

1.1 The study

The focus of this study is the acquisition of verbal morphology in Swahili, an Eastern Bantu language with heavy agglutinative prefixing and suffixing. This is the first investigation of the acquisition of Swahili, and indeed the first study of the acquisition of any eastern Bantu language. While a handful of other Bantu languages have been studied, those results differ significantly from what is reported here. Similarly, relatively few agglutinative languages have been studied, and so this study of Swahili contributes to this neglected area of research. Finally, most of what we know about the acquisition of affixes is about the acquisition of suffixes. We have a general idea that the acquisition of prefixes is problematic for children (e.g., Slobin, 1973), but very little empirical data is available. This study aims to contribute to our knowledge of the acquisition of prefixes.

In this chapter I will first discuss the fact that children invariably converge on the morphosyntactic system of their language early in the developmental sequence. Hoekstra & Hyams (1998) discuss this as Early Morphosyntactic Convergence (EMC), and Wexler (1998) describes this as Very Early Knowledge of Inflection (VEKI). I will suggest that child language falls into three classes based on how verbal inflection is used: Root Infinitive (RI) languages (such as German and Dutch, see below for a description of what an RI is), non-Root Infinitive languages (such as Italian

and Spanish), and bare verb languages (such as English). This is summarized in table 1.1 below¹:

Table 1.1 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	
French	Catalan	
	Japanese	

Hoekstra & Hyams (1998) show that there are morphosyntactic and semantic differences between RIs and bare verbs. The only bare verb language that they consider is English, but I will show that bare verb languages are considerably more frequent than commonly thought. I will outline some previous findings from other Bantu languages, focusing on the acquisition of verbal morphology. I will show that Bantu languages fall into the category of bare verb languages. I will then discuss the acquisition of two agglutinative languages, and we will see that the omission of verbal morphology is pervasive in these languages too. This typology will serve as background for our investigation of the acquisition of Swahili.

We will see that Swahili children allow bare verbs like the children of other Bantu languages. However, because Tense in Swahili is obligatory (in contrast to languages such as Sesotho and Siswati), it allows us to investigate the development of several inflectional affixes at the same time. This in fact will be the most revealing aspect of this study – that subject

agreement and tense develop differentially and have different effects on other aspects of grammar, e.g., the occurrence of overt subjects. We will see that Swahili children allow bare verbs, but also allow various other sorts of underspecified verbs. I show that these underspecifications are accounted for by a model of underspecification called ATOM (Agr-Tense Omission Model, Schütze & Wexler, 1996).

Let us begin with an overview of the types of errors children make. I will first show that errors of omission are very frequent and typical of child language. One product of such errors is the bare verb in English. I will then show that errors of commission are relatively rare in child language, with the apparent exception of Root Infinitives in languages such as German². RIs have received so much attention in the literature, and so I will describe their characteristics in section 1.2.3. With the addition of languages that allow neither bare verbs nor root infinitives (e.g., Italian), this will establish our typology in table 1.1 above.

1.2 Early Morphosyntactic Convergence

Children don't begin talking until around the first birthday, after which it takes them several years to fully acquire the grammar of their ambient language. While it may take several years for them to gain mastery of their language, during that time children's grammar does not develop at random. Children make systematic errors and progress through systematic

¹ In addition there are languages that allow bare participles, default finite forms, etc. I ignore these here for the sake of simplicity.

² I say *apparent*, because I take the position (as many in the field) that RIs and bare verbs arise from the same underlying mechanism, namely the underspecification of functional heads. See chapter 4 for details.

stages in the acquisition of language, culminating in mastery of the language they are born into. One such stage in language acquisition that begins with multi-word utterances and ends sometime around the 3rd birthday is the so-called Telegraphic Stage (Brown, 1973). Telegraphic speech is characterized by the omission of obligatory elements such as determiners, prepositions, agreement markers, etc.

1.2.1 The Frequency of Omission

Children acquiring various languages omit determiners (1), copulas (2), auxiliaries (3), subject-verb agreement (4, taken from CHILDES, MacWhinney (2000), etc.:

- | | | |
|-----|---|--|
| (1) | Determiner Omission | |
| a. | Paula play ball | Paula, 1;6 (Radford, 1990) |
| b. | Haley draw boat | Hayley, 1;8 (Radford, 1990) |
| c. | want open door | Daniel, 1;8 (Radford, 1990) |
| d. | Niekje ook boot maken
Niekje also boat make
'Niekje has also made a boat' | Dutch (Schaeffer, 1994) |
| e. | Papa heeft ook trein
Daddy has also train
'Daddy also has a train' | Dutch (Schaeffer, 1994) |
| f. | est tombé éfant
is fallen elephant
'The elephant has fallen | French (Ferdinand, 1996) |
| g. | train va tomb[e]
train go fall
'The train is going to fall.' | French (Ferdinand, 1996) |
| (2) | Copula Omission | |
| a. | I in the kitchen | English (Becker, 2000) |
| b. | Da rote ball
there red ball | German (Salustri & Berger-Morales, 2001) |

- (3) Auxiliary Omission
- a. baby talking Hayley, 1;8 (Radford, 1990)
 - b. doggy barking Bethan, 1;9 (Radford, 1990)
 - c. Mummy doing dinner Daniel, 1;10 (Radford, 1990)
- (4) Subject-Verb Agreement Omission
- a. It only write on the pad Eve, 2;0 (Brown, 1973)
 - b. Cromer have some Adam, 2;7 (Brown, 1973)
 - c. He bite me Sarah, 2;9 (Brown, 1973)

Note that the examples in (4) exemplify the fact that English children allow bare verbs in contexts in which an agreement marker is obligatory (3rd person singular –s). Not only is omission typical, it is very frequent overall. Table 1.2 below shows that the omission of 3rd person singular agreement in English occurs at rates of approximately 80%:

Table 1.2 Proportion of bare verbs in child English

Child	Age	% Bare Verbs
Eve	1;6-1;10	78
Adam	2;3-3;0	81
Nina	2;4-2;5	75

(Sano & Hyams, 1994)

We will see shortly that in a variety of agglutinative languages too, omission is extremely frequent.

1.2.2 The Paucity of Agreement Errors

While omission is frequent and widespread, errors of agreement are extremely rare. Table 1.3 (from Sano & Hyams, 1994) shows that children in a variety of languages generally make errors of agreement less than 4% of the time.

Table 1.3 Agreement errors in various languages

Child	Language	Age	n	%error	Source
Simone	German	1;7-2;8	1732	1%	Clahsen & Penke 1992
Martina*	Italian	1;8-2;7	478	1.6%	Guasti 1994
Diana*	Italian	1;10-2;6	610	1.5%	Guasti 1994
Guglielmo*	Italian	2;2-2;7	201	3.3%	Guasti 1994
Claudia	Italian	1;4-2;4	1410	3%	Pizzuto & Caselli 1992
Francesco	Italian	1;5-2;10	1264	2%	Pizzuto & Caselli 1992
Marco	Italian	1;5-3;0	415	4%	Pizzuto & Caselli 1992
Marti*	Cat/Spanish	1;9-2;5	178	0.56%	Torrens 1992
Josep*	Cat/Spanish	1;9-2;6	136	3%	Torrens 1992
Gisela*	Catalan	1;10-2;6	81	1.2%	Torrens 1992
Guillem*	Catalan	1;9-2;6	129	2.3%	Torrens 1992

(*data available on CHILDES, MacWhinney, 2000: Martin, Guglielmo, Diana corpora, Cipriani *et al.* 1991; Marti, Josep, Guillem, Gisela corpora: Serra & Solé, 1992)

Take English for example, where children hardly ever use 3rd person singular agreement morphology in 1st person singular contexts, as in the hypothetical example given below:

- (5) I eats cake UNATTESTED

Table 1.4 below shows that from the data of 10 children, Harris & Wexler (1996) identified 1724 verbs that occur in 1st person singular context, of which only 3 occur with the incorrect 3rd person singular –s suffix.

Table 1.4 Frequency of verbs in first person singular contexts

Stem	Irregular Past	-ed	-s
1349	325	47	3

Data from 10 children on CHILDES, age range = 1;6-4;1 (Harris & Wexler, 1996)

Similarly, Schieffelin (1985) shows that children acquiring Kaluli ‘very rarely put the wrong case marking on nouns’ (p.537). Smoczy_ska (1985) reports the same for Polish speaking children, and Clancy (1985) says that ‘the system of verbal inflections in Japanese emerges quite early, and errors are not frequently reported’ (p.383). Thus such errors of commission are extremely rare, while errors of omission are common.

This general paucity of such errors has led researchers to conclude that children acquire knowledge of the morpho-syntactic properties of their language extremely early, hence Hoekstra & Hyams' (1998) EMC. In the next section we will see a phenomenon that at first glance is an apparent exception to EMC, but which in fact conforms to EMC.

1.2.3 Root Infinitives

The evidence for EMC is that when inflection occurs, it occurs overwhelmingly correctly. However, the occurrence of Root Infinitives casts doubt on the validity of this generalization. A Root Infinitive (RI) is a verb that occurs with infinitival morphology in a matrix clause (Hoekstra & Hyams, 1998). Thus, instead of having finite inflectional morphology in a main clause, RIs have 'incorrect' infinitival morphology, as in the examples in (6)³. In example (6a), the child uses the infinitive form of the verb *haben* 'to have'. The appropriate finite form of this verb is *hat*. This is not an error of *omission*, but appears to be an error of *commission*⁴. Thus, this may represent an exception to EMC.

³ This view that RIs are verbs with infinitival morphology is made explicit by Hoekstra & Hyams (1998). However, others propose that RIs occur as default forms when the verb is underspecified for Tense (Wexler, 1994), Agr (Clahsen, Eisenbeiss & Penke, 1996), or both Agr and Tense (Schütze, 1997). See chapter 4 for a discussion of these proposals.

⁴ Carson Schütze points out that 3rd singular subject is not incompatible with the infinitive, and thus is not an error of commission in that sense. However, it is an error of commission in the sense that there is a mismatch between the morphology (non-finite) and the intended meaning (finite). The English bare form does not fit this criterion of being an error of commission because while the form does not match the intended meaning, there is a total absence of morphology, thus an error of omission.

- | | | |
|---|---|--------|
| (6)a. Thorstn das haben
Thorstn that have-inf
'Thorstn has that.' | b. Zahne putzen
teeth brush-inf.
'(Someone) brushes (his)
teeth.' | German |
| c. Papa schoen wassen
daddy shoes wash-inf.
'Daddy washes (the) shoes.' | d. Ik ook lezen
I also read-inf.
'I also read.' | Dutch |
| e. Fermer yeux
close-inf. Eyes
'(I have) closed (my) eyes' | f. 'Tasha ouvrir
Natasha open-inf.
'Natasha is opening/ going to
open/wants to open it.' | French |

RIs occur in languages such as German (Weissenborn, 1990; Poeppel & Wexler, 1993), Dutch (Weverink, 1989), French (Pierce, 1989), Swedish (Platzack, 1990), etc.

Table 1.5 Frequencies of RIs in French and Swedish

Language	Child	Age	%RIs
French	Nathalie	1;7-2;11	76%
French	Daniel	1;5-2;5	60%
Swedish	Freja	1;11-2;0	38%
Swedish	Tor	1;11-2;2	56%
Swedish	Embla	1;8-1;10	61%

(Adapted from Sano & Hyams, 1994)

It is important for us to discuss the various characteristics of RIs because they are such a pervasive phenomenon in several languages in Western Europe. Because RIs are attested in some child languages, a priori we expect RIs to be a possibility for Swahili children too. RIs have become an important focus of study in the acquisition of European languages, precisely because they represent an apparent departure from the EMC. However, upon closer inspection, it becomes clear that they are not an exception to the EMC. RIs have several defining characteristics. First, RIs

are optional: RIs occur side-by-side with fully inflected verbs, sometimes in immediately adjacent utterances. Examples (7-8) are taken from a single corpus of a German speaking child, Andreas (cited in Wexler, 1994)⁵.

- (7) [+finite] verbs
- a. Mein Hubsabe hat Tiere din
my helicopter has animals in it
'My helicopter has animals in it'
- b. Caesar tieg e nich ('Caesar kriegt er nicht')
Caesar get he not
'He is not getting Caesar.'
- (8) [-finite] verbs
- a. ich der Fos hab'n
I the frog have (-fin)
'I have the frog.'
- b. Thorstn das haben
Thorstn that have (-fin)
'Thrstn has that.'
- c. Zahne pussen
teeth brush (-fin)
'I (want) to brush my teeth.'
- d. tein ('kein') Zahnburste liegen
no toothbrush lie (-fin)

Behrens (1993) investigates temporal marking in the speech of Simone (Miller, 1976), a child acquiring German, and reports that RIs occur

⁵ The translations here are those provided by the authors. They represent the authors 'hypothesis as to what the child intended. We will see shortly that others have proposed that RIs generally have modal meanings associated with them. The translations given in the text would thus change so as to incorporate a modal meaning.

alongside finite verbs in the same transcripts. Her figures are given in table 1.6.

Table 1.6 RIs and finite verbs in the files of Simone

Age	RIs	Finite
2;0	72%	28%
2;6	32%	68%
3;0	16%	84%
4;0	4%	96%

(Behrens, 1993)

Second, it is well known that RIs occur in the same position as infinitival verbs in the adult language, namely in an *unraised* position. In child German, RIs occur overwhelmingly in final position, while finite verbs in child German occur overwhelmingly in second position (examples and table taken from Poeppel & Wexler, 1993):

- (9) a. Thorsten Caesar haben unraised RI (SOV)
Thorsten Caesar have-inf
'Thorsten has (the doll) Caesar'
- b. Ich hab tein Bürse raised finite verb (SVO)
I have (a) small brush

Table 1.7 Finiteness versus verb position in German

	+finite	-finite
V2	197	6
V final	11	37

(Table adapted from Poeppel & Wexler, 1993, p.7. Data is from Andreas, age 2;1, taken from the CHILDES system, MacWhinney, 2000)

Similarly, adult French has verb raising over negation and adverbs when the verb is finite, but not when the verb is infinitival. Pierce (1989) reports that French children produce RIs in the unraised position (post negation), and inflected finite verbs in raised position.

- (10) a. pas casser unraised RI (V-neg)
not break
- b. marche pas raised finite verb (neg-V)
walks not

Table 1.8 Finiteness versus verb position in French

	+finite	-finite
Verb <i>pas</i>	216	2
<i>pas</i> Verb	9	122

(Table is from Pierce, 1989; data are from four French speaking children aged 1;8 – 2;2)

This correlation between verb raising and RIs strongly refutes the view that RIs are speech errors. Because child RIs behave syntactically like infinitives in the adult language, this means that RIs are syntactic infinitives that are permitted in root clauses and they give rise to certain interpretations. Adult languages too allow RIs with particular interpretations (see Schütze, 1997; Hoekstra & Hyams, 1998 for details), but not in declarative contexts that we see RIs in child language. More importantly, however, the fact that RIs behave like adult infinitives shows that they are not exceptions to the EMC. If children have not converged on the morphological properties of infinitives (and the associated syntactic restrictions on that morphology), then they should not know the position requirements of infinitives. As the two tables above show, children clearly have knowledge that RIs are infinitives.

The third characteristic of RIs is that they generally occur with null subjects (Weverink, 1989; Pierce, 1989; Phillips, 1995).

Table 1.9 Null subjects in finite and non-finite clauses in several child languages

		Null Subjects in Finite Clauses	Null subjects in non-Finite clauses
French (Pierce, 1989)	Daniel	150/273 (55%)	166/205 (81%)
	Nathalie	90/304 (30%)	131/295 (44%)
	Philippe	182/782 (23%)	153/194 (79%)
German (Behrens, 1993; Krämer, 1993)	Simone	781/3699 (21%)	2199/2477 (89%)
	Andreas	34/263 (13%)	69/101 (68%)
Dutch (Krämer, 1993; Haegeman, 1995)	Thomas	165/596 (28%)	246/267 (92%)
	Heinz	1199/3768 (32%)	615/721 (85%)
Flemish (Krämer, 1993)	Maarten	23/92 (25%)	89/100 (89%)
Faroese (Jonas, 1995)	O.	8/52 (15%)	67/161 (42%)
Danish (Hamann & Plunkett, 1998)	Anne	366/3379 (11%)	394/667 (59%)
	Jens	742/3173 (23%)	539/937 (58%)

(Adapted from Rasetti, 2000, table 9.1, p.239).

In table 1.9 we see that the proportion of null subjects in non-finite clauses is consistently higher than the proportion of null subjects in finite clauses in a range of languages. This is reminiscent of the correlation between adult non-finite embedded clauses and null PRO subjects, and thus it has been suggested that child RI clauses are like adult infinitival clauses (e.g., Guilfoyle, 1984; Sano & Hyams, 1994; Radford, 1990; Hyams, 1996). See chapter 5 for more discussion of null subjects and their relation to underspecified clauses. Because child RIs have this property of adult infinitives, this is further evidence that RIs are not an exception to the EMC.

The fourth characteristic of RIs is that they generally occur in modal contexts (for Dutch, see Wijnen, 1996; for French see Ferdinand, 1996; for Swedish see Plunkett & Strömquist, 1990; for a review see

Hoekstra & Hyams, 1998). Examples are given below (from Wijnen, 1996):

- (11) a. Papa boek lezen
Daddy book read-inf
'(I want)Daddy to read the book'
- b. Niekje buiten spelen
Niekje outside play-inf
'Niekje wants to play outside'

Hoekstra & Hyams (1998) argue that this modality comes from the presence of infinitive morphology, which cross-linguistically is generally associated with modal meaning (Duffley, 1992; Bolinger, 1968; Stowell, 1981; Han, 2000). The fact that child RIs are associated with modal/irrealis contexts is further evidence that the infinitival morphology on RIs is not 'incorrect' use of morphology, but rather that RIs are genuine syntactic infinitives, and thus do not represent an exception to the EMC.⁶

⁶ The fifth characteristic of RIs is that they generally do not occur in wh-contexts. While they are very frequent in declarative clauses, it has been shown for several languages that RIs are very rare in wh-contexts (e.g., Kursawe (1994) for German; Haegeman (1994) for Dutch). The table below shows that in the corpus of Hein (Elbers & Wijnen, 1992), of the 721 non-finite verbal clauses, only 2 occur with wh-questions (0.2%). However, wh-questions occur at a rate of 2.3% in finite contexts. This does not directly refute the view that RIs are exceptions to EMC, but is an important characteristic that we will return to in chapter 5.

Finiteness in declaratives versus wh- questions in child Dutch

	+finite	-finite
All Clauses	3768	721
Wh- questions	88	2

Table taken from Haegeman (1994); data is from Hein (Elbers & Wijnen, 1992)

However, not all languages allow RIs. Sano & Hyams (1994) show that RIs are very frequent in languages such as German, Dutch, French, Swedish, etc., but are somewhat rarer in languages such as Italian, Spanish, Catalan, etc. (see table 1.10 below). Children acquiring these languages produce neither RIs nor bare verbs (see for example Guasti, 1992). Thus these child languages are classified as a third group, distinct from RI languages and bare verb languages.

Table 1.10 Proportion of RIs in some non-RI languages

Language	Child	Age	%RI
Italian (Guasti, 1992)	Diana	2;0	0%
	Martina	1;11	16%
		2;1	4%
Italian (Schaeffer, 1990)	Paola	2;0-2;5	7%
	Daniele	1;7-2;6	8%
	Massimo	1;7-2;6	6%
	Gabriele	1;7-2;6	7%
	Orietta	1;7-2;6	5%
	Elisabeth	1;7-2;5	10%
Spanish (Grinstead, 1994)	Damariz	2;6-2;8	5%
	Juan	1;7-2;0	12%
		2;1-2;4	10%
Catalan (Torrens, 1992)	Guillem	1;11-2;6	3%
	Marti	2;0-2;5	3%

(Adapted from Sano & Hyams, 1994)

Furthermore, there is the issue of whether English bare verbs are RIs, or whether they should be categorized separately. Wexler (1994) suggests that English bare verbs are the result of the same processes that produce German and Dutch RIs, namely the underspecification of tense. Wexler argues that one appealing reason to assimilate them to RIs is that it

would unify English (which at that time was the only bare verb language that was well known) with other Germanic languages. However, we will see that bare verbs occur frequently in the languages of the world, and thus categorizing English bare verbs as RIs obscures this distinction. While bare verbs and RIs may be the result of the same underlying mechanism, it is unclear why in some languages bare verbs are attested while in other languages RIs are attested (and still in other languages neither are attested).

I will not provide an answer to this question, but there are a number of possibilities that come to mind that may be worth pursuing in the future. One possibility is the morphological differences between Bare verb languages and RI languages. Specifically, it has been noted that infinitives in languages such as German and Dutch behave differently and are of a different morphological status than the infinitive in English. For example, in the adult RI languages the infinitive occurs as a postverbal affix while in English it occurs preverbally. Furthermore, as Schütze (1997) points out, the infinitive in RI languages occurs as a bound morpheme to the verb, with no intervening lexical material allowed. However, in English the infinitive marker has a somewhat looser relationship with the verb, since intervening lexical material is allowed, for example “To boldly go where no one has gone before”.

However, this does not answer the question of why languages such as Italian and Spanish (which have bound postverbal infinitives) do not exhibit RIs. Thus the reason for this three-way typology remains a mystery. What I will show in the next section is that Bantu languages disallow RIs and allow bare verbs, and thus are categorized as Bare Verb Languages. The infinitive in Bantu languages resembles the English infinitive in that it is preverbal and allows intervening material, suggesting that the preverbal

position and free nature of the infinitive perhaps is related to the occurrence of bare verbs.

1.3 Acquisition of Bantu Languages

In this section I will briefly discuss some previous acquisition findings from two other Bantu languages: Sesotho and Siswati. The only acquisition studies reported in the literature are of southern Bantu languages, such as the pioneering work on Sesotho by Katherine Demuth. The majority of what is discussed in this section comes from work done by Demuth and her colleagues. For an excellent review of the acquisition of Sesotho and the Bantu acquisition literature, see Demuth (1992b).

1.3.1 Sesotho

Sesotho is a southern Bantu language spoken in the nation of Lesotho. Like Swahili it has rich noun class morphology, but unlike Swahili, Sesotho is a tonal language. Sesotho also has a rich verbal complex, with the word order given in (12) below:

(12) Subject SA–(T/A)–(OA)–Verb–suffixes/perf. /pass.–Mood Object

The verbal complex consists of a subject agreement marker, followed by an optional tense/aspect marker, an optional object agreement marker, various suffixes (including grammatical function changing suffixes), an optional aspect marker, and a mood final vowel. An example is given below (taken from Demuth, 1992)⁷:

⁷ I follow Demuth’s transcription protocol, where ´ indicates high tone, and low tone is left unmarked.

- (13) Thabo ó – pheh – íl – é di–jó
 Thabo SA₁–cook–PERF–MOOD 8-food
 ‘Thabo cooked (some/the) food.’

Sesotho is a tonal language, making use of two tones (Demuth, 1993): high, and Ø (which surfaces as a default low). Unlike Swahili, the only prefix that must occur in every declarative utterance is the subject agreement prefix. As we will see in chapter 2, Swahili requires subject agreement as well as tense in almost every declarative context.

The infinitive in Sesotho (and indeed in Swahili as well) is preverbal and may occur with object agreement intervening between it and the verb. Thus morphologically it resembles the English infinitive in these ways. However, lexical material may not intervene between the infinitive prefix and the verb, suggesting a somewhat tighter relationship than in English.

The inflectional prefixes are the focus of the current study, and so I will restrict the discussion to the acquisition of verbal prefixes in Sesotho. Demuth reports that children acquiring Sesotho go through three non-discrete stages. In the first stage, children predominantly produce bare verbs, that is, verbs that are missing all inflectional prefixes. The second stage is typified by verbs which have ‘shadow vowel’ prefixes. Such prefixes are typically [a] or [e] in Sesotho. This phenomenon has been noted in other Bantu languages as well, notably Siswati (Kunene, 1979, see below). (14) is an example of an utterance with a shadow vowel from Sesotho. The first line is the child utterance, followed by the adult equivalent, a gloss and a translation.

- (14) a lahlíle
 ke – di – láhl – íl – e+
 SA_{1s} –OA₁₀–throw away–PERF–MOOD
 ‘I threw them away.’

In fact, such preverbal syllables are not uncommon in the acquisition of other languages. In the literature, these preverbal vowels have been variously labeled: Peters (2001) refers to them as filler syllables; Bottari, Cipriani & Chilosi (1993) refer to them as Monosyllabic Place Holders; Veneziano & Sinclair (2000) call them ‘additional elements’. We will return to this phenomenon in chapter 3 when we investigate child Swahili.

In the third stage, fully inflected forms predominate. Thus the developmental sequence is given in (15).

- (15) no inflection > ‘shadow vowels’ > well-formed inflection

Demuth notes that these stages are not discrete, and that utterances of all three kinds occur in all three stages. The examples below (Katherine Demuth, p.c.) all come from the same child (L, age 2;1), and are attempts at the same utterance. Examples (16a-d) are the child utterances, and example (17) is the adult target.

- (16) a. qetile
 b. ketile
 c. eketile
 d. aketile

- (17) ke – qet – il – e
 SA–finish–prf–IND
 ‘I finished’

(16a,b) are examples of bare verb stems (i.e., missing prefixes, although they do contain the aspectual suffix), while (16c,d) are examples of utterances with shadow vowels. At the same time, this child also has well-formed subject agreement markers in other utterances:

- (18) ke – i – thol – ets – e tsena
 SA–Rflex–find–APPL/PRF–IND these
 ‘I found these’

However, importantly, while bare verbs do occur in Sesotho, RIs do not. Thus we can classify Sesotho in our typology of child languages as a Bare Verb language, since there is a stage at which children allow bare verbs. These bare verbs occur alongside fully inflected verbs, and are thus typical errors of omission. Demuth reports that bare verbs are significantly less frequent than bare nouns in her corpora, but still do occur. She predicts that bare verbs are more common with younger children.⁸

1.3.2 Siswati

Kunene (1979) investigates the acquisition of noun class morphology, possessive morphology and agreement in Siswati, a southern Bantu language closely related to Sesotho. Her data comes from two children aged 2;2-3;0 and 2;11-3;6. She finds that Siswati children at early

⁸ Besides the acquisition of prefixes, Demuth reports on the acquisition of other verbal elements. For example, Demuth (1989) reports that passives occur relatively frequently in Sesotho speaking children's everyday speech. She finds that, contrary to English speaking children (cf. Borer & Wexler, 1987; Fox & Grodzinsky, 1998), Sesotho children show evidence of spontaneous production of passives as early as 2;8, as well as comprehension of passives. This is different from children acquiring Western European languages, who acquire the passive relatively late. Demuth attributes this early proficiency to the unusual frequency of passives in the input. Sesotho care-givers use an unusually large proportion of passives when speaking to children (in Demuth 1992b the proportion of passives in the care-givers utterances is reported as 23/386 (6%)). Demuth also reports on the acquisition of grammatical function changing suffixes such as the applicative (Demuth, 1998) and the tonal system of Sesotho (1993), as well as nominal morphology (Demuth, 1988, 1992, 1994). However, since the focus of the present study is the verbal complex in Swahili, I will not discuss these results. For an overview of these results see Demuth (1992b).

stages produce bare verbs exclusively. Adult Siswati requires minimally that the verb occur with a subject agreement prefix (tense and object agreement are not obligatory). The youngest child at her first data point produced only bare verbs with no inflectional prefixes whatsoever.⁹ Examples are given below.¹⁰

- (19) a. Zanele lala
 adult form: Zanele u-ya-lala
 Zanele SA-ya-sleep
 'Zanele is sleeping.'
- b. landzela mine
 adult form: ngi - ta -ku -landzela mine
 SA-future-OA-follow me
 'I will follow you.'
- c. tfwana khala
 adult form: um-tfwana u - ya - khala
 pref-child SA-ya-cry
 'The baby is crying.'

These examples are taken from the first recording of the youngest child. In the second recording, when the youngest child was 2;3, she began to produce verbs that had a "rudimentary" subject agreement marker. This "rudimentary" prefix is similar to what Demuth describes in the

⁹ Kunene does not provide quantitative data, so we do not know how many verbal utterances the child produced at this stage.

¹⁰ In two of these examples, there is a prefix 'ya' in the adult form, which Kunene describes as occurring 'when there is no adjunct following the verb.' The exact function of this prefix is unclear, since in the examples she glosses this prefix as 'ya'. I assume it is a marker of intransitivity.

development of Sesotho as a “shadow vowel”. In Siswati, these prefixes occur primarily as [i] or [a], as examples below show:

- (20) a. i/a – lhala la mlumbi
 Adult form: u – lhala la um – lumbi
 SA – sit here pref. – whiteman
 ‘The whiteman sits here.’
- b. a – buka tfombe t – a – mi
 ngi–buka ti–tfombe t – a – mi
 SA – look pref.–picture PA–PM–my¹¹
 ‘I am looking at my pictures.’

In example (20a), ‘i/a’ indicates that the child used both prefixes in free variation. In (20b), only [a] occurred as the “rudimentary” prefix.

Kunene describes the development of prefixes on the verb in Siswati as being roughly parallel to what Demuth describes for Sesotho, i.e., beginning with no inflection, then developing into a shadow/rudimentary vowel, followed by full well-formedness. We will see that early Swahili is similar to Sesotho/Siswati in that it allows ‘shadow vowels’ as prefixes at a particular stage in development. These ‘shadow vowels’ (in Swahili as well as Sesotho and Siswati) are not well-formed in that they do not conform phonologically to any target prefix. The same holds true in Swahili.

Unlike Sesotho and Siswati, Swahili obligatorily takes subject agreement and tense prefixes. Thus, the use of a single (‘shadow’) preverbal vowel (21a) can be contrasted with the occurrence of both well-formed target prefixes (21b), the occurrence of a single well-formed prefix (21c-d), and the occurrence of no well-formed prefixes (21e).

¹¹ PA = possessive agreement marker, PM = possessive morpheme

- (21) Possible prefixes in Swahili:
- a. preverbal shadow vowel
 - b. Both SA and T prefixes
 - c. Only SA prefix
 - d. Only T prefix
 - e. Neither SA nor T prefixes

This is of relevance in the current study because we will see that Swahili children in fact produce each of these possibilities.

The Bantu languages that have been studied to date are all southern Bantu languages. I have only reported on Sesotho and Siswati, but other languages that have been studied include Zulu (Suzman, 1982; 1991), Basotho (Connelly, 1984; 1987), and Chichewa (Chimombo, 1981).¹² These languages differ from Swahili in that tense is not obligatory: in Swahili, no declarative clause may occur without a tense prefix. The fact that Subject Agreement and Tense are marked as independent prefixes in Swahili in almost every declarative clause sets it apart from the other Bantu languages mentioned earlier in this chapter. The acquisition of these two prefixes will be the primary theme of this dissertation. I turn now to a brief look at the acquisition of other agglutinative languages.

1.4 Other Agglutinative Languages

I will focus on the acquisition of verbal affixes, and we will see that children generally allow omission of these affixes. We will see that children identify the bare verb stem at early ages, and morphology is added

¹² Chichewa is not from the same group as Sesotho or Siswati. However, the study cited focuses on negation, and the differences between first and second language acquisition. Other work by Chimombo has focused on tone, and thus are not of direct relevance to Swahili.

to that stem as children develop. This is the case in Sesotho and Siswati, and we will see this in Quechua and Navajo. This will establish a pattern in child language that allows bare verb stems.

1.4.1 Quechua

Courtney (1998) investigates the acquisition of Quechua in the spontaneous speech of four children aged 2;0-2;8. Quechua is an SOV agglutinative language with a verbal complex that has extensive suffixing. In adult Quechua, the order of morphemes is given in (22)¹³:
 (22) V – Derivational – OA – Prog – OA – T – SA – Num – Conditional
 The derivational suffixes include transformative, factitive, desiderative, and perdurative. OA occurs either after the derivational suffixes or after the progressive marker, and the SA marker is compositional between person and number marking. Furthermore, “a Quechua verb must bear either a person-of-subject suffix appended to the root, or the Imperative/Infinitive morpheme. That is to say, adult Quechua speakers never produce bare roots or stems.” (Courtney, 1998, p.60). For an imperative, the infinitive/imperative suffix is appended at the end. Thus the infinitive occurs as a suffix in Quechua (unlike either English or Bantu languages). As can be seen from (22) several suffixes may intervene between the verb and the infinitive suffix.

In Quechua child-directed speech, Courtney reports that of 279 identifiable verbs, 249 occur with SA as well as at least one additional suffix (p.168). Therefore almost 90% of all verbs that the children hear

¹³ Derivational = derivational suffix; OA=Object Agreement; Prog = progressive aspect ; T = Tense ; SA = Subject Agreement ; Num = number marking

occur with at least two suffixes. The remaining 30 verbs occur with SA. Importantly, no bare verbs occur in the input.

Children, on the other hand, allow bare verbs quite frequently. The youngest child (2;0-2;2) produced 81 verbs (with 48 different verb roots), of which 46 were bare.¹⁴ The second child (2;5-2;7) produced 37% bare verbs in the first half of his corpus, but in the second half produced none whatsoever. The third child (2;7-2;8) also produced no bare verbs. Thus bare verbs are a symptom of very early child language (as Demuth surmises about child Sesotho). Examples are given in (23). In each case, the child produces the bare verb (*muna*) when the adult form would have taken at least one suffix.¹⁵

- | | | |
|------|-------------|---|
| (23) | a. | Chay muna qan |
| | adult form: | chayta munanki qan
'you want that' |
| | b. | Carruta muna noqa |
| | adult form: | noqa carruta munani
'I want the car' |
| | c. | noqa carru muna |
| | adult form: | noqa carruta munani
'I want the car' |
| | d. | mana muna noqa carrupi |
| | adult form: | noqa carrupi mana munanichu
'I don't want the car' |

¹⁴ The results are as follows: 46 bare verbs (57%), 20 V+SA (25%), 19 V+ Infinitive/Imperative (23%), with the remaining 14 being of all different sorts.

¹⁵ Note that the subject and object also differ from the adult form in that case marking is often omitted. This is relevant only insofar as it further exemplifies the tendency towards omission. However, our focus here is on the acquisition of verbal morphology.

Recall that Quechua does have an infinitive marker that occurs at the end of the verbal complex. RIs, however, are not attested. 19 of the youngest child's 81 verbs occurred with the infinitive/imperative marker on it, but they were used in imperative contexts. Therefore these 19 utterances were used correctly, and not as unadult-like root clause infinitives.

1.4.2 Inuktitut

Inuktitut is a polysynthetic language spoken in arctic Quebec. The adult language does not permit bare verbs, and does not have a morphological infinitive. Swift & Allen (2002) report that normally developing Inuit children allow bare verbs that are ungrammatical in the target language. The typical verbal complex is made up of the verb root, several optional suffixes (e.g., modalization, negation, etc.), followed by an obligatory inflectional suffix. This final inflectional suffix is a portmanteau morpheme encoding person, number and mood.

Swift & Allen take their data from four Inuit children aged between 2;0 and 3;6. They found that while verbal inflection dropping occurs in adult speech, it is far more frequent in child speech than adult speech:

Table 1.11 Verbal Inflection dropping by Inuit children and adults

	Children	Caretakers
Uninflected verbs	145 (5.6%)	7 (0.29%)
Inflected verbs	2439	2409

They found that overall the omission of verbal inflection decreased with maturity, and that often the children produced verbs in both inflected and uninflected forms. The following is from the same child, aged 2;10:

- (24) a. piilaurit
 piiq – lauq – git
 remove–POL.IMP–2sS
 ‘You get off, please’
- b. pii!
 Piiq – Ø
 Remove–no.infl.
 ‘(you) get off.’

Swift & Allen conclude that the omission of verbal inflection cannot be due to a lack of knowledge of the inflection itself, but must be due to other factors. They investigate whether omission of inflection is due to discourse-pragmatic conditions, structural conditions and/or emotional conditions. They find that none of these conditions alone account for all verbal omission, suggesting that inflection omission is generally a much more complex phenomenon than is commonly assumed.

The details of their study are not crucial for our purposes here, simply the fact that Inuit children omit verbal inflection and allow bare

verbs at significantly higher rates than adult Inuktitut allows¹⁶. Thus Inuktitut is another language that can be added to the bare verb category.¹⁷

Thus we can add these languages to the inventory of languages that we saw in table 1.1 earlier.

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		

Our conclusion from this discussion then is that bare verbs are not restricted to English alone. English is thus one member of a class of diverse

¹⁶ Adult Inuktitut arguably does not allow bare verbs at all, given the rate of 0.29% cited by Swift & Allen. However, it is unclear from Swift & Allen's study whether adult native speakers disallow bare verbs in all discourse contexts.

¹⁷ Feurer (1980) describes the acquisition of Mohawk verbal morphology. The development of Mohawk verbal morphology is described as 'an expansion' from bare, minimal verbs to fully adult-like verbs. This, as we have seen in all the languages discussed so far, is the common pattern. In Mohawk, as in Navajo, the bare stem is unattested in the adult language. There is minimally one obligatory prefix which marks subject agreement. Examples are given of child utterances that are entirely missing any prefix. In addition, Feurer describes utterances that have a proto-prefix, which can be assimilated to Demuth's 'shadow vowel' (although in Mohawk such prefixes are not always vowels). The age of the child being described is 3;3 at the earliest stage, and it is only at this stage that the child produces bare verbs. I do not include a comprehensive review of Mohawk because the data and description in Feurer (1980) is limited.

a.	konó:lu	b.	kyó?teh
target form:	yo-kvnó:ru	target form:	wa-kyó?te?
	it-is-raining		I-am-working

languages, and assimilating English to RI languages misses this fact. One fact that this table fails to capture is that Sesotho, Siswati, Quechua, etc. differ from English in that the former group of languages have true infinitive markers (as RI languages). English, on the other hand, arguably does not have a real infinitive marker since 'to' behaves differently from infinitive markers in other languages (see chapter 2 for more discussion of this point, specifically the differences between French infinitives and English 'infinitives. See also Schütze, 1997). We will return to this point in chapter 5 in the conclusion.

The present study is concerned with the acquisition of Swahili, the acquisition of which we know very little about. We will see 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). Swahili may behave like the closely related language Sesotho, which allows bare verbs, but not RIs. The inflectional morphology of Swahili is particularly interesting because, unlike Sesotho, it independently and distinctly marks Agr and T. As we will see, this proves crucial in distinguishing among several theories in language acquisition.

In this study, I will describe the phenomenon of omission in Swahili, and show that in the case of Swahili, omission is primarily a syntactic process. I will show that the omission of verbal prefixes by Swahili children has effects on the occurrence of overt subjects. Specifically, we will see that the omission of tense results in the complete absence of subjects in early Swahili, a fact that follows directly from the theory of PRO in adult languages. This confirms previous accounts of null

subjects in child language that posit that null subjects are PRO (e.g., Hyams & Wexler, 1993; Bromberg & Wexler, 1995). I will also argue that in adult Swahili, the omission of Subject Agreement is possible when there is an anaphoric topic operator binding a null constant in subject position. I show that children even at early ages show evidence of having knowledge of this adult possibility. This provides evidence that children acquiring Swahili have access to universal principles of language since the evidence in the input for this construction is minimal. I will add further support to the general consensus that the knowledge that children exhibit about their language is based (at least in part) on universal principles of language.

1.5 Organization of the Dissertation

This dissertation is organized as follows. In chapter 2 I present a description of Swahili. I describe each inflectional affix in detail, as well as derivational suffixes, nominal morphology and some basic phonological facts. I present my theoretical assumptions, and then move on to two current debates in the Bantu literature that deal first with whether agreement prefixes are agreement or pronominal clitics, and second with whether tense in Swahili is a tense prefix or an auxiliary verb. These are active debates in the literature with no absolute consensus. I contribute to this debate because it is necessary in order to understand the status of the prefixes in this dialect of Swahili. I argue that the agreement prefixes (both subject and object) are both agreement and not pronominal, and that tense is a tense prefix, and not an auxiliary verb. I then go on to show that Swahili allows subject agreement to be omitted, and I present a typology of clauses that allow this. I propose that subject agreement omission occurs in the presence of a null constant bound by a topic operator, and I provide

evidence from embedding and quantified subjects. Thus, in addition to NP-traces, *pro* and PRO, adult Swahili has null constant clauses. The analysis that I propose asserts that adult Swahili exhibits several different null elements, which must be acquired by children. We will see in the later chapters that children indeed do acquire these null elements fairly early on.

In Chapter 3 I discuss the methodology, data collection, the subjects, transcription protocols, and the staging criteria. Chapter 4 deals with the emergence of the inflectional prefixes. I first outline some influential theories that deal with omission of inflectional affixes or RIs. I then present the Swahili data. This is the chapter in which most of the empirical results will be described, quantified and exemplified. I then return to the theories of language acquisition and evaluate them in light of the Swahili data.

Chapter 5 addresses the question of subject and subject agreement omission in child Swahili. I show that subject omission correlates with the various underspecified clause types. This is evidence that the omission of inflectional prefixes is not a phonological or purely morphological process, but a process that is syntactic in nature. Furthermore, the omission of subject agreement in child Swahili adheres to the principles of subject agreement omission in adult Swahili that we discuss in chapter 2, showing that subject agreement omission is not simply an error on the part of the child. I show that children exhibit knowledge of various silent elements very early on. The evidence for these silent categories is very sparse in the input, and thus I conclude that children acquire aspects of their language with little or no overt evidence in the input. This is evidence that children acquire language with the aid of universal principles of language.