOC 3SG.M(NOM) laughed-3SG.M  
‘Ali LAUGHED.’  
(Livnat 1983:94)

In conclusion, Livnat’s conclusion is incorrect; *waa* is indeed an indicator of predicate focus; 69b is inappropriate because narrow focus is preferred over broad focus, not because *waa* does not indicate predicate focus.

3.1.3 Aspects of Discourse Configurationality

More recent research has expanded the study of Somali focus structure into the realm of pragmatics and discourse. Svolacchia, Mereu and Puglielli’s (1995) article presents a more in-depth analysis of the syntax and semantics of focus structure in
Somali. There are several areas in which RRG provides a simpler and more consistent analysis than theirs.

Svolacchia et al. (1995:72) state that the focus marker *baa* is dependent on the verb, based on the following example:

(71) a. *Ma Maxamed baa teg-ay mise Cabdi baa teg-ay?*  
    INT Maxamed FOC gone-3SG.M or Cabdi FOC gone-3SG.M  
    'Has MAXAMED gone or has ALI gone?'

b. *Ma Maxamed baa teg-ay mise Cabdi?*  
   INT Maxamed FOC gone-3SG.M or Cabdi  
   'Has MAXAMED gone or has ALI (gone)?'

c. **Ma Maxamed baa teg-ay mise Cabdi baa?**  
   INT Maxamed FOC gone-3SG.M or Cabdi FOC  
   'Has MAXAMED gone or has ALI gone?'  
   (71a-c from Saeed 1984:110)

According to the authors, the ungrammaticality of 71c is due to the absence of the verb *tegay* in the second clause; therefore, the particle *baa* must be part of the Verbal Complex, because it cannot occur without the verb (Svolacchia et al 1995:72). However, there is an alternative analysis, that the ungrammaticality of 71c is due to the constraint that a focused constituent in a *baa* construction must occur in the core-initial position, and therefore is prohibited from occurring postverbally. This constraint makes it unnecessary for *baa* to be analyzed as part of the verbal complex when there is no compelling evidence to do so. In fact, there is evidence to suggest that *baa* is not part of the Verbal Complex, because elements may intervene between *baa* and the verb (see example 5b).

Svolacchia et al. note that a coreferential clitic pronoun always occurs in agreement with the topic, whether subject or object, and that a focused object NP co-occurs with the clitic pronoun, while a focused subject NP does not. They fail, however,
to note that the object pronoun always occurs, that is, that agreement with the object is always present. When the topic and the object coincide, a single clitic pronoun (in nominative form) serves both functions.

The authors’ final claim (1995:93) is that there are no specialized positions in Somali grammar for discourse roles. However, I have shown that in each construction, there is a specialized position for focus: the core-initial position. In each construction, the actual and potential focus domains are identical, i.e. the focused constituent remains in the same position in each construction.

3.2 Conclusion

Although focus structure in Somali is well-documented in the literature (Saeed 1984, Livnat 1984, Svolacchia 1995 et al), it has not been previously analyzed from a theoretical framework which allows for the integration of syntax, semantics and pragmatics. Previous analyses only examined the effects of focus markers on the constituents in focus. However, because of the integral role played by focus structure in the syntax of Somali, focus markers often influence more than the constituents they focus; they also affect the occurrence of coreferential pronoun clitics, verb paradigms, and case endings. Although much has been said about the relatively free word order in Somali, no previous studies have related this free word order to the relative rigidity of focus structure in the language. The unmarked focus position in Somali is the core-initial position. In most instances, the preferred response is the one which replaces the WH-word in its exact position; i.e. occurs in the unmarked focus position. The analysis outlined in this thesis is more concise than the previous analyses, because it does not propose that transformations are necessary in order to explain the focus structure of
Somali. For example, Svolacchia et al (1995:70) notes that constituents focused by *baa*
are always preverbal, but the authors do not recognize the fact that, despite the variation
in word order, the focused constituent remains in the same position. Within their
theoretical framework, there is no way to posit that sentences with different word orders
can contain constituents which are located in the same position, without proposing
transformations. There are constraints on the order of the constituents in Somali: the
focused constituent in the *baa* construction must appear preverbally, while in the *waxaa*
construction, the constituent which receives secondary focus may not appear preverbally.
Therefore, the unmarked focus positions for primary and secondary focus are mutually
exclusive, as is expected. In addition, there is a relationship between focus structure and
intrasentential pronominalization. Previous analyses found it difficult to explain the lack
of a coreferential pronoun with lexical NPs. The principle of intrasentential
pronominalization in RRG, however, explains this seeming anomaly.

There are many issues which arose during this study of Somali focus structure
which require further investigation. First, the interaction of focus structure and word
order deserves further exploration. Determining preferential answers to questions without
an established discourse context proved difficult for native speakers. A more detailed
experiment, involving analysis of natural texts, may shed light on the functions of
different word orders and their relationship to topic and focus. Studies of spoken Somali
and written text would also provide assistance in further detailing the function of focus in
the language. In addition, an exploration of other morphosyntactic features such as case
marking would show further effects of the object on grammatical features of the
language.
References


