

4.1 Introduction

Recall from chapter 1 we saw that children learning a variety of languages allow root bare verbs. We saw that children acquiring Sesotho allow bare verbs at early ages before producing verbs with filler ‘shadow’ syllables, and then going on to adult-like utterances (Demuth, 1992; 1994). Children acquiring Siswati follow the same developmental order (Kunene, 1979). Children acquiring Quechua, an agglutinative language spoken in Peru, use bare verbs very frequently at early stages - 57% of all verbal utterances (Courtney, 1998). We also saw that children acquiring Inuktitut produce bare verbs, which are ungrammatical in adult Inuktitut (Swift & Allen, 2002). In all of these adult languages bare verbs are unattested, and so this constitutes a genuine departure from the adult norm.

We also saw that errors of omission are very frequent in child language (bare verbs being an example of such an error), but errors of commission are extremely rare. The one exception to this generalization is the case of Root Infinitives (RIs). We saw that children learning languages such as German, Dutch, French, etc. allow root clauses with infinitival verbs. However, children learning other languages such as Italian, Spanish, etc. produce neither RIs nor bare verbs. Thus we developed a typology of languages, as summarized in table 1.12 in chapter 1, reproduced here:

Table 1.12 Summary of languages that allow RIs, disallow RIs and allow bare verbs.

True RI languages	Non-RI languages	Bare Verb Languages
German	Italian	English
Dutch	Spanish	Sesotho
French	Catalan	Siswati
Swedish	Japanese	Quechua
Icelandic		Inuktitut
Russian		

How does Swahili fit into the typology described in table 1.12?

We know that Swahili has an infinitive prefix (*ku*), and so RIs in principle should be available to children. However, Swahili may behave like Italian (which also has an infinitive, but which is not an RI language). Finally, Swahili may behave like its closely related language Sesotho in exhibiting bare verbs. The inflectional morphology of Swahili is particularly interesting because it independently and distinctly marks Agr and T. As we will see, this proves crucial in distinguishing between several influential theories in language acquisition. We will investigate the question of whether Swahili is an RI language, a bare verb language or neither, as well as the development of inflectional affixes in the Swahili verbal complex.

This chapter is organized as follows. In section 4.2 I discuss five theories of language acquisition. The first approaches the problem from a phonological perspective and attempts to account for the general tendency towards omission. The remaining four approaches are syntactic in nature. In section 4.3 I consider the predictions that each theory makes for Swahili,

and in 4.4 I introduce the results. In section 4.5 I return to these theories and re-evaluate them in light of the Swahili data. Section 4.6 includes a summary and some concluding remarks.

4.2 Theories of language acquisition

In this section we will discuss five theories of the acquisition of functional elements, each approaching the problem from varying perspectives and drawing on different data. Given that Swahili overtly represents functional material in such a clear and unambiguous manner, these theories are particularly relevant to the acquisition of Swahili. The theories fall into three basic classes. The first considers the omission of functional elements such as determiners to be a result of production constraints. Thus a full syntactic representation is postulated, with production limitations constraining the expression of the entire representation. The second theory considers child Root Infinitives and other functional omissions to be the result of truncation of higher portions of the syntactic representation. The final class of theories, often called Full Clause Theories, argues that early grammars project all the way to the CP in all cases, but that specific deficits in particular heads result in the omission of particular features. After discussing each proposal, I state in section 4.3 the predictions that each theory makes for child Swahili. It is here I believe Swahili may be most enlightening since, as we will see, each theory makes precise predictions about the acquisition of Swahili. In section 4.4. I present the data and return to these theories of acquisition in section 4.5.

4.2.1 Metrical Omission Model (Gerken 1991; Gerken & McIntosh, 1993)

It has long been noted that children learning English typically omit determiners in obligatory contexts (Brown, 1973; Radford, 1990; Hoekstra & Hyams, 1995; Hoekstra, Hyams & Becker, 1996).

- | | | | |
|-----|----|------------------|-----------------|
| (1) | a. | Eve sit floor | (Brown, 1973) |
| | b. | Hayley draw boat | (Radford, 1996) |

Gerken (1991) and Gerken & McIntosh (1993) propose that children have less processing capacity than adults, which leads to constraints on the production of phonological segments. These production constraints prevent children from expressing a full linguistic representation by forcing them to produce only syllables that conform to their language-appropriate metrical pattern. Specifically, children omit weakly stressed syllables that either fall outside the metrical pattern that is typical of their language or that occur as part of a non-typical metrical foot. For example, English has a predominance of trochaic feet, and so weak syllables that occur in pre-trochaic position or weak syllables that occur as part of an iambic foot are subject to omission (see 3 below). Gerken (1991) discusses the nature of these production limitations. She suggests that previous production accounts (e.g., L.Bloom, 1970; P.Bloom, 1989; Valian, 1989) suffer from the limitation that they must posit a system of meta-awareness in order for the child to know when to omit material. Most production-limitation accounts argue that sentence complexity adds to processing load, and thus the child omits material in complex sentences. Gerken quite rightly points out that this means that the child must have a system of meta-awareness that will tell the child when something is overly complex. This

system must also be able to pragmatically identify material that can be omitted (i.e., redundant material, old information, etc.). These additional systems presumably increase processing complexity, and thus are not desirable in a processing account of omission.

Instead, Gerken proposes that children's omissions are the result of mechanisms within the production system itself. She proposes a system of production that makes use of rules and templates. The original linguistic message goes through several ordered levels of rules and templates (semantic > syntactic > morphological > foot formation > head location > phonological rules and templates) before the utterance is produced. She proposes several principles that account for children's omissions. First, the child has limited processing resources. Second, templates draw less resources than rules do. And third, using resources earlier in the process (i.e., closer to the message as opposed to the produced utterance) is less taxing. Thus making use of templates at the semantic level is less taxing than making use of templates at the syntactic level.

There are two possible metrical patterns in languages, and individual languages differ as to their preferred metrical pattern. A trochaic foot is one in which the syllabic stress pattern is Strong-Weak (2a), while an Iambic foot has the pattern Weak-Strong (2b)¹. The trochaic foot is the basic meter used in the production of English words, while French makes use of Iambic feet as the basic metrical pattern.

¹ Gerken argues (following Hayes, 1982; Kelly, 1988) that binary feet are preferred over both ternary feet and unbounded feet (both of which are possibilities in adult languages). She thus only considers the simplest possibilities of trochaic and iambic feet. I follow her on this. Furthermore, the strength of a syllable is determined by the pitch, the stress, the length, etc. (Gerken, 1991; Gerken & McIntosh, 1993).

- | | | | |
|-----|----|---------|---------------|
| (2) | a. | monkey | Trochaic foot |
| | | S - W | |
| | b. | giraffe | Iambic foot |
| | | W - S | |

Gerken proposes that there is a trochaic template operating at the level of foot formation or head location. She claims that this may be because trochaic feet are easier to plan and produce (in terms of motor skills) than iambic feet. In fact, Allen & Hawkins, (1980) were the first to suggest this idea, and several others have argued in favor of such a position (e.g., Fletcher, 1985; Echols & Newport, 1992; Fikkert, 1992; Archibald, 1995; La Belle, 1999). Gerken points out, however, that if this were the case, it would put languages in which iambic feet predominate (e.g., French) at a serious disadvantage, and so she suggests that the trochaic template may develop from the English-speaking child noticing that trochaic feet are more common in her language.

Because of the constraints on production, children must target material for optional omission. The Metrical Omission Model (MOM) states that weakly stressed syllables are targeted for omission, never strong syllables. Furthermore, children prefer binary feet, and cannot leave any strong syllables unparsed². Because English is a trochaic language, English speaking children apply a trochaic template to their production system and maximize the number of trochaic feet in any particular string. Furthermore, in a trochaic language such as English, children tend to drop weak syllables that are part of an iambic foot (the dispreferred pattern, as in (3a)), and weak syllables that precede a fully formed trochaic foot (as in 3b).

² Weak syllables, however, may be left unparsed.

Crucially, however, weak syllables in a trochaic foot are not subject to omission (3c).

- (3) a. the ball
[W-S] Weak Syllable in Iambic foot
↑ Subject to Omission
- b. the mon key
W [S - W] Weak pretonic syllable
↑ Subject to Omission
- c. mon key
[S - W] Weak syllable in Trochaic foot
↑ Not Subject to Omission

Notice that in both (3a) and (3b) the syllables that are subject to omission are determiners. This has the desirable prediction that determiners will be particularly prone to omission because of the metrical pattern of English and the placement of determiners in English (cf. examples (1) in section 1.2.1). As the child's processing capacity improves, the omission of determiners decreases until they reach adult-like proportions.

Demuth (1994) uses the MOM to show that Sesotho children omit noun class prefixes based largely on these principles. Sesotho noun class prefixes (like Swahili noun class prefixes) are monosyllabic (usually CV) prefixes that attach to either a mono- or di-syllabic noun stem. Omission of prefixes occurs on nouns as well as agreement prefixes on other elements such as demonstratives:

- (4) [kolo] [sá-ne] Sesotho
adult form: se-kólo sá-ne (from Demuth, 1994, p.129)
'that school'

According to Demuth, in (4) the agreement prefix *se-* on the demonstrative *kólo* is omitted by the child because it occurs as a pretrochaic weak syllable. Demuth extends this analysis to account for why Sesotho children omit preverbal inflectional affixes. She shows that Sesotho has penultimate lengthening which produces trochaic feet at the end of all prosodic words. This creates a trochaic bias, which when coupled with the principles of the MOM predict omission of preverbal affixes. Following are her examples (18) from a child aged 2;1:

- (5) a. ta hâ:na
adult form: [ke - a - hán - a +]
SA_{1s} - pres - refuse - IND
'I refuse.'
- b. áy shépa
adult form: [ó - a - sháp - a +]
SA_{1s} - pres - lash - IND
'S/he is lashing.'

In (5a), the preverbal affixes have been reduced from *ke-a-* to *ta-*, while in (5b) the preverbal affixes are reduced from *ó-a* to *áy*. Demuth claims that this reduction is because of their position as weak pretonic syllables. She further suggests that children operate under a Minimal Word Constraint which she states as in (6):

- (6) A prosodic word contains a foot

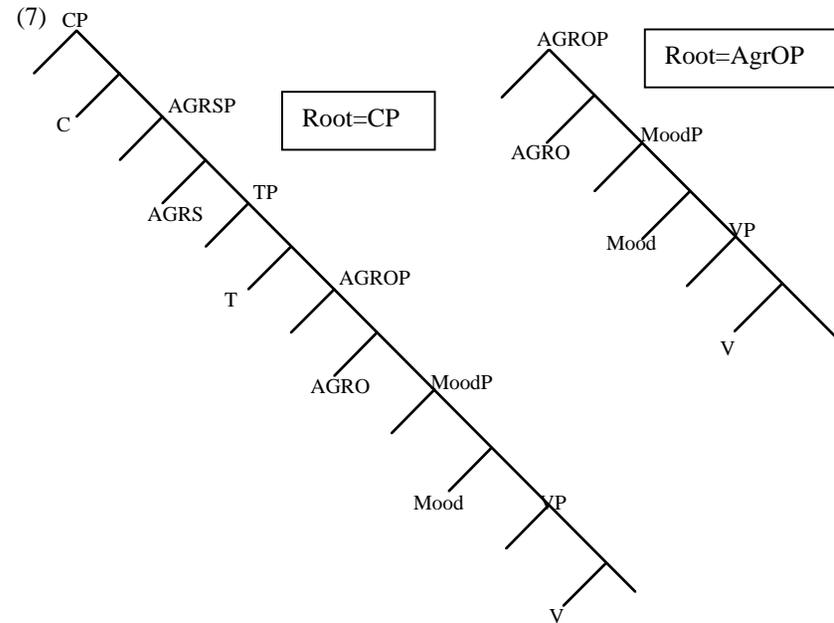
Thus the examples in (5) contain a foot, and satisfy the Minimal Word Constraint. However, Demuth claims that 'syllables falling outside of the foot will be treated as extrametrical and subject to reduction or deletion' (p.128). Thus the prefixes in (5) are reduced because they fall outside the basic trochaic foot.

These results being from Sesotho, a related Bantu language, may be indicative of what we may expect in Swahili. We will discuss the predictions MOM has for Swahili in section 4.3.1.

4.2.2 Truncation (Rizzi 1994)

Rizzi (1994) proposes that early grammars optionally fail to project the entire CP, producing truncated structures. Rizzi's primary concern involves RIs (see section 1.2.3). Recall that RIs are optional, they occur in unraised positions, and typically do not occur in wh- questions. Based on these facts, Rizzi proposes a structural account of RIs. Rizzi proposes that young children differ from adults in that the adult axiom of 'Root=CP' is not operative in child grammar. Children can optionally specify the root as any projection, and can thus have a truncated structure. While the adult must project a structure as high as the root CP for every declarative clause, the child can optionally project to a lower position, e.g. AgrOP (see 7). In the case of RIs, the child projects to a position lower than TP, and hence produces a tenseless clause.

The optionality of RIs is accounted for by the absence of the axiom root=CP. Since children can optionally specify the root as any projection in the structure, RIs can optionally occur. Furthermore, according to Rizzi, RIs occur in structures that are truncated below TP. Thus RIs occur in structures that are missing all projection above TP, including the CP. Recall that RIs rarely occur in wh- questions: a fact that is elegantly accounted for under this proposal since the presence of a wh-question necessarily entails a CP projection, hence everything below the CP (including Tense) must also be projected.



A crucial property of truncation is that when a particular projection is specified as the root by the child, all the structure up to that point is projected. For example, if the child projects up to the CP (such as when a wh- question occurs), it is not possible to leave out any intervening projections, such as TP. Hence the absence of RIs in wh- contexts. This means that any omission in the grammar of the child occurs at the higher periphery, and cannot target specific projections internal to the sentence.

The remaining three theories assume that children project each sentence all the way to CP, and so are often referred to as Full Clause Hypotheses. However, in order to account for the particular unadult-like characteristics of child speech, specific heads are assumed to be

underspecified³. Thus the difference between adults and children is localized to a particular head or heads. I will now discuss the three versions of underspecification that I believe have the most relevance to Swahili.

4.2.3 Underspecification of Tense (Wexler 1994)

Wexler (1994) points out that infinitives in adult languages like German and Dutch occur in embedded clauses, they license PRO, they can be used gerundively, etc.

- (8) a. John tried [PRO to eat all the plums]
 b. PRO to run everyday is good for your health

These properties are thought to derive from the absence of tense in infinitives. Wexler argues that children have the option to underspecify tense in main clauses, giving rise to RIs. This leads to the natural correlation of null subjects and RIs discussed in chapter 1, as PRO is licensed in non-finite contexts.

There are three pieces of evidence that suggest an underspecification of Tense in child RIs:

- (i) The form that occurs is an infinitive, which in the adult language is tenseless;
- (ii) RIs occur in unraised positions, which correspond to tenseless verbs in the adult language;
- (iii) RIs correlate with null subjects, which by hypothesis are PRO licensed in tenseless environments.

³ Underspecification means that a feature matrix on a particular head is empty. Often the feature matrix is an entire feature bundle (e.g., $\bar{\bar{}}$ -features), but sometimes it may be a single feature (e.g., a case feature). Alternatively, this could mean that the feature has no value.

Furthermore, Wexler argues that English bare verbs as in the examples in (9) and RIs in German, French, etc. result from the same underlying mechanism.

- (9)
- | | | | |
|----|--------------------------|----------------|---------------|
| a. | He fall down | Nina (2;1.29), | Suppes, 1971 |
| b. | Hayley draw boat | Hayley (1;8), | Radford, 1996 |
| c. | He bite me | Sarah (2;9) | Brown, 1973 |
| d. | It only write on the pad | Eve (2;0) | Brown, 1973 |

Wexler groups English bare verbs, bare participles / null auxiliaries and null copulas with RIs and refers to them as Optional Infinitives, arguing that they are all the result of the underspecification of Tense. Unlike Rizzi's Truncation Hypothesis, Wexler proposes that the difference between child grammar and adult grammar is specific and minimal rather than global in nature. The nature of the difference lies in the particular nature of an individual head - Tense.

4.2.4 Underspecification of Agr (Clahsen *et al.* 1996)

Clahsen *et al.* are concerned with the correlation between the acquisition of lexical knowledge (in terms of inflectional morphology) and certain syntactic effects. Putting aside the question of lexical learning versus full clauses⁴, let us focus on their underspecification proposal. Clahsen *et al.* are concerned with accounting for four pieces of evidence that are established in Clahsen (1990) for child German, listed in (10):

⁴ They are primarily concerned with the Lexical Learning Hypothesis which states that the acquisition of inflectional morphology drives the development of syntax in general. Thus they investigate the acquisition of subject-verb morphology and try to draw parallels in real time with the development of other syntactically related phenomena. This is not crucial to what follows, and so I shall put it aside.

(10)

- a. Subject-verb agreement, accusative case and dative case have not been acquired.
- b. Finite verbs raise to first or second position, while non-finite verbs remain in clause-final position.
- c. Subjects never intervene between the finite verb and negation (hence, in contrast to adult German, subjects never raise).
- d. wh-elements and complementizers are completely unattested.

They propose that children have a single functional projection above VP which is specified for +finite. This position cannot be identified with IP or AgrSP because the specifier of this position is not restricted to subjects. This position cannot be identified with CP because at this stage German children never produce wh- elements or complementizers (Clahsen, 1990). Therefore Clahsen *et al.* refer to this projection as FP – a general functional projection that is specified as <+finite>. They describe the feature <+finite> as a syntactic feature that allows its head to function as the landing site for a finite verb. The only reason they posit such a projection is because finite verbs quite robustly raise to first or second position. So FP accounts for the position facts in early German (finite verbs raise while RIs remain in final position)⁵. The absence of any other features and structure also accounts for the absence of wh- questions and complementizers. They go on to show that as the subject-verb agreement paradigm is acquired, syntactic processes associated with AgrSP develop, i.e., subject raising to [spec, AgrSP].

⁵ They show that a small proportion of finite verbs remain in final position (approximately 12% of all finite verbs), but almost no nonfinite verbs ever raise. They account for this by adding the specification of +V to the lower verbal position, making it compatible with both finite and nonfinite verbs. Therefore only finite verbs are allowed to raise but all verbs may remain low. This is not relevant to our discussion, so I will not discuss it further.

I interpret Clahsen *et al.*'s FP position as most closely resembling TP. According to Clahsen *et al.*, the FP projection carries <+finite> features⁶, which can include temporal features, as well as agreement, aspectual and mood features (Rizzi, 1997). However, since subject-verb agreement is seen as being integral to AgrSP (which has not developed, according to Clahsen *et al.*), finiteness here cannot include agreement features. Aspect and mood are generally seen not to affect verb raising in German, and so it is unlikely that FP includes aspectual or mood specification. We can therefore interpret finiteness in this case as being tense specification. Thus, another way of interpreting Clahsen *et al.*'s claim is that AgrSP (as well as AspectP and MoodP) is underspecified in early

⁶ They describe this feature as being hosted by FP, and allowing the hosting head to act as the landing site for finite verbs. However, 'finiteness' is a general descriptive term that is usually composed of several distinct features. For example, Rizzi (1997) describes finiteness as including tense, aspect, mood and sometimes agreement features. I assume this definition of finiteness. What is important for Clahsen *et al.* is that there be one and only one position to which verbs and only verbs may raise. Therefore this position cannot be associated with subjects (hence AgrSP is ruled out), and it cannot be associated with wh- elements or complementizers (hence CP is ruled out). This is why I say that their FP can be thought of as TP.

Table 4.1 Summary of possible INFL features for ATOM

	Verb form	Subject	Examples
Tense = present, +agreement	-s	NOM	He cries
Tense = present, -agreement	OI	ACC	Him cry
Tense = past, +agreement	-ed	NOM	He cried
Tense = past, -agreement	-ed	ACC	Him cried
-Tense, +agreement	OI	NOM	He cry
-Tense, -agreement	OI	GEN?	His cry

(Table taken from Schütze, 1997, p.232)

Schütze & Wexler’s argument holds appeal for two reasons. First, it draws attention to an additional syntactic effect of underspecification: Case. Thus case effects are now added to the list of syntactic effects that correlate with OIs: null subjects, no verb raising, and absence in wh-questions. This adds to the evidence that RIs are not speech errors or the result of production problems, but a true syntactic phenomenon. Second, they provide a theory of underspecification that is as precise as the other theories, but that allows for a greater amount of underspecification. Thus ATOM has an increased descriptive power because it allows more than a single underspecified head, but less predictive power because it introduces an additional degree of freedom into the model.

Swahili, as we saw in the previous chapter, marks tense and subject agreement as independent bound prefixes, making Swahili an ideal language to test these theories on. This is especially true for the underspecification theories since they each make claims about an inflectional head that is marked in Swahili as a unique and distinct prefix: SA and T. In the next section I will discuss the specific predictions that each theory makes for Swahili.

4.3 Predictions for Swahili

None of the theories reviewed in the previous section were intended to account for Swahili. Thus we do not intend this to be a criticism of the theories. However, as generalizability is a desirable property of any theory, Swahili allows us to test how generalizable each theory is. Furthermore, each of the theories discussed in the previous section make precise predictions about the kinds of things Swahili children should allow.

4.3.1 Metrical Omission Model

The Metrical Omission Model is important for us to consider for two reasons. First, Swahili has a trochaic bias, as we saw in section 2.2. Furthermore because the inflectional morphology is primarily prefixed, these prefixes are potentially in the contexts that Gerken identifies as prone to omission. We will discuss this in more detail below. The second reason that this model is important for us to consider is because Demuth (1994) uses it to account for Sesotho noun class omission and Sesotho verbal prefix omission/reduction. She argues that because Sesotho is a language that has penultimate lengthening and a bias for trochaic feet, verbal and noun class prefixes are usually either in pre-trochaic position or weak syllables in iambic feet. Thus verbal and noun class prefixes should be subject to omission/reduction by the Metrical Omission Model. Because Swahili is also a trochaic language, we expect the same pattern of omission as we see in English and Sesotho:

- Omission of pretonic weak syllables, e.g., **W** [S – W]
- Omission of weak syllables in iambic feet, e.g., [**W** – S]
- No omission of weak syllables in trochaic feet, e.g., [S – W]

We can use this as a guide to assign metrical structure to a typical Swahili verbal complex. The minimal verbal complex in Swahili consists

(18)	S	W	W	S	W	→ [S – W] W [S – W]	Applicative suffix		
syllabic	a	–	li	–	pi	–	ki	–	a
morphemic	a	–	li	–	pik	–	i	–	a
	SA _{3s} –psat–cook–applic–IND								
	‘He cooked for (someone).’								
(19)	S	W	W	S	W	→ [S – W] W [S – W]	OA prefix		
syllabic	a	–	li	–	ni	–	pi	–	ga
morphemic	a	–	li	–	ni	–	pi	–	ga
	SA _{3s} –past–OA _{1s} –hit–IND								
	‘He hit me.’								

Consider first the cases in which the verb stem is more than 2 syllables long. Example (16) shows a stem composed of three syllables, and (17) shows a stem composed of four syllables. Using the rules of the MOM (14), the only possible metrical analysis is given to the right of each example. Notice that in both cases weak syllables are left unparsed between two trochaic feet. According to the MOM these medial weak syllables are subject to omission. Therefore unlike verbs that have disyllabic stems, longer verb stems do give rise to potential omission through the MOM. However, the syllables that are subject to omission in both cases are onset syllables of the verb stem. Therefore the inflectional prefixes are not subject to omission under the MOM.

Similarly in the case of a syllabic suffix, as in (18), a medial weak syllable between two trochaic feet occurs, which is subject to omission. But in this case, as in the previous two cases, the weak syllable that is subject to omission is the onset syllable of the verb stem, not the inflectional prefixes. And finally the case of OA prefixes (19): the weak medial suffix that is subject to omission is the OA prefix. Therefore according to the MOM, OA should always be subject to omission because it is always a pretonic weak syllable. However, as in all the previous cases, the inflectional suffixes SA and T are never subject to omission.

Let us summarize the predictions of the MOM:

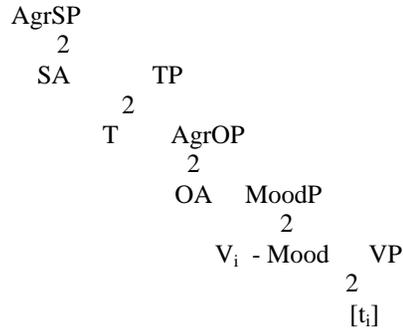
- (20) I. SA should never be omitted
(because it carries secondary stress).
- II. T should never be omitted
(because it is the weak syllable in a trochaic foot).
- III. OA should always be subject to omission
(because it is always a pre-trochaic weak syllable).

4.3.2 Truncation Hypothesis

Swahili provides us with a language that is particularly clear in its morphological structure. Not only does it spell out three inflectional morphological heads as separate, independent morphemes, it does so in the order that corresponds to their hypothesized hierarchical order (Baker’s, 1987 Mirror Theory; see Cinque, 1999 for discussion). Furthermore, there is very little phonological interference among inflectional morphemes, rendering surface forms fairly reliable indicators of underlying syntactic structure. Given this, Truncation provides a theory that has clear predictions with respect to early Swahili. Recall that truncation provides the child with the opportunity to specify the root as any projection in the syntactic tree. Let us assume the following structure of the Swahili verbal complex¹¹:

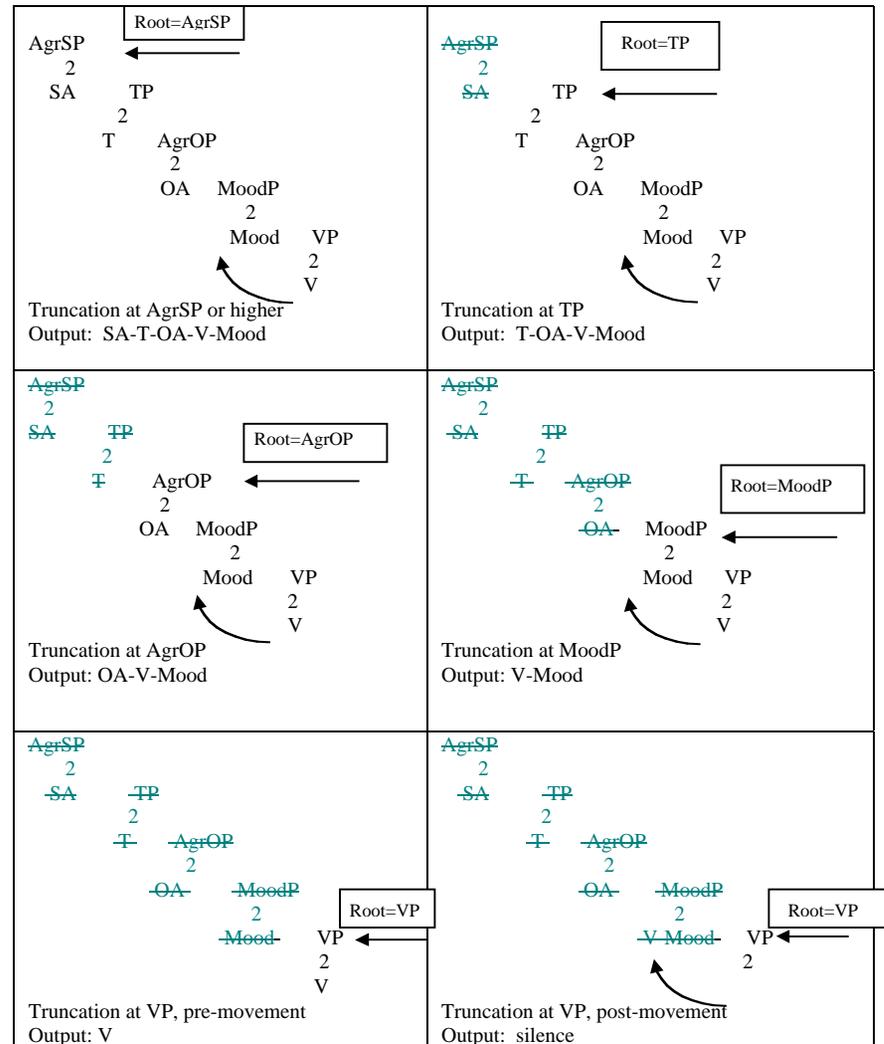
¹¹ In chapter 2 I gave a derivation of a Swahili simple tensed clause that involved head movement of the verb to Mood and remnant movement that re-ordered constituents into their surface order. I use a simpler structure here for the sake of exposition. Were we to use the more complex remnant structure the results would be the same. I also ignore specifier positions, but this again is for the sake of exposition.

(21)



In child grammar, according to Rizzi, the root may be specified as any projection above (and including) VP. Below are all the structures that the child could produce if truncation were to occur at each possible level. Truncation is silent on what level of representation truncation occurs at, i.e., it is not clear whether truncation occurs before movement or after movement. Therefore the status of Mood is unclear because prior to movement Mood is higher than the verb, but after the verb moves to Mood this is no longer the case. The only times this makes a difference is if truncation occurs at the VP level, in which case pre-movement truncation will yield a different result from post-movement truncation. I have given both possibilities below. I have indicated the “deleted” portion of the trees in lighter shade and strikethrough. Below each structure I give the output from that structure.

Recall that OA in Swahili is dependent on specificity of the object. This means that in all cases of a non-specific object, OA will not occur. Because of the difficulty in determining obligatory contexts for OA and for the sake of exposition, I will not consider it further. See section 4.3 for more details on OA in early Swahili.



Summarizing the results, we find that truncation predicts the following structures should be possible in early Swahili (ignoring the trivial option of silence).

- (22) Possible clauses according to truncation:
- | | | |
|------|------------|--------------|
| I. | SA-T-V-IND | Full Clause |
| II. | Ø-T-V-IND | [-SA] clause |
| III. | Ø-Ø-V-IND | Bare Stem |
| V. | Ø-Ø-V-Ø | Bare Root |

Furthermore, Truncation makes predictions regarding what sorts of clauses should NEVER occur in Swahili. Recall that while the root may be specified as any projection in the structure, all projections below the root *must* be projected – the child cannot omit any intervening material.

Therefore the following clause types are excluded by truncation:

- (23) Excluded clause types, according to truncation:
- | | | |
|------|------------|---------------------------|
| I. | SA-Ø-V-IND | [-T] clause with SA |
| II. | SA-T-V-Ø | [-Mood] clause |
| III. | SA-Ø-V-Ø | [-T] clause without Mood |
| IV. | Ø-T-V-Ø | [-SA] clause without Mood |

[-T] clauses are excluded as well as any clause that is missing mood but has a prefix. Thus mood may only be omitted when all other prefixes are omitted as well. Before moving on, we should note that supporting evidence for Truncation would include some or all of the possible clauses in (22) being attested, i.e., not all of the clause types in (22) need be attested. However, crucially, NONE of the excluded clauses in (23) can be attested. Any substantial (non-speech error) occurrence of the excluded clauses in (23) constitutes counter-evidence for Truncation.

4.3.3 Underspecification Theories

The underspecification theories make simple predictions. Wexler's underspecification of Tense theory predicts that children should allow the following clause types:

- (24) Permissible Clause types, according to Wexler (1994):
- | | |
|-----|------------|
| I. | SA-T-V-IND |
| II. | SA-Ø-V-IND |

Crucially, Wexler's theory does not allow SA to be omitted, and it does not allow Mood to be omitted. On the other hand, Clahsen *et al.*'s underspecification of Agr theory predicts a very different result:

- (25) Permissible Clause types, according to Clahsen *et al.* (1996):
- | | |
|-----|------------|
| I. | SA-T-V-IND |
| II. | Ø-T-V-IND |

Again, this underspecification theory predicts that SA may be omitted optionally, but T may not be omitted and Mood may not be omitted either. Finally, Schütze's (1997) ATOM predicts that the following options should be possible:

- (26) Permissible Clause types, according to Schütze's (1997) ATOM:
- | | |
|------|------------|
| I. | SA-T-V-IND |
| II. | Ø-T-V-IND |
| III. | SA-Ø-V-IND |
| IV. | Ø-Ø-V-IND |

ATOM predicts that SA and T may be optionally and independently omitted, but that Mood may not be omitted.

Table 4.2 Summary of Predictions

MOM:	SA and T should never be omitted, but OA and onset syllables of long verb stems are subject to omission.
Truncation:	Tense should never be omitted when SA occurs, and Mood should never be omitted when any prefixes occur. Other clause types are permitted.
Underspecification of T:	The only clauses permitted are full clauses and [-T] clauses.
Underspecification of Agr:	The only clauses permitted are full clauses and [-SA] clauses
ATOM:	Mood omission is excluded, otherwise all four combinations of prefix omission are permitted: Full clause, [-SA] clause, [-T] clause, Bare Stem

On more general grounds, based on what we know about children cross-linguistically, we might expect Swahili children to do two things:

- omit affixes
- Produce RIs

What would these look like in child Swahili? Let us consider the RI possibility first. Swahili has an infinitive marker that occurs in the position of T: *ku*. Recall from chapter 2 that the infinitive occurs in complementary distribution with all other tense markers, occurs in typical infinitival contexts (embedded clauses, gerunds, etc.), licenses PRO, etc. A child RI would include the infinitive marker, an indicative final mood vowel, but no SA (as SA and *ku* do not cooccur in adult Swahili). Thus we may schematize an RI as follows:

(27) INF-V-IND

Turning to possible omission of affixes, below are some possible clause types that we expect Swahili children to produce as well as labels that I assign the particular clause types. Because OA and the grammatical

function changing suffixes are optional, I ignore them here. I also do not consider mood omission because mood is less relevant to the theories that we discussed earlier.¹² We will return to mood omission in section 4.6. Furthermore, I consider only indicative clauses in this chapter because subjunctive clauses have a different structure: tense is obligatorily absent in subjunctive clauses. The relative omissions of SA and T are crucial to our evaluation of the theories presented earlier (especially the underspecification theories). Because subjunctive clauses do not contribute to this discussion I will not consider them in this chapter. See Deen & Hyams (2002) for details on the use of subjunctive and indicative clauses.

(28) Logically Possible Omission types in child Swahili:

SA – T – V – Mood	Full Clause
∅ – T – V – Mood	[-SA] Clause
SA – ∅ – V – Mood	[-T] Clause
∅ – ∅ – V – Mood	Bare Stem

Thus these four clause types plus RIs are all logical possibilities in early Swahili. Of these four clause types, full clauses are the only clause type that is permitted in prescriptive Standard Swahili, while Nairobi Swahili permits full clauses and [-SA] clauses. Recall from chapter 2 that [-T] clauses and bare stems are virtually unattested in adult speech in indicative contexts (recall that imperatives occur as bare verbs. For this reason I will exclude imperatives in my analyses; see below).

¹² Mood omission could in principle occur with each of the clause types given in the text. Therefore we can also consider [-Mood] full clauses, [-SA, -Mood] Clauses, [-T, -Mood] clauses, [-Mood] Bare stems, etc. We will see very soon that Mood is in fact almost never omitted. Thus, because these possibilities are unattested, for the sake of avoiding additional complexity I do not consider them.

In the next section I will present the results of analyses of omission of prefixes. In section 4.5 we will return to these theories and evaluate their predictions with respect to the Swahili data.

4.4 Results

In this section I will present the results in the following order. First we will see which clause types are attested in the corpus. I will discuss the details of categorizing utterances into these discrete classes, and how conflicting criteria were resolved. I will then discuss in detail the development of each clause type across stages. We will see that, quite expectedly, unadult-like clause types diminish across time, while the adult-like clause types increase in proportion. However, the relative proportions of each clause type and the differential rates of development are of great interest. In section 4.5 we will move to the evaluation of the five theories of language acquisition presented in section 4.2. Finally, we conclude with section 4.6.

4.4.1 Attested Clause Types in Child Swahili - Overview

CLAN programs were used to run automated counts of the clause types. The program COMBO was used because it allows for searches that combine text, e.g, it allows for a search of the combination of the codes SA and T, or SA and Ø, etc. A first level of analysis revealed the following facts:

- (29) General facts about early Swahili:
- SA omission is extremely frequent
 - T omission is also frequent
 - Mood Omission is almost completely absent
 - RIs are almost completely absent

In the entire corpus, the mood final vowel is omitted a total of 12 times out of a total of 2662 verbal utterances (an omission rate of 0.45%). Upon inspection, these 12 tokens are clearly speech errors because they usually occur in broken speech (i.e., when communication breaks down and the child is struggling for her next utterance), or the final vowel is assimilated to the following vowel making it difficult to determine whether it is present or not. The twelve examples are listed in appendix 4A. Thus I conclude that the omission of mood is unattested, and I will not consider it further.

Similarly, RIs are virtually unattested. Of all the 2662 utterances, there are a total of 14 RIs, which represents an RI rate of 0.52%. These 14 RIs are presented in appendix 4B. I assume that RIs are not part of the grammar of early Swahili, a fact that represents a marked difference between Swahili and languages such as German, French, Dutch, etc.

Having eliminated mood from consideration (although see Deen & Hyams, 2002), and concluded that RIs are not permissible in early Swahili, the only remaining possibilities are listed in (28) above. Again, automated CLAN programs were used to count the number of each clause type, and the results are presented below. I first present the data by child in separate tables, after which I present the data by stage. For a month-by-month breakdown of the clause types for each child, see appendix 4C.

Table 4.3 Clause types in Hawa's files

Clause Type	Frequency
Full Clause	13 (13%)
[-SA] Clause	20 (20%)
[-T] Clause	18 (18%)
Bare Stem	47 (48%)
Total	98

Table 4.4 Clause types in Mustafa's files

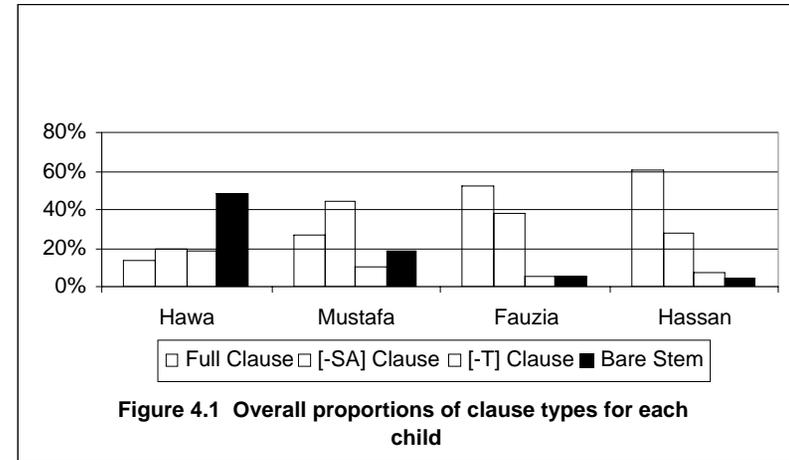
Clause Type	Frequency
Full Clause	136 (27%)
[-SA] Clause	225 (44%)
[-T] Clause	53 (10%)
Bare Stem	92 (18%)
Total	506

Table 4.5 Clause types in Fauzia's files

Clause Type	Frequency
Full Clause	183 (52%)
[-SA] Clause	135 (38%)
[-T] Clause	17 (5%)
Bare Stem	17 (5%)
Total	352

Table 4.6 Clause types in Hassan's files

Clause Type	Frequency
Full Clause	225 (61%)
[-SA] Clause	104 (28%)
[-T] Clause	26 (7%)
Bare Stem	15 (4%)
Total	370



As tables 4.3 – 4.6 show, all four possible clause types are attested in early Swahili (for examples of each clause type, see the sections on each clause type below). Let us start with a comparison of the four children. Recall we saw that Hawa is the least mature of the four children based on our three measures of linguistic maturity (MLU, V Ratio, %MPH), followed by Mustafa, Fauzia and Hassan. Beginning with Hawa in table 4.3, we see that the most frequent clause type in her files is the bare stem. Surprisingly, full clauses are the least frequent. This is surprising because full clauses are the most common indicative clause type in adult Swahili. We saw in chapter 2 that 95% of all indicative clauses in Swahili are full clauses. The other two clause types are represented in Hawa's data at between 18% and 20%, which is not insignificant. Turning to Mustafa, we see that bare stems are not the most frequent clause type in his files. Rather, Mustafa uses more [-SA] clauses than any other clause type. Furthermore, the second most common clause type is the full clause. These two clause types are the

only two that are permitted in the adult language. If we assume that Mustafa is at a stage in acquisition which is in advance of where Hawa is, then Mustafa is exhibiting a movement towards the adult target. These two clause types together represent 71% of all his indicative clauses (as compared to 33% for Hawa).

The remaining two children continue this move towards the adult norm in that they produce more full clauses than any other clause type. The second most frequent clause type in their files is the [-SA] clause, with full clauses and [-SA] clauses combining for 89% and 90% for Hassan and Fauzia respectively. However, while adults permit [-SA] clauses only 5% of the time, they occur in the speech of these children at rates of 28% and 38%. The trend then is that the less mature children are less adult-like in the clause types that they permit, while the more mature children are not fully adult-like, but closer to the adult norm than the others. This validates our staging procedure, because the children that we placed in lower stages based on three different measures exhibit less mature grammars than those placed in higher stages.

Let us turn to an analysis of the data by stage. This will unify the figures from tables 4.3-4.6 and allow us to see the development of each clause type in a much clearer fashion. Below are the frequencies of the four clause types by stage in table 4.7, followed by the proportions by stage¹³:

¹³ Monosyllabic placeholder used in unambiguous position (e.g., MPH-T-V-IND) were excluded from these counts. I determined whether the prefix was an MPH or a fully formed adult-like prefix based on context and pronunciation (see chapter 3). If an utterance included an MPH, it was not included in any of these counts.

Table 4.7 Frequencies of clause types, by stage

	Stage 1	Stage 2	Stage 3	Stage 4	Total
Files from	Hawa 2;2-2;6, Mus 2;0-2;3	Mus 2;4-2;8	Fau 1;8-2;2, Mus 2;9- 2;10	Has 2;10- 2;11	
Full Clauses	39	58	235	225	557
[-SA] Clauses	60	154	166	104	484
[-T] Clauses	42	25	21	26	114
Bare Stems	67	55	34	15	171
Total	208	292	456	370	1326

Table 4.8 Proportions of clause types, by stage

	Stage 1	Stage 2	Stage 3	Stage 4
Files from	Hawa 2;2-2;6, Mus 2;0-2;3	Mus 2;4-2;8	Fau 1;8-2;2, Mus 2;9-2;10	Has 2;10-2;11
Full Clauses	18.8%	19.9%	51.5%	60.8%
[-SA] Clauses	28.8%	52.7%	36.4%	28.1%
[-T] Clauses	20.2%	8.6%	4.6%	7.0%
Bare Stems	32.2%	18.8%	7.5%	4.1%
Total	208	292	456	370

These proportions in table 4.8 are the sum of each clause type in a particular stage divided by the total number of indicative clauses in that stage (i.e., they pool data across children). We see that the description given earlier about the less mature children permitting more unadult-like clauses is starkly evident in the staged data in table 4.7. Below is a line graph of the proportions. For comparison, I have included the relevant proportions from the adults' speech in this corpus (proportions taken from chapter 2) on the right-hand side of the figure.

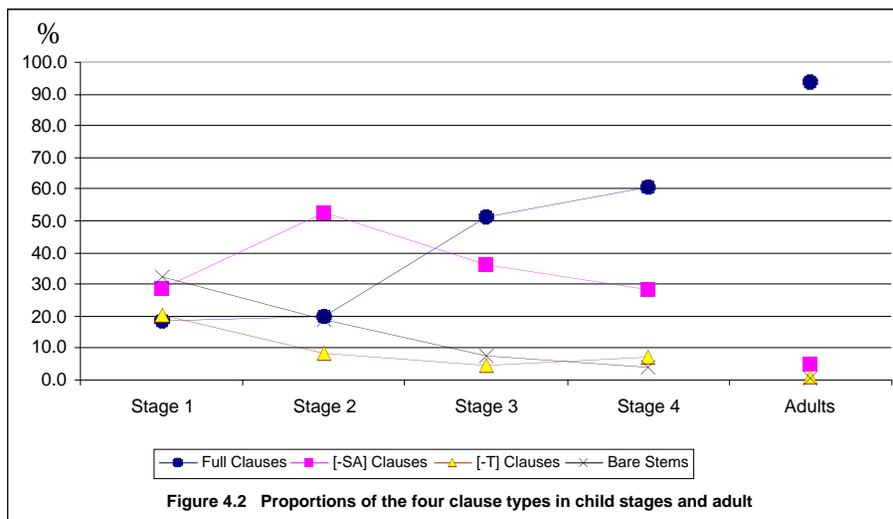


Figure 4.2 Proportions of the four clause types in child stages and adult

4.4.2 A note about Object Agreement

Before we consider each clause type in detail, a note about OA is required at this point. As discussed in chapter 2, OA is dependent on the specificity of the object: when the object is specific, OA is obligatory, but when the object is non-specific, OA is obligatorily absent. This raises a problem for determining obligatory contexts for OA because it is difficult to determine from context whether a child intends a specific or non-specific reading. However, some tentative findings are presented below. We will present results from two analyses that both suggest that children have knowledge of OA. In fact, the results will suggest that children's knowledge of OA may be ahead of their knowledge of SA inasmuch as we can tell from naturalistic data.

Each child did use OA, as table 4.9 shows:

Table 4.9 Number of OA in each child's files

Child	Number of OA
Hawa	35
Mustafa	40
Fauzia	95
Hassan	65
Total	235

The overall number of OA is lower than SA in the child files. A raw frequency count of the number of SA tokens in the child files reveals that SA occurs 978 times, while OA occurs 235 times. These numbers are not very meaningful because OA can only occur in transitive clauses, whereas SA can occur in all clauses. Children produced a total of 957 intransitive clauses and 1605 transitive clauses, as shown in tables 4.10-4.11:

Table 4.10 Total number of transitive clauses

Child	Number of transitive clauses
Hawa	283
Mustafa	557
Fauzia	466
Hassan	299
Total	1605

Table 4.11 Total number of intransitive clauses

Child	Number of intransitive clauses
Hawa	184
Mustafa	454
Fauzia	151
Hassan	168
Total	957

I first investigated whether children know that OA is only possible in transitive clauses¹⁴. I calculated the proportion of OA that occurred in transitive clauses and the proportion of OA that occurred in intransitive clauses, and the results are presented below¹⁵.

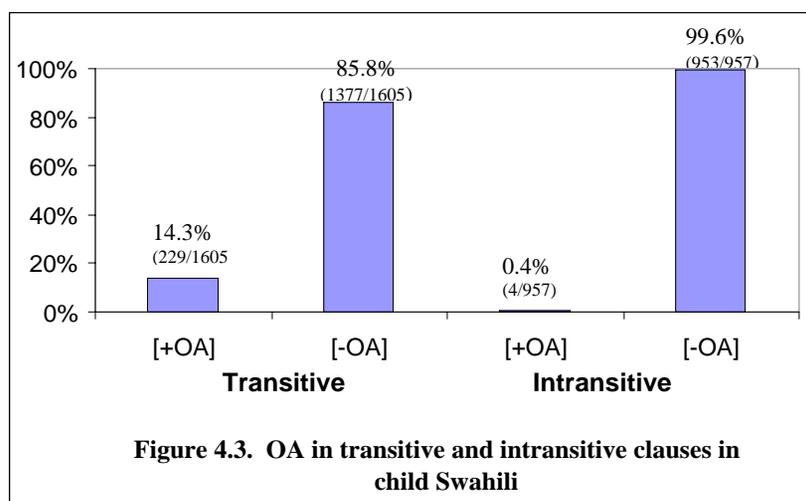
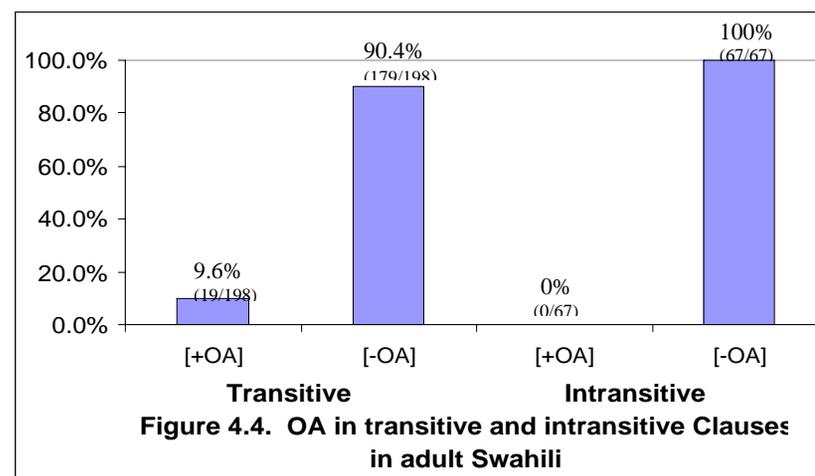


Figure 4.3 shows that while OA may be relatively infrequent in child Swahili, when OA occurs it is never on intransitive verbs. From this we can conclude that they at least know the function of OA in Swahili, that it is concerned with the presence of an object. As comparison, below is the figure for the adults in this corpus. Notice the striking similarity in

¹⁴I used the subcategorization frame of the verb to determine transitivity, since object drop of overt object DPs is grammatical in Swahili.

¹⁵The number of OA presented in this figure do not add up to 235 because this does not include distransitive clauses. When distransitive clauses are added in, the proportions are not greatly affected.

proportions. Not only are intransitives unmarked for OA, but OA in transitive clauses occurs at a similar rate as in child language.



The second analysis involves OA in contexts in which OA is obligatory. As mentioned earlier, obligatory context for OA is extremely difficult to determine from context, and so a comprehensive analysis of OA is not possible. However, the object may be obligatorily specific in some cases, e.g., when the object is a proper name.

The coding system unfortunately did not distinguish the grammatical roles of names, i.e., proper names were coded as names, with no distinction made between names that occur as subjects, names that occur as objects, names that occur as adjuncts, or names that occur as vocatives.

The total number of names in the corpus is 963, of which 189 are

postverbal¹⁶. These 189 utterances were inspected by hand, and it was found that only 27 of these 189 are unambiguously objects that require OA (the majority of the remaining names were vocative, easily identified by intonation). Of these 27, a total of 25 carry the appropriate OA (92.6%). The remaining two clauses that are missing OA are reproduced below:

- (29)
- a. h – u – Ø – beb – a Fauzia ? Has02, line 1012
 target: h – u – m – beb – a Fauzia?
 neg-SA_{2s}-OA_{3s}-carry-IND Fauzia
 ‘You don’t carry Fauzia?’
- b. Ø – on – a huyu Madua Fau07, line 1232
 target: m – on – a huyu Madua
 OA_{3s}-see-IND him Madua
 ‘See (him) Madua.’

In examples (29a-b) the OA has been omitted in obligatory contexts. Example (29a) is a negative utterance in which the object is the name Fauzia, and (29b) is an imperative.

The overall numbers are not very large, and so more data is required in order to verify this result. However, if these results are shown to be valid, we must conclude the following. First, Swahili children never use OA in contexts in which it should not be used (i.e., intransitive clauses). Thus errors of commission involving OA are unattested. Second, errors of omission involving OA are also rare. There are only two cases of OA omission out of 27 obligatory contexts – an omission rate of 7.4%. I acknowledge that the frequencies are small and so I reserve full judgment

¹⁶ Limiting the inspection to postverbal context was done for the sake of convenience. Children at this age do topicalize objects, and so this undoubtedly resulted in a reduction in tokens counted.

on this issue, but I present the data simply as suggestive and as a point of comparison with SA omission.

4.4.3 Full Clauses

Full clauses (FCs) are easy to identify in the corpus: all verbs that have SA and T. This includes those with and without OA. FCs occur with a full range of tense markers (examples 30a-f) as well as a wide range of SA (1,2,3 singular, and 1,3 plural), as well as more complex constructions. Example (30h) is a complex tense construction (grammatical in adult Swahili) and is composed of an auxiliary verb *kuwa* ‘to be’ that is fully inflected with SA and T, followed by a main verb that is also fully inflected with SA and T¹⁷. Example (30i) is a biclausal utterance with two fully inflected clauses (grammatical in adult Swahili).

- (30)
- a. ni – na – zim – a tuu ile Haw06, line 1160
 SA_{1s}-pres-finish-IND only that
 ‘I just finish that’
- b. a – ta – la – a Mus10, line 1512
 SA_{3s}-fut-sleep-IND
 ‘She will sleep’
- c. mimi ...ni – li – tow – a Mus12, line 1546
 I SA_{1s}-past-remove-IND
 ‘I...removed (it)’

¹⁷ Notice that the SA on both verbs are identical, an interesting feature of complex tenses that argues for an agreement analysis of SA since it would be unusual to have two identical subjects in the same clause. See chapter 2 for more discussion on this point.

- d. sku gani u – ta – end – a ? Fau03, line 226
 day which SA_{2s}–fut–go–IND (Future tense)
 ‘On what day will you go?’
- e. ah weh u – sha – um – a Fau07, line 1090
 you SA_{2s}–pr.prf.–hurt–IND (Past perfective)
 ‘Ah, you have already hurt (me)’
- f. a – ka – end – a Has01, line 547
 SA_{3s}–cont–go–IND (Continuative tense)
 ‘He (then) went.’
- g. wa – na – kul – a pilau Has01, line 485
 SA_{3pl}–pres–eat–IND pilau (Plural SA)
 ‘They eat Pilau.’
- h. tu – li – kuw – a tu – na – kul – a pilau! Has01, line 794
 SA_{1pl}–past–be–IND SA_{1pl}–pres–eat–IND pilau (Plural SA and
 ‘We were eating Pilau.’ complex tense)
 (lit: We were we are eating Pilau)
- i. Baba, u – li – sem – a Has01, line 333
 Father SA_{2s}–past–say–IND (Embedded clause)
- u – ta – nunu – a nini ?
 SA_{2s}–fut–buy–IND what
- ‘Father, what did you say you will buy?’
 (lit: Father, you said you will buy what?)

Each of these examples is completely adult-like. Occasionally utterances occur that have a reduced SA, as in example (31) below, where [ni] is reduced to a syllabic nasal. Such cases were categorized as FCs because in each such case the reduction was consistent with adult spoken language.

- (31) n – na – on – a Haw06, line 1720
 prefix – pres – see – IND
 unreduced target: ni – na – on – a
 SA_{1s} – pres – see – IND
 ‘I see.’

Furthermore, no cases occurred in which the prefix did not phonologically resemble the well-formed target. For example, no cases of [] or [∅] were used instead of [ni]. Therefore such cases of reduction were not counted as MPHs, but as full SA prefixes.

Finally, FCs occur with overt as well as null subjects, with and without OA, and they occur with grammatical function changing suffixes. In this respect, they are adult-like (see sections 2.4.4 (OA), 2.4.6 (suffixes) and 2.9 (null subjects) for discussion of adult Swahili).

4.4.4 [-SA] Clauses

[-SA] clauses are the most frequent clause types in the early stages, accounting for more than half of all the indicative clauses in stage 2. There are several features of [-SA] clauses that I will illustrate: They occur with overt subjects as well as null subjects; with OA and without OA; and with grammatical function changing suffixes such as applicative or passive. In each case [-SA] clauses occur with various T markers.

[-SA] clauses occur with overt subjects (32) as well as with null subjects (33). For a full discussion of this phenomenon, see chapter 5. Irrespective of whether the subject is overt or null, a wide range of tense markers occur in these clauses.

- (32) **Overt subjects:**

a. mimi Ø – na – ruk – a Mus22, line 478
 target: mimi ni – na – ruk – a (Present tense)
 I SA_{1s}-pres-jump-IND
 ‘I am jump down.’

b. baba Ø – li – tok – a ku – jeng – a Has02, line 255
 target: Baba a – li – tok – a ku – jeng – a (Past tense)
 Baba SA_{3s}-past-leave-IND inf-build-IND
 ‘Father left to build’
 (Father is a construction worker)

c. mimi Ø – li – m – chap – a Fau04, line 116
 target : mimi ni – li – m – chap – a (Past tense)
 I SA_{1s}-past-OA_{3s}-slap-IND
 ‘I slapped him.’

d. ndege Ø – me – end – a juu Mus21, line 172
 target: ndege i – me – end – a juu (Pres.perfect)
 bird SA_{inanim}-p.perf.-go-IND up
 ‘The bird has gone up.’

e. mimi Ø – ta – ingi – a apa ndani Fau08, line 160
 target: mimi ni – ta – ingi – a hapa ndani (future tense)
 I SA_{1s}-fut-enter-IND here inside
 ‘I will enter here inside.’

(33) **Null Subject:**

a. Ø – na – lal – a Haw07, line 681
 target: ni – na – lala – a (Present tense)
 SA_{1s}-pres-sleep-IND
 ‘I am sleeping.’

b. Ø – li – tumbu – a Mus18, line 378
 target: a – li – ni – sumbu – a (Past tense)
 SA_{3s}-past-OA_{1s}-disturb-IND
 ‘He disturbed me.’

d. Ø – ta – po – a Mus20, line 431

target: ni – ta – po – a (Future tense)
 SA_{1s}-fut-relax-IND
 ‘I will relax.’

e. Ø – me – tok – a juu Mus21, line 155
 target: i – me – tok – a juu (Pres. Perfect)
 SA_{inanim}-p.perf.-leave-IND up
 ‘It has left (and gone up).’

f. Ø – ka – nunu – a soda Fau09, line 1015
 target: a – ka – nunu – a soda (Continuative)
 SA_{3s}-cont-buy-IND soda
 ‘He then bought soda.’

OA occurs in [-SA] clauses, although as mentioned earlier, OA in general is less common than SA. Therefore the variety of tense markers that we see is not as wide.

(34)

a. Ø – ta – ku – pig – a Mus20, line 135
 target: ni – ta – ku – pig – a
 SA_{1s}-fut-OA_{2s}-hit-IND
 ‘I will hit you.’

b. Ø – ta – i – pig – a Fau08, line 801
 target : ni – ta – i – pig – a
 SA_{1s}-fut-OA_{inanim}-hit-IND
 ‘I will hit it.’

c. Ø – ta – ni – p – i – a mimi ? Has03, line 1624
 target: u – ta – ni – p – i – a mimi?
 SA_{2s}-fut-OA_{1s}-give-applic-IND me
 ‘Will you give me (it)?’

Finally, [-SA] clauses occur with grammatical function changing suffixes such as the applicative (35a and 35b) and the passive (35c and 35d).

Example (35e) is the only case of a [-SA] clause occurring with stacked

suffixes. This extremely complex example has an applicative, causative and a passive suffix.

- (35)
- a. Ø – ta – nunu – li – a hii soda? Fau06, line 528
 target: u – ta – nunu – li – a hii soda?
 SA_{2s}–fut–buy–applic–IND this soda
 ‘Will you buy (for me) this soda?’
- b. weh Ø – na – imb – i – a ...kalamba Fau09, line 1096
 target: Wewe u – na – imb – i – a Kalamba?
 You SA_{2s}–pres–sing–applic–IND Kalamba
 ‘Will you sing Kalamba (for me)?’
- c. Ø – na – it – w – a Tafa Fau03, line 808
 target: a – na – it – w – a Tafa
 SA_{3s}–pres – call – passive – IND Tafa
 ‘He is called Tafa.’
- d. gubi Ø – me – dung – w – a Mbize ? Mus23, line 46
 target: Kumbikumbi i – me – dung – w – a Mbize?
 White ant SA_{3s}–p.perf.–dig–passive–IND Mbize
 ‘Has the white ant dug (into the ground), Mbize?’
- e. Ø – na – vaa – li – sh – w – a Fau08, line 1649
 target: a – na – vaa – li – sh – w – a
 SA_{3s}–pres–wear–applic–causative–passive–IND
 ‘He is dressed’
 (lit.: He is caused to have the clothes to be put on.)

The majority of [-SA] clauses, however, are simple mono-prefixal clauses (i.e., tense) with no suffixes other than the mood final vowel. In some cases there was some difficulty in categorizing an utterance as either a [-SA] clause or something else. Recall from section chapter 3 that Swahili children produce MPHs (mono-syllabic place holders, Bottari *et al.*, 1993).

Refer to section chapter 3 for a discussion on how MPHs were distinguished from other well-formed prefixes.

Turning now to the development of [-SA] clauses, we see that in stage 1 they occur at 28.8%. They increase in stage 2 to over 50% and then diminish, but by stage 4 they are still very common, with 28.1% of all indicative clauses missing SA. Notice that the starting proportion (28.8%) and the final proportion of [-SA] clauses (28.1%) is approximately the same. There are several way to interpret this data, but one thing that this suggests is that while there is a surge in the number of [-SA] clauses in stage 2, this may be attributable to factors other than development. For example, it may be that Mustafa (the only child in stage 2) uses an unusually high proportion of [-SA] clauses. Indeed, a count of Mustafa’s files reveals that [-SA] clauses (across all 11 months) account for 44.5% (225/506) of all his indicative clauses. This compared to 18% for Hawa, 38% for Fauzia and 28% for Hassan. Therefore, while the staging process is useful in allowing us to see overall trends, we must take into account individual variation among children. The conclusion is that [-SA] clauses show little development across the four stages. In adult Swahili [-SA] clauses occur at a rate of approximately 5%, which is in stark contrast to child Swahili (at any stage or for any child). Therefore the principles that are responsible for the reduction of [-SA] clauses in child language develop at a stage beyond the last files in this corpus. This is an area for future research.

4.4.5 [-T] Clauses

[-T] clauses are the least common clause type amongst the four, and are the first to diminish to under 10%. They begin at a rate of over

20%, but immediately fall to under 10% in stage 2, and remain under 10% through to stage 4. Therefore I consider stage 2 the point at which [-T] clauses cease to be a possibility in early Swahili. Recall that [-T] clauses are completely ungrammatical in adult Swahili, not occurring in naturalistic discourse nor being accepted in elicitation by native speakers (see chapter 2).

Although a few [-T] clauses occur with OA (as in example (36) below), they primarily occur without OA.

(36) a – Ø – mw – on – a Haw06, line 1613
 target form: a – na – mw – on – a
 SA_{3s}–pres– OA_{3s}–see–IND
 ‘He sees her.’

Categorizing utterances as [-T] clauses occurred through a process of elimination. [-T] clauses occur with a single prefix, and so there is a possibility of misclassifying a [-T] clause as [-SA], a clause with an MPH, or a subjunctive clause with SA as the single prefix:

(37) a. [-T] clause: SA – V – IND
 b. [-SA] clause: T – V – IND
 c. MPH clause: prefix – V – IND
 d. Subjunctive: SA – V – SUBJ

For the most part, [-SA] clauses occurred with a single syllabic prefix on the verb. In most cases the verb stem itself was marked for mood and nothing else (i.e., no additional suffixes such as applicative, passive, etc.). Categorizing utterances as [-SA] or [-T] was relatively easy since the respective morphemes are very distinguishable. However, distinguishing between [-T] clauses and verbs with an MPH prefix is sometimes more difficult because there are two SA markers that are vowels with no onset consonant (2nd person singular [u] and 3rd person singular [a]). However, MPHs are distinguished from SA markers by the fact that they are reduced

vowels or single nasals, none of which occur as SA (see earlier discussion of MPHs and the method used to determine whether a prefix is an MPH or not).

[-T] clauses have the morphological structure given in (37a) above. They contain SA immediately followed by the verb root and the indicative mood final vowel ([a]). A morphologically similar clause is the subjunctive, schematized in (37d). Notice that the only difference between these two clause types is the final vowel ([a] for indicative, [e] for subjunctive). I have been considering clauses of the kind in (37a) as full clauses with a missing T. However, it is possible that they are in fact subjunctive clauses with a mispronounced final vowel.

If [-T] clauses are subjunctive clauses with a mispronounced final vowel, this represents an error of commission. Recall that errors of commission are very rare in child language in general. On the basis of what we know about other child languages, it is unlikely that children mispronounce the final vowel in [-T] clauses. So on conceptual grounds we reject this analysis, and I show empirically below that this cannot be the case.

Deen & Hyams (2002) show that children use indicative/subjunctive morphology correctly in Swahili. In all the indicative clauses in this corpus, children misuse indicative morphology 23/1436 (1.6%) times. In other words, 1.6% of verbs that are marked as indicative are used in subjunctive contexts. In adult Swahili, indicative marking co-occurs with temporal marking. Because of this relation between mood and tense, we speculate that this low percentage of mood errors in child Swahili may be due to the presence of T in some of those cases (i.e., T may be forcing a temporal, non-subjunctive interpretation), and so we calculate the

proportion of misused indicative morphology in verbs that are missing T: bare stems. Of 164 bare stems that are marked indicative, 5 occur in subjunctive contexts (3%). Thus it is not the case that the presence of T forces an indicative interpretation. Our conclusion is that Swahili children do not misuse indicative morphology when they intend to use subjunctive morphology, whether T is present or not.

Similarly, of the 114 [-T] clauses, there are 16 cases in which it is unclear what the intention of the child was. Therefore these 16 cases are not considered in this particular analysis (but are included in the general proportions presented thus far). Of the remaining 98 [-T] clauses, 8 are compatible with a subjunctive-type meaning, i.e., they could have been intended as a request, or a suggestion, or an expression of possibility or desire, etc.¹⁸. This represents an ‘error’ rate of 8/98 (8.2%). When we combine these numbers with other clauses that are missing T (i.e., bare stems), the overall ‘error’ rate is 13/262 (4.9%). Thus I conclude that errors of this sort are rare. Therefore, because the interpretation of [-T] clauses is consistent with indicative morphology, I assume that [-T] clauses are not final vowel errors, but represent errors of omission in which T has been omitted from a full clause.

A variety of [-T] examples are given in (38) below. Examples (38a-c) are from Fauzia. Notice that each utterance has a different SA – 1st, 2nd, and 3rd person singular. Examples (38d-e) from Hawa also have different SA markers, as do the examples from Hassan and Mustafa. This shows that SA is used productively in [-T] clauses and reference is not limited to any particular person. Furthermore, a wide range of intended

temporal meanings are represented: present tense (38a, d, g and h), past (38b, c, e, f and j), and present perfect (38i). Therefore temporal reference appears to be free in [-T] clauses.

- (38)
- | | | |
|---------|---|------------------|
| a. | ni – Ø – kw – ambi – a | Fau04, line 95 |
| target: | ni – na – kw – ambi – a
SA _{1s} –pres–OA _{2s} –tell–IND
‘I am telling you.’ | |
| b. | u – Ø – kingi – z – a ? | Fau07, line 1044 |
| target: | u – li – kimbi – z – a?
SA _{2s} –past–run–causative–IND
‘Did you make (him) run?’ | |
| c. | a – Ø – sem – a nini ? | Fau04, line 464 |
| target: | a – li – sem – a nini?
SA _{3s} –past–say–IND what
‘What did he say?’ | |
| d. | ni – Ø – chun – a | Haw05, line 682 |
| target: | ni – na – chun – a
SA _{1s} –past–pinch–IND
‘I am pinching!’ | |
| e. | alafu a – Ø – rud – i | Haw07, line 905 |
| target: | alafu a – li – rud – i
then SA _{3s} –past–return–IND
‘Then he returned.’ | |
| f. | alafu a – Ø – kuj – a | Has01, line 1178 |
| target: | alafu a – li – kuj – a
then SA _{3s} –past – come – IND
‘Then he came.’ | |
| g. | u – Ø – on – a nimó | Has01, line 1396 |
| target: | u – na – on – a Nimó?
SA _{2s} –pres – see – IND Nimó
‘Do you see Nimó?’ | |

¹⁸ One had a desiderative interpretation, one was used as a request, two expressed deontic necessity, and four were suggestions.

- h. u – Ø – gong – a hapa Mus04, line 1150
 target: u – na – gong – a hapa
 SA_{2s}– pres – knock – IND here
 ‘You are knocking here.’
- i. ni – Ø – pit – a Mus05, line 316
 target: ni – me – pit – a
 SA_{1s}– p.perf. – pass – IND
 ‘I have passed (through).’
- j. a – Ø – tup – a Mus03, line 2163
 target: a – li – tup – a
 SA_{3s}– past – throw – IND
 ‘He threw (something).’

As we will see in more detail, subjects generally do not occur in [-T] clauses. I will delay discussion of this point, and refer the reader to chapter 5 for a full discussion of the details and reasons for this.

4.4.6 Bare Stems

Bare stems are verbs that are missing SA and T, but have a mood final vowel. In stage 1 they are the most common clause type of the four attested indicative clause types. They quickly diminish, going from 32.3% in stage 1, to 18.8% in stage 2, to under 10% by stage 4. By stage 4 bare stems occur a mere 4% of the time. Examples of bare stem clauses from each child are presented below. In (39a-e) we have examples from Hawa’s files. Notice that the interpretation includes past, present and present

perfective. Also, the persons are 1st or 3rd person¹⁹. The other children show similar characteristics, with example (39j) from Hassan’s files showing a continuative reading.

- (39)
- a. Ø – Ø – anguk – a hivi Haw07, line 210
 target: ni – li – anguk – a hivi
 SA_{1s}–past–fall–IND like this
 ‘I fell like this.’
- b. Ø – Ø – tak – a tuwadh–a Haw07, line 5
 target: ni – na – tak – a tuwadh–a
 SA_{1s}–pres–want–IND bathe–IND
 ‘I want to bathe.’
- c. Ø – Ø – ndik – a Haw05, line 2946
 target: ni – me – andik – a
 SA_{1s}–p.perf.–write–IND
 ‘I have written.’
- d. Ø – Ø – ka – a hapa Haw04, line 130
 target: a – na – ka – a hapa
 SA_{3s}–pres–live–IND here
 ‘She lives here.’
- e. Ø – Ø – end – a job Haw03, line 128
 target: a – me – end – a job
 SA_{3s}–p.perf.–go–IND job
 ‘He has gone to work.’

¹⁹ Deen & Hyams (2002) find that the bare stems in Swahili do not have a modal meaning, parallel to English bare stems (Hoekstra & Hyams, 1998). Of the 164 bare stems in the corpus, only 5 (3%) occurred in contexts that were compatible with a modal meaning. This supports the idea that infinitival morphology licenses MoodP. Thus the irrealis meaning of RIs (Hoekstra & Hyams’ Modal Reference Effect) is not a result of a simple lack of temporal specification, but rather arises from the presence of infinitival morphology.

f.	aya, Ø – Ø – annguk – a target: aya, ni – na – anguk – a SA _{1s} -pres – fall – IND 'Aya, I am falling.'	Mus19, line 930
g.	Ø – Ø – um – a target: a – li – ni – um – a SA _{3s} -past - OA _{1s} -hurt – IND 'He hurt me.'	Mus17, line 632
h.	Ø – Ø – kata – a target: ni – na – kata – a SA _{1s} -pres-refuse-IND 'I refuse.'	Fau06, line 863
i.	Ø – Ø – kasilik – a target: ni – me – kasirik – a SA _{1s} -p.perf.-get angry-IND 'I am angry.' (lit: I have gotten angry)	Fau01, line 1709
j.	Ø – Ø – shuk – a chini target: a – ka – shuk – a chini SA _{3s} -cont-align-IND down 'He then got down.'	Has01, line 687
k.	Ø – Ø – kul – a target: ni – na – kul – a SA _{1s} -pres-eat-IND 'I am eating.'	Has03, line 855
l.	baba Ø – Ø – end – a hoa target: Baba a – me – end – a hoa Baba SA _{3s} -p.perf.-go-IND hoa 'Father has gone home.'	Has03, line 1309

As we did for the other clause types, we will now consider the coding of bare stems. I have presented bare stems as having underspecified Tense and Agr. However, given the frequency of RIs in languages such as

German and French, it is possible that Swahili bare stems are actually RIs with the infinitival prefix missing. After all, an infinitival verb (INF-V-IND) would have the correct metrical structure so as to allow metrical omission of the INF prefix (assuming that SA is syntactically omitted).

The infinitive prefix, being in a preverbal position, would take weak stress. Being in a pre-trochaic position, it would thus be subject to omission by the MOM.

(40)	w s w	→ w [s – w]
syllabic:	ku-en-da	
morphemic:	ku – end – a	
	inf – go – IND	
	'to go'	

Wexler (1994) suggests that English bare verbs be assimilated to RIs. Under this view English bare stems and Swahili bare stems are considered RIs that appear as bare verbs because of language-specific morphology. The appeal of this idea is that it unifies the analysis of English bare verbs, Swahili bare stems and RIs. Under this view, children acquiring English produce bare verbs while German, Dutch and French children produce RIs because of morphological differences between these languages: the infinitive in English is not a true infinitive in that it is not a bound affix on the verb (material may occur between the verb and the particle: 'to boldly go where no one has gone before'). Furthermore, the infinitive marker in English is preverbal while in RI languages it is a suffix. Moreover, adult English uses bare verbs in contexts in which adult RI languages use infinitives, e.g., as complements to causatives, or perception verbs.²⁰

²⁰ Thanks to Carson Schütze for discussion on this point and the French example.

- (41) Il nous a fait chanter.
 French
 He us has made sing-infin
 'He made us sing.'

However, true RIs generally have a modal or irrealis meaning (Wijnen, 1987; Blom & Wijnen, 2001). This has been shown to be true for Dutch, French, German and Swedish, although the evidence for other RI languages is not so clear. However, English bare verbs, as noted by Hoekstra & Hyams (1998), tend to have a temporal meaning, usually present tense here-and-now, but also past tense (Deen, 1997; Madsen & Gilkerson, 1999; Torrence, 2002). Hoekstra & Hyams attribute the irrealis meaning associated with RIs to the presence of infinitival morphology, which they note, cross-linguistically is generally associated with irrealis or modal meaning in adult languages. Therefore there is a clear semantic difference between true RIs and English child bare verbs. This argues against an RI analysis of English bare verbs.

As we have seen in this section, Swahili children produce bare verbs, as English-acquiring children do. We saw in section 4.4.5 that the interpretation of bare verbs in Swahili is overwhelmingly non-modal (5/164 bare verbs occur in irrealis contexts), and thus Swahili bare verbs pattern with English bare verbs. Moreover, as in English, RIs are completely unattested in Swahili. Thus, at first glance Swahili appears to be parallel to English in many ways. However, unlike English, the infinitive in Swahili is a true infinitive. It is a bound affix on the verb. The only other material that can occur between it and the verb is OA, which is expected given that AgrOP occurs below TP. Finally, complements to perception verbs occur in the subjunctive or as fully inflected verbs, but not as bare verbs. Thus

Swahili does not have the same properties of English that purportedly result in bare verbs in child language.²¹

Hoekstra & Hyams (1998) showed that significant differences exist between RIs and bare verbs in child language. The Swahili facts presented here show that despite having a true infinitive in the adult language, bare verbs are a genuine possibility in child Swahili.

4.4.7 Tense Omission versus SA omission

Thus far we have seen that Swahili children permit four clause types: full clauses, [-SA] clauses, [-T] clauses and bare stems. We saw in chapter 2 that Swahili adults only allow two clause types: full clauses and [-SA] clauses. Therefore the crucial difference between adult Swahili and child Swahili is the omission of T. Let us therefore examine clauses that are missing T as a single class. Below is a graph showing the occurrence of tenseless clauses ([-T] clauses and bare stems) across the developmental

²¹ Additionally, several of the syntactic correlations that have been noted in RIs in true RI languages (German, Dutch, French, etc.) do not hold in English. For example, null subjects predominate in non-finite contexts in RI languages, but not in English (see section 1.2.3). Also, RIs generally do not occur in wh- questions in languages such as German and Dutch, but in English wh- questions occur with bare verbs (see Roeper & Rohrbacher 1994 and Bromberg & Wexler, 1995, who look at wh- fronting and null subjects in finite and non-finite contexts). Therefore English bare verbs differ from true RIs in some significant syntactic ways. These tests cannot be applied to Swahili because Swahili allows null subjects and wh- words remain in situ.

stages and the occurrence of those clauses in which SA is omitted (including [-SA] clauses and bare stems, labeled ‘SA-less’).²²

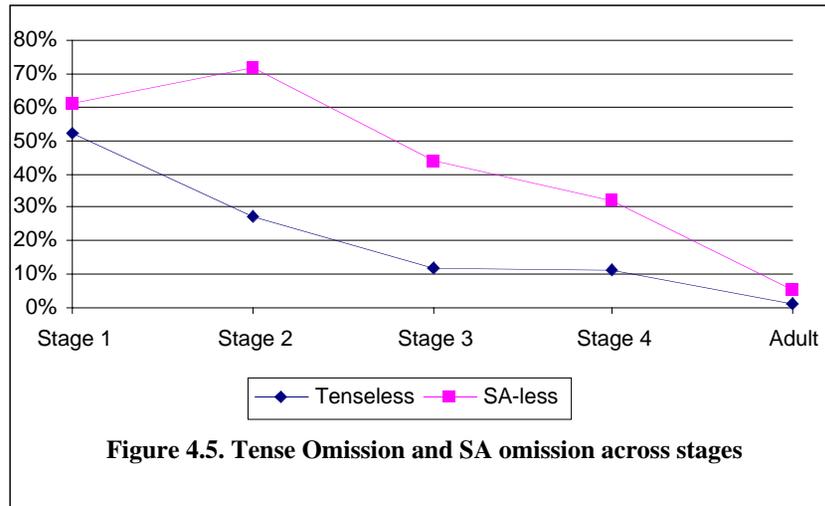


Figure 4.5. Tense Omission and SA omission across stages

Overall, the omission of SA and the omission of T tend to diminish as the children mature. However, tense omission ceases to be a possibility in early Swahili much earlier than SA omission. By stage 3, tense omission is rare (12%), while SA omission at that same stage occurs at a rate of more than 40%. Similarly while T omission in stage 4 is rare, SA omission is still above 30%. Another way to view this is that SA omission is tolerated longer than T omission. We will return to this point in the section in which

²² Including bare stems in both categories is necessary because I wish to compare SA omission on the one hand and T omission on the other. Bare stems fall into both categories.

we review ATOM’s (Schütze & Wexler, 1996) applicability to the Swahili data.

4.4.8 Summary

We investigated the occurrence of various verbal clauses in the speech of four Swahili speaking children, and found that not only do they allow clause types that adults allow, they also allow two clause types that are unattested in adult speech. We found that full clauses are used relatively sparingly at early stages in development, but develop into the majority clause type by stage 4. We also found that [-SA] clauses are used very frequently, and remain a genuine possibility beyond stage 4. We saw that the use of tense in these clauses is fully productive, and the reference of the missing SA is free. Overt subjects and null subjects occur in both full clauses as well as [-SA] clauses.

In addition to full clauses and [-SA] clauses (both of which are attested in adult Swahili), children produce verbal clauses that are missing T (which we call [-T] clauses) and verbal clauses that are missing both T as well as SA (which we call bare stems). We saw that [-T] clauses cease to be a possibility fairly early in development, falling to under 10% by stage 2. Bare stems also diminish, but do not fall to under 10% until stage 4. We saw that [-T] clauses allow a range of temporal references and occur with a range of SA markers, while bare stems allow for a range of temporal meanings as well as persons. In both cases we found that they were generally incompatible with a subjunctive-type meaning, and so we concluded that they were indicative clauses with missing T and/or SA.

Finally, we noted that T omission is unattested in adult Swahili, but is a possibility in child Swahili. SA omission, on the other hand, is

possible in both adult Swahili as well as child Swahili. Seemingly aware of this difference, Swahili children tolerate the omission of SA even into stage 4, but T omission ceases to be a significant phenomenon in stage 4.

In the next section, we will re-evaluate the theories presented in section 4.2 in light of what we now know about child Swahili. We will see that none of the theories fully explain everything that we have noted, but Schütze & Wexler's (1996) ATOM captures the general facts of Swahili. There are several facts, however, which remain unexplained even by ATOM.

4.5 Revisiting Acquisition Theories

In section 4.2 we discussed five influential theories of the acquisition of early morphosyntax. We discussed the Metrical Omission Model which proposes that children omit syllables that occur in particular metrical configurations. We also discussed Rizzi's (1994) Truncation Hypothesis, which proposes that children can optionally specify the root of the clause as a lower projection than CP. We then discussed three underspecification theories: an underspecification of T theory (Wexler, 1994), an underspecification of Agr theory (Clahsen *et al.* 1996) and a theory of underspecification of both T and Agr (Schütze & Wexler, 1996). We will evaluate each of these five theories in the following sections.

4.5.1 Metrical Omission Model

At first glance, the MOM appears to have promise. The morphemes that are optionally omitted in Swahili are individual syllables, and so can be individually targeted for metrical omission. The inflectional prefixes in Swahili are preverbal, and we know from chapter 2 that Swahili

is a trochaic language. However, recall our predictions from section 4.2.6. We saw that in verbal complexes with disyllabic verb stems, MOM predicts no omission (42a)²³. In verbal complexes with trisyllabic verb stems, MOM predicts omission of only the onset syllable in the verb stem²⁴. Therefore no omission of prefixes is predicted.

- (42) a.
$$\begin{array}{ccccccc} S & - & W & - & S & - & W & & \rightarrow & [S - W] & [S - W] \\ \vdots & & \vdots & & \vdots & & \vdots & & & & \\ ni & - & ta & - & pik & - & a & \text{ (chakula)} & & & \\ SA_{1s} & - & fut & - & cook & - & IND & \text{ (food)} & & & \\ \text{'I will cook food'} & & & & & & & & & & \end{array}$$
- b.
$$\begin{array}{ccccccc} S & - & W & - & W & - & S & - & W & & \rightarrow & [S - W] & W & [S - W] \\ \vdots & & & & & \\ ni & - & ta & - & anguk & - & a & & & & & & & \\ SA_{1s} & - & fut & - & fall & - & IND & & & & & & & \\ \text{'I will fall'} & & & & & & & & & & & & & \end{array}$$

²³ One objection could be that the addition of OA may affect the parsing of the string such that metrical omission becomes relevant. However, the crucial fact is that SA carries secondary stress, and hence is a strong syllable. Two side-by-side strong syllables are dispreferred, and so because SA is a strong syllable, T is a weak syllable. Therefore SA and T will always form a trochaic foot. The addition of OA will only add a weak syllable between the trochaic foot of the prefixes and the verb stem. Therefore MOM may have predictive power for the omission of OA, but it cannot explain the omission of SA or T. The addition of any other affixes (e.g., the applicative suffix, the passive suffix, etc.) will only affect the metrical structure of the latter portion of the string and will leave SA and T in a trochaic foot.

²⁴ In fact, there is sporadic evidence that this may be correct: verbs such as *anguka* occur frequently as [uka] – the onset [a] being omitted. This occurs when the verb is bare, as well as when the verb has some or all prefixes. Thus this is evidence that MOM actually does apply as Gerken (1991) proposes, but it does not explain the omission of the inflectional prefixes.

This is plainly contradicted by the data. Not only are prefixes omitted by children, at stage 1 over 80% of indicative verbs are missing either SA, T or both. The overwhelming frequency of prefix omission thus is not compatible with MOM. This is in contrast to Demuth's (1994) conclusion that the MOM (at least in part) accounts for the omission of noun class prefixes in Sesotho.

4.5.2 Truncation Hypothesis

In section 4.2.6 we discussed the predictions that Truncation makes for Swahili and we listed the clause types that we expect if Truncation applies in early Swahili. We saw earlier that Mood is never omitted. In terms of Truncation we interpret this as truncation always occurring above Mood (see below). We also saw earlier that it is difficult to determine precisely whether OA is omitted or not (although the little evidence that there is suggests that it isn't omitted). Putting aside OA for this reason, Truncation makes the following predictions:

(43) Predictions for Truncation:

- (i) Full clauses should optionally occur
- (ii) [-SA] clauses should optionally occur
- (iii) Bare stems should optionally occur
- (iv) [-T] clauses should be impossible.

Clearly predictions (i) – (iii) hold, which provides evidence for Truncation. However, prediction (iv) is false, as we saw that in early stages [-T] clauses occur at substantial levels (over 20% in stage 1). Truncation can not account for [-T] clauses because they contain SA (which necessarily entails that the root has been specified at or above AgrSP), but are missing T (which occurs below AgrSP). Because the root is specified as AgrSP or

above, the structure must project at least up to SA, including TP. Therefore whatever the reason for T omission in [-T] clauses, it cannot be due to truncation.

Furthermore, we must not dismiss the fact that Mood is obligatory in every utterance. Truncation makes no distinction between functional projections other than their respective hierarchical positions. Without the stipulation of a lower limit of what can be specified as the root, Truncation holds no potential to explain why Mood is obligatory and not any of the other functional affixes.

4.5.3 Underspecification of T

Wexler (1994) argued that in English and other RI languages children have the option to underspecify T. For Swahili, as we saw in section 4.2.6, the underspecification of T hypothesis predicts that child grammar permits full clauses and [-T] clauses. While these clause types are attested, they are by no means the only clause types that occur in child Swahili. In fact, at stage 1, full clauses and [-T] clauses account for only 39% of all indicative verbs, suggesting that the underspecification of T is only partly responsible for the typology of child clauses.

4.5.4 Underspecification of Agr

Clahsen *et al.* (1996) argued that AgrS is optionally underspecified. We saw that their theory predicts that Swahili children should allow full clauses and [-SA] clauses. However, they cannot account for the occurrence of [-T] clauses and bare stems. In fact, neither underspecification of T nor underspecification of Agr can account for bare stems.

4.5.5 ATOM

The natural conclusion from the preceding discussion is that the underspecification of T as well as Agr is permitted by child grammar – precisely the proposal of Schütze & Wexler (1996). They propose that both Agr and T can be optionally and independently omitted, resulting in the predictions that Swahili children should allow full clauses (no underspecification), [-SA] clauses (AgrS underspecification only), [-T] clauses (T underspecification only) and bare stems (underspecification of both AgrS and T). Thus our results are consistent with ATOM.

4.6 Remaining questions

While ATOM does predict the four clause types that occur in early Swahili, there are several aspects of the Swahili data which ATOM does not directly explain. We will now consider three results that ATOM does not address directly.

The first issue that ATOM does not account for is the difference between SA omission and T omission. We saw earlier that T omission is less frequent than SA omission at all times, and ceases to be a possibility well before SA omission ceases to be a possibility. ATOM does not predict this difference, but it is not incompatible with ATOM either. There is nothing in ATOM that either predicts this or rules this out.

The second issue is that of bare stems. According to ATOM, underspecification of Agr and T are independent. Therefore the omission of both SA and T (i.e., bare stems) should be the product of the independent omission SA and the independent omission of T. That is, the rate of bare stems should be related to the rates of [-SA] clauses and [-T] clauses since bare stems occur through the independent omission of SA and T. This is

clearly not the case in the data. In every stage except stage 4, bare stems outnumber [-T] clauses, which is unexpected if bare stems are a product of [-T] and [-SA] clauses.

However, Carson Schütze (p.c.) points out that this prediction holds only on individual child data and not on staged data. ATOM says nothing about aggregated data, and must be evaluated on a file-by-file basis. I include such data in Appendix 4C, and an inspection of this data reveal that bare stems in each file outnumber [-T] clauses. While this is not incompatible with ATOM, it strongly suggests that Agr omission and T omission are not entirely independent, and that there may be other factors that contribute to the omission of inflectional prefixes in Swahili.

The final issue relates to Mood. Deen & Hyams (2002) investigate the occurrence of subjunctive and indicative verbs and their interpretation. Their overall finding is that indicative verbs from early on occur in prototypically realis contexts while subjunctive verbs occur in prototypically irrealis contexts. However, they also found that in stages 1 and 2, Mood is not as productive as it is in stages 3 and 4. Not only are the overall frequencies lower in the earlier stages, but the type-token ratios are lower, and the proportion of subjunctive-to-indicative verbs is lower.

Table 4.12 Types/tokens of verbs expressing irrealis mood

	Stage 1	Stage 2	Stage 3	Stage 4
Type	8	5	26	20
Tokens	9	7	50	37

Table adapted from Deen & Hyams (2002)

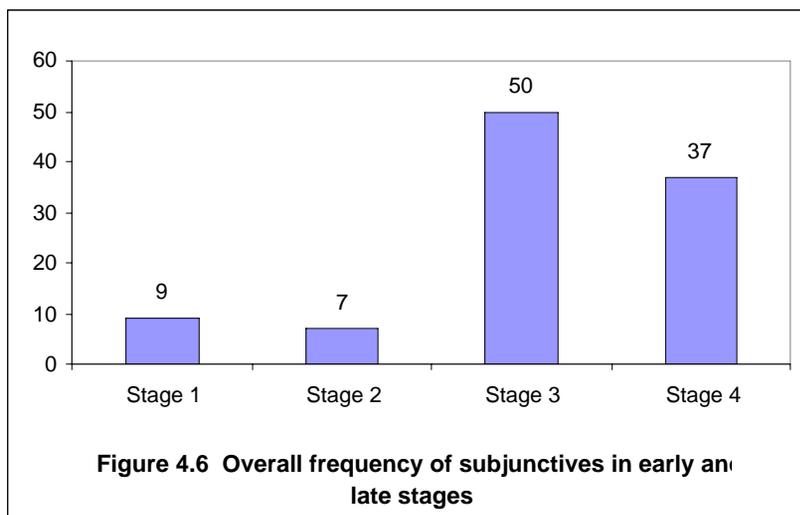


Table 4.13 Proportion of subjunctives to indicative verbs in early stages and late stages.

	Stages 1+2	Stages 3+4
Subjunctives	16	87
Indicatives	500	826
Ratio	0.03	0.10

Table 4.13 and figure 4.6 show the raw frequency of subjunctives in each stage. Table 4.13 then shows the ratios of subjunctives-to-indicative verbs in the early stages and the late stages. We see that in the early stages, subjunctive verbs occur at a rate of 0.03. This means that for every 100 verbs in these stages, 3 are subjunctive. In stages 3 and 4 the ratio jumps to 0.10, which means that proportionately there are three times as many subjunctive verbs in the later stages than the early stages. Deen & Hyams suggest that stage 3 is when Mood becomes productive.

Alternation in the final mood vowel is crucial in showing productivity in Mood in early grammar. The paucity of subjunctives in the

early stages, therefore, suggests that Mood too may be underspecified. Notice also that the stage in which Mood becomes productive (stage 3) co-occurs with the stage in which Tense becomes obligatory in the child grammar (see figure 4.4 in which we saw that T omission diminishes to around 10% by stage 3)²⁵.

Agr and T may be underspecified, and this results in omission of SA and T. If Mood is underspecified, then how then do we explain the earlier finding that the mood final vowel is virtually always present? There is good reason to believe that the phonotactic system of a particular language is acquired very early by children, perhaps before the first word is even spoken (e.g., Jusczyk *et al*, 1993). We saw in chapter 2 that Swahili generally does not allow coda consonants, but requires that every word end in a vowel. This is typically visible in loan words that take on a thematic final vowel, e.g., *lazima* (from the Arabic ‘ladhim’), *wiki* (from the English ‘Week’), etc. Children undoubtedly are attuned to this phonotactic requirement. Let us assume that Mood is underspecified. This would result in the child omitting the mood final vowel, just as SA is omitted through all four stages and T is omitted primarily through stages 1 and 2. Thus a target utterance such as (44a) with Mood underspecified would have an intermediate representation as in (44b). However, knowledge of the phonotactic requirement that coda consonants are ungrammatical would prevent (44b) from surfacing, and the form in (44c) would surface.

- (44) a. ni – ta–anguk–a
 SA_{1s}–fut–fall–IND
 ‘I will fall down.’

²⁵ See Deen & Hyams (2002) for some speculations about the connection between T and Mood.

- b. ni – ta – anguk
SA_{1s}–fut–fall–∅
- c. ni – ta – anguk – a
SA_{1s}–fut–fall–thematic vowel

However, as Carson Schütze points out, this predicts that the final vowel on the verb should be subject to the same principles that govern epenthetic vowels in loan words. The final epenthetic vowel in loan words is related to the preceding vowel in the stem, i.e., when the vowel in the stem is [high, front], the epenthetic vowel is harmonized as [high, front]. When the preceding vowel is [low, back], the epenthetic vowel is [low, back]. Thus the final vowel on the verb should alternate according to verb stems. This is clearly not the case, with the majority of verbs being marked with the indicative *a-*, including verbs that have [high, front] vowels in the stem, e.g., *ingi-a* ‘enter’, *pig-a* ‘hit’, etc. Thus the naturalistic data on this point are not conclusive, and this remains a point for future research.

4.7 Conclusion

In this chapter we investigated the acquisition of Swahili verbal inflection. We saw that overall Mood is rarely omitted, although we suggested that this may be due to phonotactic requirements. We saw that OA too is rarely omitted. We found that Swahili children omit SA and T in significant proportions at early stages, and slowly develop towards the adult norm. Swahili children independently omit both Agr as well as T. I showed that neither a metrical theory of omission nor a Truncation theory accounts for this. I also showed that underspecification of a single functional head is inadequate to account for all the data, and thus the

Swahili data lend support to Schütze & Wexler’s ATOM which states that Agr and T may be independently and optionally underspecified.

We will see in the next chapter that these inflectional omissions are not simply morphological or phonological omissions. While Swahili cannot be examined in the way that other European languages have been examined (because Swahili does not have wh- movement, null subjects are allowed, verb raising does not occur over negation or adverbs, there is no V2 effect, etc.), I will show that there are syntactic effects of this underspecification. I will show that different sorts of subject omission correlate with the different clause types, suggesting that the omission of inflectional prefixes is syntactic in nature.