

# **The Morphosyntax of Mood in Early Grammar with Special Reference to Swahili\***

---

\* Our thanks to Al-Amin Kimathi and Zein Shatry, who provided immeasurable help with the data collection. Our thanks also to Ginny Gathercole and two anonymous reviewers for comments and suggestions. This work was partially supported by a UCLA Faculty Senate Grant to N. Hyams.

## **Abstract**

In this paper we explore the development of the morphosyntax – semantics interface by comparing development in 4 typologically diverse languages, Dutch (a Germanic V2 language), Greek, Italian (a Romance pro-drop language) and Swahili (a Bantu language), with particular emphasis on Swahili, a relatively under-studied language whose morphosyntactic structure is particularly relevant to the questions we address. We show that children acquiring these different languages all adhere to a morphosyntax- semantics mapping principle that forces a complementarity between the expression of mood and the expression of tense -- the Semantic Opposition Hypothesis (SOH), following Hyams (2002). Our findings support the hypothesis that linguistic development is guided by universal principles that lie at the interface of semantico-conceptual structure and morphosyntax.

## 1. Introduction

Much recent acquisition research within the principles and parameters framework has focused on early morphosyntactic development. One notable result from this line of research is the finding that children converge on the core morphosyntactic properties of the adult “target” very early in the acquisition process. Language-specific morphology is acquired with few errors, and parameters of Universal Grammar (UG), for example those related to word order and null subjects, seem to be correctly set from the earliest multiword utterances. Thus, the early grammar of English is essentially English, early German is essentially German, and so on. These findings of an early morphosyntactic convergence (Hoekstra and Hyams 1998) raise in a perspicuous way the question of whether language development follows a universal course. In this paper we explore this question by comparing development in 4 typologically diverse languages, Dutch (a Germanic V2 language), Greek, Italian (a Romance pro-drop language) and Swahili (a Bantu language). Since the acquisition of Swahili has been far less studied than the other languages, our discussion of Swahili is considerably more detailed.. The sections on Dutch, Greek and Italian summarize previous research and provide a background and theoretical context within which to approach the Swahili data.

We will focus in particular on the semantico-conceptual categories of tense and mood and their realization in the morphosyntax of these different languages. We will propose that a universal principle – the Semantic Opposition Hypothesis – mediates the mapping of tense/aspect/mood (TAM) categories onto the morphosyntax of specific languages during acquisition.<sup>1</sup> We hope to show that despite differences in the morphosyntax of particular (child) languages, children’s grammar development is guided by universal principles and share certain fundamental properties.

The paper is organized as follows. In the following section we review several studies of the realis-irrealis mood distinction as it is manifested in the various European child languages noted above and we lay out our assumptions concerning the relation of realis and irrealis mood to the morphosyntax of tense and finiteness. In sections 3 and 4 we examine the acquisition of tense and mood in Swahili. Section 3 provides some background on the morphosyntax of Swahili and in section 4 we show that the Swahili child data are consistent with the finding of the European languages. Section 5 presents our conclusions.

## **2. Mood and finiteness in early grammar**

Recent research within the generative framework has revealed a great deal about the morphosyntax and semantics of finite and non-finite structures in early language, especially the properties of root infinitives -- root clauses containing an infinitival verb, which occur robustly in various child languages (See Wexler 1994; Sano and Hyams 1994; and Hoekstra and Hyams 1998 for review). Hyams (2001, 2004) observes that in various child grammars a semantic opposition between realis and irrealis mood (cf. also Stephany 1986; Wagner 1998) is realized in the morphosyntax through the specification/ underspecification of I(nflectional) features. Specifically, in each of the languages to be discussed below there is a (arguably) non-finite clause that is realized in irrealis contexts, which we define as contexts of volition, direction and intention. The non-finite irrealis clause stands in opposition to well-formed clauses with [+/-past] temporal reference that occur in realis contexts. The specific form of the non-finite verb differs depending on the particular morphosyntax of the target language. In some languages, Dutch and German, for example, irrealis mood is expressed by an RI. In the Romance pro-drop languages such as Italian, which do not show an RI stage, the imperative is the functional analogue of the RI, showing up in (at least a subset) of irrealis contexts (Salustri and Hyams

2002), and in Greek a bare perfective form occurs in irrealis contexts (Hyams 2002, Stephany 1986; 1995).

The generalization that emerges from these various child languages is given a more precise formulation as the Semantic Opposition Hypothesis (SOH) in (1).

- (1) *The Semantic Opposition Hypothesis*  
The expression of irrealis mood in the early grammar excludes a tense specification.

In other words, the morphosyntactic expression of mood and tense are in complementary distribution in the early stages of grammar development.

We begin in the following section by discussing root infinitives and we will also provide a structural explanation for the generalization in (1).

## 2.1 Root infinitives

Of the various non-finite forms that children produce special attention has been devoted to the RI phenomenon. RIs are robustly found in many child languages, for example, German, Dutch, French, Swedish, Russian, among others. The percentage of RIs (out of all verbs) ranges between 40% and 80% depending on the particular language and child and level of development (see Blom 2003; Hyams 2001; Sano and Hyams 1994; Wexler 1994). Examples are given in (2).

- |     |    |   |         |
|-----|----|---|---------|
| (2) | a. | Niekje buiten spelen<br>Niekje outside play-in                | Dutch   |
|     | b. | Pas aller à l'eau<br>not go to the water                      | French  |
|     | c. | Auf Teddy fenster gucken<br>also Teddy window look-inf.       | German  |
|     | d. | Jag också hoppa där å där.<br>I also hop-inf. there and there | Swedish |

In this paper we define the RI stage as the stage in which children alternate between finite and non-finite verbs, what Wexler refers to as the Optional Infinitive stage (Wexler 1994). In

some languages, Dutch for example, there is an earlier stage during which children use RIs exclusively (Blom & Wijnen 2000). We will have more to say about the ‘RI-only’ stage when we turn to the Swahili data in section 4. (See also note 2). Blom (2003), following de Haan (1987) and others argues that the earliest finite forms, typically modals, are not true finite verbs, a point we substantially agree with. We return to this below.

While children in the RI stage produce both RIs and adult-like finite verbs, RIs have a very different distribution than their finite counterparts. For example, RIs fail to undergo typical verb movement processes such as V2 in the Germanic languages or V to I raising in French, as illustrated in (2a) and (2b), respectively. Also, RIs typically occur with null subjects (though not always, as illustrated by the examples in (2)); they fail to occur in *wh* questions and topicalized sentences; and they do not occur with subject clitics in French (See Hoekstra and Hyams 1998 for review of the relevant properties.)

The different distributions of finite verbs and RIs suggest that two different structures are involved and also that the early grammar allows non-finite root clauses as a grammatical option. If RIs were the result of a performance limitation we would not expect the grammatical contingencies noted above. If children were simply reproducing surface forms that they hear in the input (Ingram & Thompson 1996) then we should wonder why their reproductions are so selective, why don’t they reproduce the modal as well? We thus adopt the hypothesis that in the early grammar non-finite verbs are grammatical in root contexts.

There has been considerable debate over the nature of the mechanism that licenses RIs in early grammar. Rizzi (1994), for example, proposes that children have the option of truncating the clause below CP; truncation at the VP level yields a root infinitive. Other accounts hold that one or more functional head (Tense, AGR, Number) can be left underspecified in the early

grammar and the underspecification results in an absence of finiteness (e.g. Wexler 1994; Schütze and Wexler, 1996; Hoekstra and Hyams 1995 among others). A somewhat different underspecification account is proposed by Clahsen, Eisenbeiss and Penke (1996), the lexical learning hypothesis, which holds that some functional structure is initially absent in the child's grammar and is built up as the child acquires the agreement paradigms of the language. The account we will propose shortly bears some formal similarity to this hypothesis, but it also differs from it in crucial respects (cf. note 7). More recently, Blom (2003) has proposed that the RI phenomenon is to be captured strictly at the morphological level. Much in the spirit of distributive morphology (Halle and Marantz, 1993), Blom proposes that infinitives are a default verb form inserted as a 'placeholder' until children learn the relevant finite form (cf. also Ferdinand 1996). This proposal is discussed further in section 2.2.

While much has been learned about the morphosyntax of non-finite clauses in early language, less is known about the meaning of these structures. Increasingly, however, studies are investigating the semantic properties of RIs and other non-finite clauses. Hoekstra and Hyams (1998), based on work by Wijnen (1997) and others, noted that RIs in Dutch and other languages have an essentially modal or irrealis meaning – that is, they express the child's needs, wishes, intentions with respect to some eventuality (cf. also Blom 2003 for Dutch; Plunkett and Strömquist 1990 and Joseffson 2002 for Swedish; Labelle 2000 for French; Ingram and Thompson 1996, Becker and Hyams 1999; Lasser 1997 for German.). Hoekstra and Hyams referred to this finding as the Modal Reference Effect (MRE), stated in (3).

- (3) *The Modal Reference Effect*  
With overwhelming frequency RIs have a modal/irrealis meaning.

Tables 1a and 1b (adapted from Wijnen 1997) report the referential properties of RIs and finite verbs from 4 Dutch children (Josse 2;0 –2;6; Matthjis 1;11-2;8; Niek 2;7-3;2; Peter 1;9-2;1, available on CHILDES, MacWhinney 2000). Approximately 86% of RIs have a modal or irrealis meaning while finite verbs have temporal reference, usually present.<sup>2</sup>

[Insert Table 1a here]

[Insert Table 1b here]

We use the terms ‘modal’ and ‘irrealis’ interchangeably to refer to event denotations that are not situated with respect to speech time, unlike past and present events, but rather include categories such as intention, direction, volition (and others that are not relevant to the acquisition data we consider). Thus, the modal/irrealis utterances express the child’s wish, intention or need with respect to some as yet unrealized state of affairs. We also follow Giorgi and Pianesi (1997) and others (e.g. Portner 1994; Farkas 1992) in assuming that modals and mood are different grammatical manifestations of modality. Modals (e.g. *will*, *must*) are a lexical realization of modality (merged directly under MoodP in a syntactic tree) while mood is an inflectional realization (e.g. subjunctive, imperative mood) of modality that checks the features of MoodP either through movement of the verb (Move) or Attract (cf. Chomsky 1995). We return to the syntax of irrealis mood below.

In Hoekstra and Hyams (1998) it is proposed that the modality of the RI is specifically associated with the infinitival morphology. To the extent that this is true, children have zeroed in on an essential property of adult grammar since infinitives typically have irrealis meaning in



adult grammar as well, as argued by Stowell (1982), Bolinger (1968), Duffley (1992), and most recently Han (1998). This is illustrated by examples such as (4):

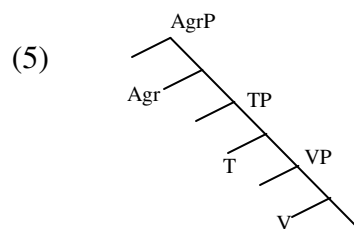
- (4) a. John remembered to bring the wine.
- b. Niet parkeren hier  
          NEG Park-inf here  
          'No parking here'
- c. What to do?/ Che fare?
- d. Non tornare a casa troppo tarde.  
          Not come-inf. home too late  
          'Don't come home too late'

As pointed out by Stowell, in sentences such as (4a) the infinitive *to bring* is unrealized at the moment of remembering, jussives as in the Dutch example in (4b) express necessity, expressions such those in (4c) *what to do*, *where to eat*, etc. in English and many other languages typically have a modal meaning, roughly 'what should we do', 'where should we go', and so on, (4d) illustrates that in Italian and many other varieties of Romance negative imperatives are formed with the infinitive (as discussed by Zanuttini 1997 and others).<sup>3</sup>

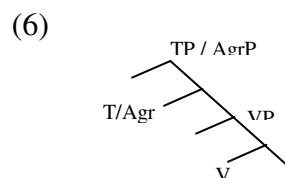
The appearance of an infinitive in these various kinds of irrealis contexts is unlikely to be a coincidence and the child data strongly suggest that whatever principle is responsible for the irrealis meaning in adult language is also present in child grammar. Following Hoekstra and Hyams (1998) (cf. also Han 1998), we will assume that the infinitival morpheme has an irrealis feature. In line with most current syntactic research (Chomsky 1995 and others), we also assume that features in the verb must match the categories projected in the syntax: Thus, tense features on the verb are "checked" against a Tense category; mood (i.e. irrealis) features are checked against a Mood category; aspect features against an Aspect projection, and so on. Conversely, functional structure, which gives rise to temporal/aspectual/modal meanings, must be licensed by corresponding features in the verb. Thus, T(ense)P must be licensed by tense features in the verb and so on. Alternatively, a syntactic projection may be licensed by a lexical item inserted into its

head, for example a modal in the head of MoodP. In short, the model of grammar we assume is one in which inflectional features, which may be morphologically (or lexically) realized, license syntactic structure and the interpretation is given by the structure.

A final crucial assumption we make is that languages vary with respect to how they bundle inflectional features into syntactic categories (Giorgi and Pianesi 1997; cf. also Bobaljik and Thrainsson 1998). For example, in Italian and other languages tense and agreement features can be realized independently. The past tense form *am-av-o* (love-past-1<sup>st</sup> per.) consists of a past tense morpheme and an agreement morpheme. The same is true in German, e.g. *du lieb-te-st* (you love-past-2<sup>nd</sup> per.). The features in these independent morphemes are checked against separate Tense and Agr(eement) heads in the syntactic structure, as in (5).



English, on the other hand, expresses either tense (e.g. *loved*) or agreement (e.g. 3<sup>rd</sup> per. *loves*) but never both (*\*he loveds*). Giorgi and Pianesi (1997) argue that the reason English has no such form is that Tense and Agreement features are bundled together into a unitary category in this language, as illustrated in the tree in (6).



Since English has only one head, only one set of features (either tense or agreement) can be morphologically realized, but not both. In terms of its temporal reference the specification of Agr features (-s) entails the unmarked [-past] value.

Thus, languages differ with respect to whether particular pairs of inflectional features form independent or unitary categories, and hence the child must learn how different features are grouped together in the syntax of their particular language. Learnability considerations thus require that separate categories be projected only when there is positive evidence (in the form of multiple morphemes or multiple specifier positions (cf. Giorgi and Pianesi 1997)); otherwise a unitary category is assumed. In other words, unitary categories are the default, hence initial hypothesis for the child.

Just as Tense and Agr can constitute a single category, so can Tense and Mood (Giorgi and Pianesi 1997). The English modals, in contrast to German or Italian modals, for example, are unmarked for tense and agreement features (e.g. *\*musted, musts*). This follows from the structure of the English clause in which MoodP/TP constitute a unitary category (cf. also Roberts 1993).<sup>4</sup> Another example is provided by the subjunctive and imperative, which express certain kinds of irrealis meanings. The English subjunctive (for speakers who accept it) is possible only in the absence of finite morphology (viz. *I demand that he go/ \*I demand that he goes*). Similarly, the imperative is uninflected for tense (*Leave now! \*Left now!*).

The complementarity of mood and tense features is illustrated most clearly in a language like Swahili, which has overt subjunctive morphology. As shown in the examples in (7) the Swahili verb must be marked for tense, unless it is subjunctive, in which case it bears only a subjunctive morpheme. We return to Swahili in section 3.

- (7) a. A – ta – fik – a kesho? [+T] Indicative  
 SA<sub>3s</sub>–fut–arrive–IND tomorrow  
 ‘Will he arrive tomorrow?’
- b. \*A – fik – a kesho? [-T] Indicative
- c. Ni – fik – e kesho? [-T] Subjunctive  
 SA<sub>1s</sub>–arrive–SUBJ tomorrow  
 ‘Should I arrive tomorrow?’
- d. \*Ni – ta – fik – e kesho? [+T] Subjunctive

On the other hand, in a language like Italian tense and mood can be realized independently as in the past subjunctive (e.g. *Magari andasse via* (I wish (he) go-past-subj. away) or the past tense of modal verbs such as *potere* (can) or *dovere* (must). (e.g. *dov-e-va/pot-ev-a.-* (must-past-3<sup>rd</sup> per/can-past-3<sup>rd</sup> per.).

Although rather sketchy, this review of our background assumptions is sufficient to allow us to lay out a structural explanation for the SOH in (1). The essence of our proposal is that the child’s initial assumption is that MoodP/TP is a unitary category. Hence, in the child’s grammar a verb may check tense features or mood features but not both (in the same structure). It follows from this hypothesis that the grammatical expression of modality in the early grammar, including volition, direction, and intention is realized by a non-finite form with appropriate features. If the verb has an irrealis feature to check against the head of MoodP, then it cannot also check tense features and TP is not licensed, for the reasons outlined above.<sup>5</sup>

## 2.2 Alternative accounts of the modality of RIs

It has been suggested that the modal meaning associated with RIs arises not from any irrealis specification or functional structure as we have claimed, but is rather a "default" option. One such approach is to say that the modal meaning is the default *meaning* in the absence of a tense specification or temporal anchor (K. Wexler, p.c.). A second approach is to say that the RI is a

default *form* that occurs in the absence of a more specific form, in this case the modals (Blom 2003). We formulate these hypotheses as in (8a, b):

- (8) The modality of RIs arises because:
- (a) Irrealis mood is a default interpretation in the absence of tense
  - (b) RIs are a default form in the absence of modals

The hypothesis in (8a) is rather counterintuitive on its face. Why would the default temporal meaning be the most semantically complex one? Why wouldn't the default meaning simply be the 'here and now' associated with the speech time (e.g. Smith 1980)? Moreover, (8a) is also empirically insufficient insofar as there are other non-finite clauses that occur in early language that do not have an irrealis meaning. For example, in Italian and other languages, children produce bare participles, as in the examples in (9). Though non-finite, these structures have the perfective/completive meanings typically associated with participles and never have an irrealis meaning.

- (9)
- |    |   |  |
|----|---|--|
| a. | Caduta<br>fell                          | Intended meaning: 'I have fallen'<br>Adult equivalent: sono caduta                               |
| b. | Buttata via, guarda<br>thrown away look | Intended meaning: 'I have thrown (it) away, look'<br>Adult equivalent: l'ho buttata via, guarda' |

Indeed, Blom (2003), following Lasser (1997), analyzes these participial forms as [-tense, +completed] while RIs are analyzed as [-tense, -completed]; in other words the [-tense] feature specification alone is not sufficient to guarantee a modal meaning.

English is another case in point. Hoekstra and Hyams (1998) show that the English bare form (arguably a non-finite form – cf. Wexler 1994) most often (i.e. 89% of the time) has a temporal meaning, past or present, as illustrated by the examples in (10) and the quantitative data in table 2. .



The analysis proposed in this paper is also consistent with this RI/modal trade-off, as we assume that the finite modals (e.g. with past tense morphology) constitute a potential trigger for the Dutch child to learn that TP and MoodP are not a unitary category in Dutch. We thus expect that as the modals become productive, leading to the independence of TP and MoodP, RIs will decrease.

The Dutch modals raise an interesting question for our proposal, as a reviewer points out to us. As noted by De Haan (1987) and others, the earliest finite verbs include modals (*kan*, ‘can’ *moet* ‘must’, *mag* ‘may’) and these modals show the syntactic behavior of finite verbs in that they appear in second position. If TP/MoodP is a unitary category at the initial stage this is unpredicted. De Haan argues that these early modals express modality but they are not in fact true finite forms in that do not show systematic marking for number or person. A similar claim is put forth in Hoekstra and Jordens (1994), who observe that the early modals occur almost exclusively with negation (e.g. *kannie* ‘cannot’, *magnie* ‘may not’). Hoekstra and Jordens propose that these elements express negative modality, but are not analyzed as true finite forms. Following the proposals of de Haan and Hoekstra and Jordens, we will assume that the early Dutch modals license MoodP but not TP and are therefore entirely consistent with the SOH.

Assuming a compositional semantics for natural language, a more conceptual issue arises for the hypotheses in (8), namely, if the modality of RIs is not a feature of the verb or a feature of the structure, how then is the semantics of the sentence computed? If the modality does not come from the sentence itself, we are forced to assume that the child’s semantics is of quite a different sort than the adult’s. It seems to us that a more principled explanation is one that avails itself of an independently motivated grammatical theory and assumes that children construct meanings according to compositional principles, just like adults,

In the following sections we briefly review some further support for the SOH from Greek and Italian before turning to Swahili.

### 2.3 The Greek bare perfective

Greek is a language without an infinitive and yet Greek children produce a form that shares many of the temporal properties of RIs (Hyams 2002). Stephany (1981, 1986), and more recently Varlokosta, Vainikka & Rohrbacher (1998), note that Greek children at around age 2 frequently produce utterances such as in (11) (cf. also Tsimpli 1992).

- (11) Ego katiti (Spiros 1;9)  
I sit-PERF- 3rd sing.  
'I am going to/want to sit'

In terms of its morphological form, the verb is a present perfective. We refer to this construction as the 'bare perfective' since the child's verb is not marked for past tense nor is there a future or subjunctive particle, as is required in the adult grammar. These utterances are thus ungrammatical in the adult language. Indeed, present perfectives are universally ungrammatical because of a semantic incompatibility between completive aspect and the here-and-now value of the present tense. (See, for example, Giorgi and Pianesi's (1997) punctuality constraint').

Varlokosta et al. (1998) argue that such sentences are non-finite, noting that the verb often fails to agree with the subject, occurring with the 3<sup>rd</sup> person *-i* ending, though the intended subject is not 3<sup>rd</sup> person, as illustrated in (12).

- (12) Pàri γυγάκι  
Take-perf. 3<sup>rd</sup> sing. Piggy  
'May I take the piggy?'

The rate of agreement errors in Greek for the 2 Stage 1 children (Spiros, 1;9, and Janna, 1;11 from the Stephany corpus (Stephany 1997) in the CHILDES database, MacWhinney 2000)



examined by Varlokosta et al and Hyams is around 39% (38% and 40% for Spiros and Janna, respectively), a result that is quite out of line with other child languages, where the rate of agreement error is under 4% by most counts (cf. Hoekstra and Hyams 1998 for discussion). However, the lack of consistent agreement is not surprising if the verb is non-finite. In the absence of tense the verb appears in the default 3<sup>rd</sup> person form.

According to Stephany (1981, 1986), Tsimpli (1992) and others, the interpretation associated with utterances like (12) is clearly modal or irrealis, expressing intentions, wishes, and obligations. Thus, with respect to interpretation, it is very close to the Dutch RIs. As pointed out by Stephany, young Greek children have two classes of verbal utterances, those which describe ongoing and past events, for which adult-like present imperfective and past perfective forms are used.. Opposed to these are utterances that have modal meanings, such as the examples in (11) and (12), in which the verb is a bare perfective. Table 3a presents the percentage of perfective verbs with modal reference as compared to the percentage of imperfective verbs with modal reference for the 2 Greek children discussed above, based on data from Stephany (1997) and p.c.

[Insert Table 3a here]

Table 3b compares the aggregate Greek results to the aggregate Dutch results presented in table 1a. In Greek, 77% of the bare perfectives have modal reference, slightly lower than the 86% of Dutch RIs with modal reference, but significantly more than the imperfectives of which only 2% occur in irrealis contexts.

[Insert Table 3b here]

It is easy to see how the categories described by Stephany fit the generalization in (1); In Greek as well we find a non-finite form that expresses irrealis mood. The Greek results are significant because we are finding RI-type effects in a language that does not have an infinitive. This tells us that the mood-(non-)finiteness relation, despite its generality across (typologically distinct) child languages, is sensitive to the particular morphosyntax of the target language. The SOH is by hypothesis a developmental principle that occurs at the interface of semantics and morphosyntax: Universal semantic oppositions (realis-irrealis) find their realization in the specific morphosyntax of a language. We thus expect the semantic oppositions to play out differently in typologically distinct languages. With this in mind, let's turn to Italian.

#### **2.4 The Italian imperative**

Italian, unlike Greek, does have an infinitival form. It is well known, however, that Italian children (like children acquiring other Romance null subject languages) do not go through an RI stage, in marked contrast to the Dutch (and German, Swedish, etc) children discussed earlier (Sano and Hyams 1994; Guasti 1993/4). Yet, the logic of the SOH leads us to expect that there will be an RI analogue in Italian (and the other Romance null subject languages). Salustri and Hyams (2003) provide evidence from 3 monolingual Italian children and one bilingual Italian-German child that there is indeed such an analogue, which they identify as the imperative.

Salustri and Hyams adopt Han's (1998) description of the imperative as a form that is marked with an 'irrealis' feature that contributes an unrealized mood interpretation (and a 'directive' feature encoding directive illocutionary force). Thus, the imperative shares the essential properties of RIs. It is irrealis in that it expresses a direction or wish to bring about a state of affairs that is unrealized at speech time and it is tenseless. In contrast to RIs and bare perfectives,

imperatives are fully grammatical in adult language and so the simple appearance of imperatives in the child's language is not remarkable. If, however, the imperative in child null subject languages represents an RI analogue, it is expected to have some distinguishing properties. Salustri and Hyams make the following two predictions: (i) In null subject languages imperatives will occur significantly more often in child language than in adult language; and (ii) in child language imperatives will occur significantly more often in null subject languages than in the RI languages. The logic behind (i) is that over time some imperatives will be replaced by adult-like modal expressions (as occurs in Dutch in which modals replace RIs, see previous section). A similar reasoning gives rise to (ii): if RIs and imperatives both express irrealis mood, then RIs might bleed imperatives in the RI languages.

Salustri and Hyams show that both of these predictions are borne out by the data. The rate of imperatives in adult child-directed Italian is about 15% (based on adult tiers in the CHILDES files). The rates for the 4 children (Martina, 1;10-2;7; Diana, 1;8-2;6; Viola, 2;1-2;7; Denis, 1;5-2;2; CHILDES, Calambrone corpus, MacWhinney 2000; Leonini 2002) peaked at around 40% somewhere between the ages of 2;0 and 2;4 and then dropped to adult-like levels by about age 2;6 or 2;7.

As predicted, the rate of imperatives is considerably higher for Italian children than for Italian adults, even in child-directed language. Moreover, the "overuse" of imperatives happens at the same age as the RI stage in the RI languages (roughly between the ages of 2;0 and 2;6). Notice as an aside that despite the salience of imperatives in the input (e.g. special prosody, first position) (Newport, Gleitman and Gleitman 1972), the child frequencies do not at all mirror the adult frequencies, as would be predicted by statistical learning models.

To test the second prediction, Salustri and Hyams compared the rate of imperatives in early Italian and early German. As predicted, the 3 monolingual German-speaking children (Caroline, 1;3-2;6; Kerstin, 2;0; Simone, 2;0-2;7; CHILDES Nijmegen corpus, MacWhinney 2000) use far fewer imperatives during the relevant period (between the ages of 1;6 and 2;7) (and far more RIs). The overall rate of imperatives in early German – about 10% -- was quite low as compared to German adults (who had about 36% imperatives) and as compared to Italian children. Also, the frequency of imperatives remained fairly constant across all data points with no peak during the first half of the 3rd year as was found for the Italian children.

Interestingly, the difference in imperative/RI use was also found in a bilingual German-Italian child, Leo, between the ages of 2;0 to 2;7. The bilingual child is the perfect controlled experiment and his data perfectly mirror the cross-linguistic differences found in the monolinguals. Leo's overall rate of imperatives in Italian (47%) was far higher than in German (3%); conversely his use of RIs was substantially higher in German (72%) than Italian (4%). We take this as strong support for the hypothesis that the imperative is the RI analogue in Italian, and further support for the SOH.

An important question that we do not address here is why different non-finite forms instantiate the irrealis mood in the different child languages. Why doesn't early Italian have a true RI stage, instead of the imperative? Conversely, why doesn't the imperative rather than the RI instantiate irrealis mood in German, and so on? In Salustri and Hyams it is proposed that the RI is the unmarked (or most economical) option for irrealis mood (because it has the simplest syntax) unless there is some independent reason why the infinitive is blocked. In Greek there is no infinitive, so this is not an option. The Italian infinitive, on the other hand, has (abstract)

agreement features that must be checked in TP (Belletti 1990), making finiteness obligatory in root clauses (cf. also Rizzi 1994).

Central to the SOH is the requirement that irrealis mood be expressed by an element with an irrealis feature. This is a property that is shared by the infinitive, the imperative and the subjunctive (Han 1998), precisely the forms children adopt. In Dutch and other RI languages, etc., irrealis mood is instantiated by the infinitive, in Italian by the imperative, and in Greek by a verb form that occurs under a subjunctive particle in the adult language. The fact that the imperative has a narrower range of meanings than the RI, which expresses intention and volition in addition to direction, is entirely consistent with the analysis proposed here, which is based on the formal requirement that Mood be licensed by a verb with a relevant, i.e., irrealis, feature, and not on functional requirements that the RI and RI-analogues all express the same range of meanings.

We turn finally to Swahili. Our claim is that the same opposition that we see in the languages just discussed between irrealis and realis mood expressed by a non-finite and finite form, respectively, also shows up in Swahili. We will see that once again the children avail themselves of a non-finite irrealis form, the subjunctive

### **3.0 Mood and finiteness in Swahili**

Before presenting the Swahili child data, we first describe the structure of the Swahili verbal complex. Swahili, an eastern African Bantu language, is an SVO language in which the subject and the object are optionally null. The verb occurs in a verbal complex that contains inflectional material as well as grammatical function changing suffixes. The structure of the verbal complex is given in (13).

(13) Subject Agr. – Tense – Object Agr. – Verb–Derivational suffixes– Mood

There are several tense/aspect markers (Ashton 1947; Vitale 1985), given in table 4.

These markers occur in the preverbal position indicated in (13) above, and occur only when the verb is in the indicative mood.

[Insert Table 4 here]

Mood is specified by the final vowel in the verbal complex. There is a 3-way alternation between the unmarked indicative [a], the subjunctive [e] and the negative [i]. The unmarked form occurs with on-going actions/states, present habitual actions, past actions/states, future actions/ states and imperatives. Examples are given in (14).

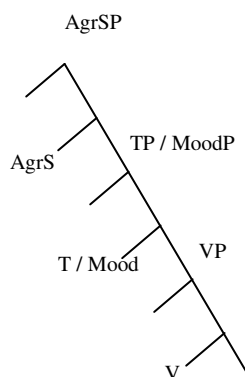
- |      |    |  |                      |
|------|----|--|----------------------|
| (14) | a. | Juma alimfuata Mariam<br>Juma a – li – m – fuat – <b>a</b> Mariam<br>Juma SA <sub>3sg</sub> -past-OA <sub>3sg</sub> -follow-IND Mariam<br>'Juma followed Mariam' | Indicative, past     |
|      | b. | Tafadhali nipatie kalamu<br>Tafadhali ni – pat – i – <b>e</b> kalamu<br>Please OA <sub>1sg</sub> -give-applic-SUBJ pen<br>'Please give me a pen'                 | Subjunctive, request |
|      | c. | ah, sasa nilale<br>sasa ni – lal – <b>e</b><br>now SA <sub>1s</sub> -sleep-SUBJ<br>'Ah, now I (want to) sleep'   | Subjunctive, desire  |
|      | d. | Anataka kufua dafu<br>A – na – tak – a <b>ku</b> – fu – a dafu<br>SA <sub>3sg</sub> -pres-want-IND INF-husk-IND coconut<br>'He wants to husk a coconut'          | Embedded infinitive  |
|      | e. | Soma!<br>som – <b>a</b><br>Read-IND<br>'Read!'   | Imperative           |

The sentence in (14a) is a transitive indicative sentence. The verbal complex shows 3<sup>rd</sup> person singular subject agreement with *Juma*, contains the past tense 'li', and 3<sup>rd</sup> person singular

object agreement with *Mariam* (a specific DP). The verb is in the unmarked indicative form. (14b) is a polite request and the verb is subjunctive. Notice that when the subjunctive final vowel is used, the tense marker is absent. This complementary relation between tense and mood is very relevant to the analysis of non-finite forms in early grammar proposed in the previous sections and we will return to this issue below. The verb in (14c) is also subjunctive, this time expressing a desire. The sentence in (14d) contains an embedded infinitive. Note that the Swahili infinitive marker ‘ku’ is a prefix with the same distribution as the tense markers in tensed clauses. The final example is an imperative form, which is also an unmarked (indicative), tenseless form.

We assume that in adult Swahili MoodP and TP form a unitary category (as in Giorgi and Pianesi 1997), as shown by the complementarity of tense and mood in the examples in (7) above. The clausal structure that we assume for Swahili is shown in (15).<sup>6</sup>

(15)



In the next section we will present the results of an analysis of the acquisition of inflectional elements in child Swahili, focusing in particular on tense. We will see that the development of tense morphology parallels the development of subjunctive mood morphology and thus supports the analysis of Tense and Mood as a unitary category and the early subjunctive as an RI analogue.

### **3.1 Subjects**

Turning now to the acquisition of Swahili, the data collection was conducted over a period of 11 months in Nairobi, Kenya. Biweekly audio recordings were made of naturalistic speech in the homes of four children of differing ages. The children were recruited through personal contacts and referrals. Seven families were considered for inclusion in the study, of which three were deemed unsuitable (on the basis of the linguistic environment of the home, practical issues such as distance and safety of the home, etc.). Table 5 provides for each child the age range, the number of recordings, and the MLU range (calculated as morphemes per utterance).

[Insert Table 5 here]

### **3.2 Method**

Each session was audio recorded using an external microphone placed in the center of the room. Audio recording was the preferred method of data collection (as opposed to video recording) because of the relative ease of transporting the equipment to the homes and the relatively inconspicuous nature of the recorders (as opposed to a video camera on a tripod).

One caregiver (and occasionally a second caregiver) and the investigator were present in the room with the child during all recording sessions (all of which took place at the home of the child). The investigator took contextual notes, while the caregiver engaged the child in speech about a free topic. These contextual notes were later used in transcribing and coding.

The data were transcribed using CHAT format (MacWhinney, 2000) and each transcript was then checked by a native speaker assistant. Morphological coding was inserted into the



transcript (in CHAT format) during the transcription process. As mentioned earlier, contextual notes were used to aid in the process of determining the function of particular morphemes. As a method of post-hoc verification of the coding process, each transcript was checked in the following manner. First, a random sample of the transcript that contained a substantial number of codes (at least 100 codes) was selected and all codes removed. Then a native speaker assistant re-coded the transcript, and comparisons were made to the original coding. It should be noted that the majority of errors that occurred were due to typographical errors or misunderstandings of the coding procedures, not interpretation of the data. See below for more details on determining the interpretation of utterances in the child data.

Because of various social and economic difficulties, it was not possible for all four children to remain in the study for the duration of the project. Thus, the following method was used to compare data from the four children. Each child was assigned to a particular stage or stages according to 3 measures of grammatical development: MLU, verbs per utterance (Valian 1991) and proportion of protosyntactic devices (Bottari, Cipriani, and Chilosi 1993/1994; Peters, 2001). The criteria used to establish the stages are given below in table 6.

[Insert Table 6 here]

Each file was examined with respect to the three measures of grammatical development and the stage into which each file fell was determined according to a points system, e.g., an MLU of 2.0 received one point, while an MLU of 2.7 received two points, and an MLU of 3.2 received three points, etc. The stage to which each file was assigned was determined on the basis of the combined points that each file scored across all three grammatical measures. We

then pooled the data from each stage. (See Deen 2002 for further details on the children and the establishment of the stages). Because the data from Mus come from such a lengthy time period (almost an entire year), data from this child was divided into three of the four developmental stages (stages 1, 2, and 3). We will refer to these four stages for the initial analyses of Swahili.

### 3.3 Data

Our first analysis was to determine the number of indicative, subjunctive and negative sentences. This was done with a computerized count using CLAN programs (MacWhinney 2000). Imperatives, repetitions, imitations and formulaic utterances were excluded from this analysis and the results are presented by stage in table 7.

[Insert Table 7 here]

Focusing on the indicative clauses first, we found that Swahili children produce verbal complexes of the kinds listed in (16):

- |      |    |                 |            |
|------|----|-----------------|------------|
| (16) | a. | Full Clause     | SA – T – V |
|      | b. | [-SA] Clause    | ∅ – T – V  |
|      | c. | [-T] Clause     | SA – ∅ – V |
|      | d. | Bare Stem       | ∅ – ∅ – V  |
|      | e. | Root Infinitive | INF – V    |

We see from these data that in Swahili, subject agreement and tense may be independently omitted (a fact that is consistent with Schütze & Wexler's (1996) ATOM; See Deen 2001; 2003b for discussion of the development of tense, subject agreement and subjects in child Swahili). However, the relative rates of each clause type differ markedly, as indicated in table 8.

[Insert Table 8 here]

Table 8 shows that full clauses increase over time from 18% to 60%, while [-T] and bare stems diminish over time from a combined 52% (20% -T and 32% Bare Verb) in stage 1 to 11% in stage 4. Interestingly, [-SA] clauses fluctuate across time but show no strong downward movement during the observation period, in marked contrast to the forms that are underspecified for tense. A variety of tenses occur in stage 3 in various contexts and on various verb stems. Thus not only does tense become obligatory in stage 3, the variety of tense markers also indicates productivity in the stage 3 grammar. See Deen (2002) for more details on the variety of tense markers that occur in the child data. Figure 1 shows the decrease in [-tense] sentences across the 4 periods.

[Insert Figure 1 here]

Thus, the most significant development that takes place is the shift from a grammar in which tense specification appears optional to one in which it is obligatory. As we will see in section 4, a second, parallel development occurs in the domain of mood. During the initial stages children use only the default indicative form; at stage 3 subjunctive morphology becomes productive. We will argue that stage 3 is the point at which the semantic opposition between realis and irrealis is mapped onto the relevant morphosyntax.

It is also clear from these data that Swahili children do not produce RIs. In fact, at all stages fewer than 2% of all indicative utterances are RIs. Note that this is the percentage out of indicative utterances, not all verbal utterances. So compared to RI rates given for other

languages, the rate we find in Swahili is exceedingly low, lower even than languages such as Italian and Spanish where the RI rate is around 8%. Swahili, then, is clearly not an RI language of the German or Dutch sort. We will show below, however, that the development of the semantics-morphosyntax mapping in Swahili parallels its development in RI languages such as Dutch.

#### **4.0 The Semantic Opposition Hypothesis**

The SOH holds that the realis-irrealis opposition is a primitive one in early grammar that is realized in the morphosyntax of the particular target language. Since Swahili is not an RI language, the question arises as to how irrealis Mood is licensed in the early grammar if not through infinitival morphology. One obvious candidate is the bare stem, since it is ostensibly a non-finite form. As shown in table 8 above, bare stems occur at a rate of 32% in stage 1, diminishing across the stages to 4% in stage 4 – exactly the developmental profile that we would expect from the SOH. However, an analysis of the bare stems in the corpus shows that the vast majority of bare stems carry a temporal, realis interpretation (see table 14 below). Thus, the bare stem, though non-finite is not an irrealis form, and hence not a possible RI-analogue within the framework proposed here. We will return to this issue shortly.<sup>7</sup>

Given the morphosyntax of Swahili, a more plausible candidate for the expression of irrealis mood in early Swahili is the subjunctive form, as in the adult grammar.<sup>8</sup> Accordingly, we investigated the use of subjunctives in the child data. The following generalizations emerge: First, overall, children seem to use fewer subjunctives than adults do. In child-directed adult speech within this corpus the rate of subjunctives is almost 25%. However, since adults use a disproportionate number of polite requests with children, this figure is most likely not representative of normal adult-adult interaction. In the child data we find between 2% and 12%

of their verbs are subjunctive. Importantly, we noticed a developmental trend. In stages 1 and 2, the children use subjunctives at the rate of approximately 2%. However, in stage 3 there is a substantial jump in the use of subjunctives to almost 10%, which reaches 12% at stage 4. Figure 2 shows the rate of subjunctives across the 4 stages and in the adult grammar.

[Insert Figure 2 here]

In addition to the increase in frequency of subjunctives at stage 3, there is a marked increase in the number of subjunctive verbs, as well as the contexts in which subjunctive is used. Table 9 shows that during stages 1 and 2, subjunctive morphology appeared on only 8 and 7 different verbs, respectively. In stage 3, however, the subjunctive was used with 29 different verb types.

[Insert Table 9 here]

As also shown in table 9, the range of irrealis meanings that children expressed increased from stage 2 to stage 3: in stages 1 and 2, children expressed only 3 of the five possible irrealis meanings – desire, request and suggestion, whereas in stage 3 all of the possible 5 meanings emerged - desire, possibility, request, permission and suggestion. As we will see shortly, the verbs expressing irrealis meaning are almost always correctly inflected with subjunctive morphology.

It is sometimes proposed that children acquire morphology in an item-based manner (e.g., Tomasello, 2000; Pine, Lieven & Rowland, 1998). The thrust of this approach is that children acquire whole segments of language as unanalyzed chunks, and slowly decompose those chunks

into specific constructions. According to this view, inflectional morphology is not acquired within a system of grammar, but is associated with specific lexical items. We have proposed that subjunctive (irrealis) morphology is productive at stage 3. In order to show that Swahili-speaking children do indeed have productive control of morphology at this stage and rule out the possibility that they simply use a set of unanalyzed verb forms, we did an item analysis of the verbs at this stage. If children use certain verb roots only with indicative mood vowel and other verb roots only with subjunctive mood vowel, that would be consistent with an item-based analysis in which the repertoire of unanalyzed learned forms has simply increased in stages 3 and 4.<sup>9</sup>

We investigated this possibility by inspecting the verb roots on which subjunctive morphology occurs. We first listed all cases of subjunctive verbs, and then within the transcript in which each verb occurred, we searched for utterances that contained that same verb root. On the whole, we found that verb roots that occur in the subjunctive also occur in the indicative or the imperative (which carries the default indicative final vowel) within the same transcript. For example, in stage 3, there are a total of 34 verbs that occur with the subjunctive (see table 10 below). Of these 34 verbs, 23 occur in both the subjunctive and the indicative moods within the same transcript (labeled ‘Productive’ in tables 10 and 11). Of the remaining 11 verb roots, 6 occur in the subjunctive in one transcript and in the indicative in a subsequent transcript (labeled ‘Likely Productive’ in tables 10 and 11). The remaining 5 were either solitary subjunctive tokens, or occurred repetitively without change across all contexts (labeled ‘Unlikely’ in tables 10 and 11). An example of this final category of verbs (which we consider truly formulaic and unproductive) is *nipe* (SA<sub>1s</sub>-give-SUBJ) ‘give me (it)’. This utterance was used repetitively across many contexts, much like the English *wanna* or *gimme*.

An analysis of individual children within stage 3 shows that both children in this stage are productive in both their use of verb roots and mood morphology. Of the 23 Productive verb stems in table 10, Mus produced 19 and Fau produced 21 (with overlap of the more common verb stems). Of the Likely productive verb stems, Mus produced 4 verbs stems and Fau produced 3, and of the Unproductive verb stems, Mus produced 2 and Fau produced 3.

In stage 4 (data of which comes from a single child, Has, shown in table 11), a total of 21 different verb roots occur with the subjunctive vowel, of which 17 occur with the indicative vowel within the same transcript (and are thus Productive). The remaining 4 verb roots occur in both moods in consecutive transcripts (Likely productive), with no verbs appearing only once or in formulaic contexts (Unproductive). These data are presented in tables 10 and 11 below

[Insert Table 10 here]

[Insert Table 11 here]

It is possible that item-based learning is only evident at very early stages in acquisition, e.g. our stages 1 and 2.<sup>10</sup> Our goal in this section is to rule out the possibility that the productivity observed in stages 3 and 4 is due to some lexically specific associations between certain verb roots and particular mood morphology. The question of whether children's very earliest subjunctive verbs (i.e. in stages 1 and 2) are acquired on an item-by-item basis is not at issue here. Nevertheless, in the interest of completeness, we conducted an item-analysis on stages 1 and 2. Our results reveal that 6 of the 9 verbs in stage 1 that occur in the subjunctive also occur with indicative morphology in the same transcript. The remaining 3 verbs are the only tokens of

these verb roots throughout the entire corpus. In stage 2, 7 of the subjunctive verbs occur within the same transcript as an indicative-marked variant of that verb. Of the remaining two tokens, one was the only token of its verb-type in the corpus, and the other occurred several transcripts after an indicative variant. Thus, even at the earliest stages of development we find no support for item-based learning. However, given the very limited amount of data in stages 1 and 2 we cannot draw any further conclusions.

So we conclude that mood oppositions become productive at stage 3, precisely the point at which a tense specification becomes obligatory. Thus, stage 3 appears to be a crucial turning point in the development of Swahili. We propose that at this point the subjunctive morpheme with its irrealis feature is available to license MoodP. In the absence of a mood specification, TP is active, hence the obligatoriness of tense at this point.

The development that takes place at stage 3 in Swahili is analogous to the “modal shift” described by Blom & Wijnen (2000) for child Dutch (see note 4). Dutch children’s first verbs are always infinitival. These early RIs occur in both irrealis and realis contexts. In a subsequent stage, when finite verbs appear, RIs occur predominantly in modal contexts while the finite verbs are restricted to temporal contexts. Blom and Wijnen argue that at the first stage Dutch children have not yet analyzed the infinitival morphology of the RI (there is no contrasting finite form), and hence, in our terms, the infinitival morpheme is not available to license MoodP. Thus, the effects of the SOH show up only when the nonfinite form is analyzed by the child. This is the case for the Dutch RI, as described by Blom and Wijnen, and we propose it is also true for the Swahili subjunctive.

The development of mood oppositions in the various languages discussed and especially the parallels between Dutch and Swahili, two languages that are quite different in terms of their



surface morphology, provide support for the general hypothesis that there are universal principles mediating the development of semantics and morphosyntax, and specifically the SOH.

In the next section we discuss a final observation concerning the development of the Swahili mood system, which is that when children begin to use the subjunctive they overregularize the semantics -morphosyntax mapping in a way which points to the emergence of an unmarked realis-irrealis split, even though the Swahili system is marked in certain respects. This again argues for the presence of universal principles in the acquisition of temporal and modal categories.

#### **4.1 Subjunctive and indicative morphology**

Much crosslinguistic work has been done investigating the semantic contexts of subjunctive forms. Subjunctive is the morphology that is often associated with irrealis mood (Bybee, Perkins, and Pagliuca 1994; Chafe 1995; Givón 1994). Givón, for example, shows that irrealis morphology generally occurs with suggestions, the expression of desires, making of requests etc. Realis or temporal forms, on the other hand, occur in past and present contexts. Chafe (1995) describes these different semantic contexts as occurring in a “gradient” of markedness, as in figure 3, with the unmarked realis contexts and irrealis contexts on opposing ends of a continuum. In the unmarked case, imperatives pattern as realis forms and have indicative morphology, while future and intentionals pattern like irrealis forms and have subjunctive morphology. However, the morphology of these three categories – imperatives, futures and intentionals – is subject to cross-linguistic variation (see also Giorgi & Pianesi, 1997, for a discussion of the relation between irrealis and the subjunctive form).

[Insert Figure 3 here]

Swahili adheres to the unmarked realis-irrealis distinction with the exception of future and intentionals, which have indicative morphology and hence represent a marked option. Therefore, the Swahili system is a marked mood system in that it departs from the prototypical realis/irrealis mapping. The Swahili mood system is summarized in table 12. (The bold face indicates that the assignment of future and intentional to the indicative category is a marked option:

[Insert Table 12 here]

#### **4.2 Subjunctive and indicative morphology in early Swahili**

The final mood vowel is obligatory in Swahili and children never omit this vowel. This could be due simply to phonotactic requirements, as all words in Swahili end in a vowel. If the child were inserting a default vowel at the end of a verb in order to satisfy phonotactic requirements, we would not expect to find consistent form-meaning relationships. In other words, we would not expect children to respect the semantic restrictions on the use of particular mood vowels, but instead would expect a significant rate of errors. In order to test this, we investigated the form-meaning contingencies in the children's indicative and subjunctive clauses.

To determine the meanings of children's utterances, we considered the previous and following utterances of the child, as well as the parents' utterances, intonation, hand gestures and eye-gaze (if available from contextual notes) and in some cases follow-up questions with the parents. Examples (17) and (18) below show a child's utterance that was coded as 'desire'. In example (17) the previous utterance served as the major clue as to the intention of the child. In

example (18), both the previous as well as the following adult utterance indicate that the child's intention is to express a desire.

(17) \*ADU: u – na – tak – a kw –end – a wapi?  
 SA<sub>2s</sub>–pres–want–IND INF– go–IND where  
 ‘Where do you want to go?’

\*CHI: mw – on – a Haw04, line 1872  
 OA<sub>3s</sub> – see – IND  
 ‘See him/her.’

\*ADU: ku – mw – on – a nani ?  
 INF – OA<sub>3s</sub>–see–IND who  
 ‘To see whom?’

(18) \*ADU: u – na – tak – a ku – shuk – a ?  
 SA<sub>2s</sub>–pres–want–IND INF– alight–IND  
 ‘You want to get down?’

\*CHI: ruk – e Mus19, line 804  
 jump – SUBJ  
 ‘Jump (down).’

\*ADU: u – a – tak – a ku – shuk – a u – n – end – a wapi ?  
 SA<sub>2s</sub>–pres–want–IND INF– alight–IND SA<sub>2s</sub>–pres–go–IND where  
 ‘You want to get down and go where?’

Example (19) below shows a child utterance that was coded as a ‘suggestion.’ Contextual notes indicate that the child made a hand gesture indicating that she and the adult go together, which indicates a suggestion. Additionally, intonation and the following adult utterance were useful in making this determination. Example (20) shows request by a child (marked by the subjunctive vowel), which differs from an imperative in intonation and the final vowel. The intonation indicates a request (as opposed to an order or a question/suggestion), and the response by the adult is appropriate to a request.

(19) \*CHI: tu – end – e ... mw – on – a Wali Haw05, line 246  
 SA<sub>1pl</sub>–go–SUBJ OA<sub>3s</sub>–see–IND Wali  
 ‘Let’s go...see Wali.’

- \*ADU: mm? Haya, tu – end – e.  
ok SA<sub>1pl</sub>–go–SUBJ  
'Mm? Ok, let's go.'
- (20) \*CHI: Charlie
- \*ADU: hmm ?
- \*CHI: wek – e kwa hile duo yako Has02, line 11  
put–SUBJ in that drawer yours  
'(Please) put (it) in that drawer of yours.'
- \*ADU: Ni – wek – e wapi?  
SA<sub>1s</sub>–put – SUBJ where  
'Where should I put it?'

Determining the intention of the child is a difficult task, and care was taken to ensure accuracy. As mentioned earlier, the codes were checked by a native speaker assistant, which did result in some discrepancies in coding. All such cases that were unclear were discarded from the counts.

Our initial analysis was of all indicative verbs, and the results are presented in Table 13.

[Insert Table 13 here]

Ignoring the unclear cases, we have 1436 indicative clauses. Of these, a total of 23 are used incorrectly to express desires or make suggestions, etc. (The errors are bold faced in the table.) This represents an error rate of 1.6%. The remaining 98.4% of indicative clauses are used correctly. Note that included in this count are all indicative verbs (excluding imperatives), whether tensed or tenseless. However, it is possible that the use of tense forces an indicative interpretation, and so a second analysis looked at the interpretation of bare indicative stems,

which you will recall, are not marked for tense. The results (excluding unclear cases) are presented in table 14.

[Insert Table 14 here]

We have 164 clauses containing indicative bare stems, of which 5 are used incorrectly to express desire, suggestion or request. This is an error rate of 3.1%. The remaining 96.9% of indicative bare stems are used correctly.

One noteworthy result is that the overwhelming majority of bare forms occur in temporal (i.e. non-modal) contexts. As we mentioned earlier, this argues against the hypothesis in (8a), that the default interpretation of non-finite forms is modal. It also argues against Blom, Krikhaar, and Wijnen's (2001) hypothesis that the 'default' form has both a modal as well as temporal meaning. Rather, the Swahili bare forms pattern with the English bare form in being non-modal. Our hypothesis is that the Swahili bare verb lacks the irrealis feature to license Mood, as in English.

Blom & Krikhaar (2001) and Blom (2003) also claim that a significant factor determining the lack of irrealis bare verbs in English is that only 3<sup>rd</sup> person contexts are considered. This is because in English a present tense specification is only visible on 3<sup>rd</sup> person. They claim that the preponderance of RIs in Dutch occur in first and second person. However, the Swahili results in table 14 include verbs of all grammatical persons, thus showing that in Swahili the bare stem has temporal meaning irrespective of person.

Blom (2003) also suggests that bare verbs in early English may result from simple (phonological) omission of *-s* and is therefore really a hidden finite form. As such, it would not

have modal meaning but would have the temporal meaning associated with finite verbs. A reviewer notes that a similar analysis may hold of Swahili, that is to say, the bare verb may be a hidden finite verb and hence have temporal rather than modal meaning for that reason. Torrence and Hyams (2003) show that in child English the bare verb typically denotes ongoing events (e.g. Mommy dance = Mommy is dancing) while the finite *-s* verbs have a habitual meaning (as is correct in the adult language). It is therefore unlikely that the finite verb is the source for the bare verb in English. With respect to the Swahili data we are considering, we do not know at this point whether a similar difference in meaning exists between the bare verb and the inflected verb. However, Deen (2003a) shows that the rate of overt subjects in bare stem contexts is significantly lower than the rate of overt subjects in full clause contexts because of the different structural properties associated with the omission of Subject Agreement and Tense (e.g., case marking properties). Thus it is unlikely that the Swahili bare verb results simply from the phonological omission of finite prefixes (see Deen 2002; 2005 for more details).

Turning now to the subjunctive clauses, as shown in table 15, of the 105 subjunctive forms, there are a total of 18 errors (bold face numbers), that is, 18 instances in which subjunctive appears on verbs that would be indicative in the adult language. This is an error rate of 17%. Importantly, however, of these 18 cases 10 occur in future or intentional contexts (the bold face numbers with asterisks).

[Insert Table 15 here]

Recall that in the crosslinguistically unmarked (or prototypical) case, future and intentional contexts are irrealis. According to Chafe and Givón, adult Swahili future/intentional

contexts are indicative and are hence marked. By (over-)extending subjunctive morphology to the future and intentional contexts, the Swahili children are using the indicative-subjunctive morphology according to the unmarked mapping. By hypothesis the Swahili children have not acquired the marked characteristic of Swahili according to which future and intentionals have indicative morphology. If we put aside these 10 cases, as representing a principled departure from the adult grammar, the number of subjunctive errors drops to 8 – that is, only 7.6%. The corrected form-meaning contingencies are presented in table 16.

[Insert Table 16 here]

Thus, there are 3 sources of evidence pointing to stage 3 as the point at which mood morphology is analyzed by the child. First, the number of subjunctive verbs increases; second, the variety of subjunctive verbs and irrealis meanings increases; and third, Swahili children mark irrealis according to the unmarked mapping shown in figure 3, given by UG. It is not obvious how statistical learning models or item-based learning can account for such overextensions.

Summing up, we have thus identified two significant developments at stage 3: The specification of tense becomes obligatory, as shown in figure 1. Second, mood oppositions become productive, as just discussed. We thus propose that stage 3 is the point at which the realis-irrealis opposition maps onto Swahili morphosyntax, in accordance with the SOH.

## **5.0 Leaving the RI/RI analogue stage: the shift to independent MoodP and TP**

We have seen this complementarity between Mood and Tense is not an option reserved for child grammar. As we have discussed, in adult Swahili there is no tense specification in

subjunctive clauses. The same is true for various structures in other adult languages, for example English modals, as discussed in section 2.1. Thus, the complementarity that we see in early grammar between temporally anchored finite forms and non-finite irrealis forms is based on more general principles that hold in adult languages as well. This is in accordance with the Continuity Hypothesis that holds that child grammars do not fall outside the grammatical range permitted by UG (White 1982; Hyams 1983).

There are, however, adult languages in which Tense and Mood do not form a unitary category, for example, Dutch and Italian and Greek. Children acquiring these languages must eventually learn that Mood and Tense are independent. What is required, therefore, is positive evidence of separate functional heads. This can be provided by morphology – forms that carry both tense and mood morphemes. As discussed earlier, in Italian there are past subjunctive forms, as well as past tense modals (cf. section 2.1). These forms tell the child that there must be two separate inflectional heads to check the two sets of features. Similarly, in Dutch (and other RI languages) modals have a past tense form. In Greek, finite verbs occur under the modal particle *na*, for example *na εφvyε* (*na* left-3rd sing ‘Can it be that she left?’), again providing evidence for distinct heads for Mood and Tense. If these modal forms serve as the trigger for the splitting of MoodP and TP, then we expect that children will exit the RI (or RI analogue) stage as they acquire the modals. As discussed in Hyams (2002; 2003) and Varlokosta et al (1998) this is the case. In Dutch child language the proportion of RIs decrease as the proportion of modals increases (cf. Blom 2003) and in early Greek the bare perfective decreases as the proportion of *na* structures increase.



In Swahili, a language without modals, no such evidence is forthcoming and thus the Swahili speaking child will stay with a unitary category and the complementarity of mood and tense remains as per the SOH (cf. 8).

## **6.0 Concluding Remarks**

In this paper we have argued for a universal principle that governs the mapping between semantics and morphosyntax in the early grammar, the semantic opposition hypothesis (SOH). The SOH states that the opposition between realis and irrealis mood is grammatically realized in an opposition between finite and non-finite morphosyntax. More precisely, it states that the expression of irrealis mood in the early grammar excludes a tense specification. We derived this generalization by proposing that in the early grammar (as in some adult grammars such as Swahili) Tense and Mood form a unitary category (Giorgi and Pianesi 1997). We proposed that the SOH gives rise to the RI stage in the Germanic languages, and we saw that in several other typologically distinct child languages -- Italian, Greek and Swahili -- other non-finite forms that map onto irrealis mood .

We paid particular attention to Swahili, a language in which mood excludes tense in the adult language as well. We showed that there is a point in the development of Swahili in which a clear semantic distinction between realis and irrealis mood emerges and that this semantic opposition is mapped onto appropriate finite/non-finite morphology, parallel to what is observed in the European languages.

We argued that the irrealis meaning of RIs does not arise as a default interpretation of non-finite forms since not all non-finite forms in early language occur in irrealis contexts, for example, English bare verbs, Italian bare participles, and Swahili bare verbs do not have modal meaning. Thus, the absence of a tense specification is a necessary but not sufficient condition for

the expression of irrealis mood. In addition, there must be an irrealis feature that licenses a Mood projection. Our crosslinguistic survey showed that children in fact avail themselves of the 3 verbal forms that have an irrealis feature – infinitives, imperatives and subjunctives, and that the choice of one or the other is determined by the morphosyntactic requirements of the target language.

On the basis of the languages that we have surveyed, including Swahili, a little studied language, we conclude that there are indeed universal principles that govern language development. The child's early convergence on the specific morphosyntax of the target language does not preclude commonalities in grammatical development. Our results suggest the universals of development may reside at the interface levels, between semantico-conceptual structure and morphosyntax, and possibly between the articulatory-perceptual interface and morphosyntax.

## References

- Ashton, Ethel Osteli. 1947. *Swahili Grammar*. London: Longmans, Green and Co.
- Barbiers, Sjef. 1995. *The Syntax of Interpretation*. HIL/Leiden University: PhD dissertation.
- Becker, Misha. 2000. *The Development of the Copula in Child English: The Lightness of be*.  
UCLA: Doctoral Dissertation.
- Becker, Misha, and Nina Hyams. 1999. Modal reference in children's root infinitives.  
*Proceedings of the 30<sup>th</sup> Child Language Research Forum*. ed. Eve Clark. Stanford: CSLI.
- Belletti, Adriana. 1990. *Generalized verb movement*. Turin : Rosenberg and Selier.
- Blom, Elma, and Evelien Krikhaar. 2001. Where to use a nonfinite clause? An experiment on the use of verb forms in early Dutch and English child language. *Proceedings of GALA 2001* eds. João Costa, Maria João Freitas (Eds). Palmela, Portugal.

- Blom, Elma, E. Krikhaar, and Frank Wijnen. 2001. Root infinitives in Dutch and English child language: an experimental approach. *Proceedings of BUCLD 25*. Somerville: Cascadilla Press.
- Blom, Elma, and Frank Wijnen. 2000. How Dutch children's root infinitives become modal. *Proceeding of BUCLD 24*. Somerville: Cascadilla Press.
- Blom, Elma. 2003. *From Root Infinitive to Finite Sentence*, University of Utrecht/LOT: PhD dissertation.
- Boboljik, Jonathan, and Thrainsson, Hoskuldur. 1998. Two heads aren't always better than one. *Syntax* 1:37-71.
- Bolinger, Dwight. 1968. Entailment and the meaning of structures. *Glossa* 2:118-127.
- Boser, Katherine, Lust, Barbara, Santelmann, Lynn, and Whitman, John. 1992. The syntax of CP and V2 in early child German. *Proceedings of NELS* 23:51-63.
- Bottari, Piero, Paola Cipriani, and Anna Maria Chilosi. 1993/1994. Protosyntactic devices in the acquisition of Italian free morphology. *Language Acquisition* 3 (4):327-369.
- Brown, R. (1973) *A First Language: The Early Stages*, Cambridge: Harvard University Press.
- Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Chafe, Wallace. 1995. The realis-irrealis distinction in Caddo, the Northern Iroquoian Languages, and English. In *Modality in Grammar and Discourse*, edited by J. Bybee and S. Fleischman. Philadelphia, PA: John Benjamins Publishing Company.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.

- Chomsky, Noam. 1993. The minimalist program. In *The view from building 20*, eds. Ken Hale and Jay Keyser. Cambridge: MIT Press.
- Cinque, Guglielmo. 1999. *Adverbs and Functional heads*. New York: Oxford University Press.
- Clahsen, Harold, Eisenbeiss, Sonja, and Penke, Martina. 1996. Lexical learning in early syntactic development. In H. Clahsen, ed. *Generative Perspectives on Language Acquisition*. Amsterdam: J. Benjamins.
- Deen, Kamil Ud. 1997. The interpretation of root infinitives in English: is eventivity a factor? *Unpublished UCLA Manuscript*.
- Deen, Kamil Ud. 2001. The acquisition of Swahili verbal morphology. Proceedings of the GALA 2001. eds. João Costa, Maria João Freitas (Eds). Palmela, Portugal, pp.41-48.
- Deen, Kamil Ud. 2002. *The Acquisition of Nairobi Swahili: The Morphosyntax of Inflectional Prefixes and Subjects*. Unpublished Doctoral Dissertation, UCLA.
- Deen, Kamil Ud. 2003a. Underspecified verb forms and subject omission in Nairobi Swahili. In Beachley, Barbara, Amanda Brown & Frances Conlin (Eds), *BUCLD 27 Proceedings*. Somerville, MA: Cascadilla Press, 220-231.
- Deen, Kamil Ud. 2003b. The Acquisition of Inflectional Prefixes in Nairobi Swahili. *Annual Review of Language Acquisition*, 3, 139-179.
- Deen, Kamil Ud. 2004. Object Agreement and Specificity in Swahili. In A Brugos, L. Micciula and C.E. Smith (Eds) *BUCLD 28 Proceedings*. Somerville, MA: Cascadilla Press, 129-140.
- Deen, Kamil Ud. 2005. *The Acquisition of Swahili*. Amsterdam, Netherlands: John Benjamins Publishing.

- Demuth, Katherine, and J. Gruber. 1995. Constraining XP sequences. In *Niger-Congo Syntax and Semantics 6*. eds. by V. Manfredi and K. Reynolds. Boston: Boston University African Studies Center.
- Duffley, Patrick. 1992. *The English Infinitive*. London: Longman.
- Farkas, Donca. 1992. The semantics of subjunctive complements. In *Romance Languages and Modern Linguistic Theory*. eds. P. Hirschbuhler and K. Koerner. Amsterdam: John Benjamin.
- Ferdinand, Astrid. 1996. *The Acquisition of the Subject in French*, HIL/Leiden University:PhD dissertation.
- Giorgi, Alessandra, and Fabio Pianesi. 1997. *Tense and Aspect: from Semantics to Morphosyntax*. Oxford: Oxford University Press
- Givón, Talmy. 1994. Irrealis and the Subjunctive. *Studies in Language* 18 (2):265-337.
- Guasti, Teresa. 1993/4. Verb syntax in Italian child grammar: finite and nonfinite verbs. *Language Acquisition* 3:1-40.
- Haan, Ger J. de. 1987 A theory bound approach to the acquisition of verb placement. eds. Ger de Haan and Wim Zonneveld. *Formal Parameters of Generative Grammar* 3, 15-30. University of Utrecht.
- Halle, Morris & Marantz, Alec. 1993. "Distributed Morphology and the Pieces of Inflection," in K. Hale and S.J. Keyser, eds., *The View From Building 20*, Cambridge, Mass.: MIT Press, 111-176.
- Han, Chung-hye. 1998. *The Structure and Interpretation of Imperatives: Mood and Force in Universal Grammar*. University of Pennsylvania: PhD dissertation.
- Hoekstra, Teun, and Nina Hyams. 1998. Aspects of root infinitives. *Lingua* 106 (1-4): 81-112.

- Hoekstra, Teun, and Hyams, Nina. 1996. The syntax and interpretation of dropped categories in child language: a unified account. *Proceedings of WCCFL XIV*, Stanford: CSLI.
- Hoekstra, Teun and Jordens, Peter. 1994. From adjunct to head. In Teun Hoekstra and Bonnie Schwartz, eds. *Language Acquisition Studies in Generative Grammar*. Amsterdam: J. Benjamins
- Hyams, Nina 1983. *The Acquisition of Parameterized Grammars*. CUNY PhD dissertation.
- Hyams, Nina. 2001. Now you hear it, now you don't: the nature of optionality in child grammar *Proceedings of BUCLD 25*. Somerville: Cascadilla Press.
- Hyams, Nina. 2002. Clausal structure in early Greek: A reply to Varlokosta, Vainikka and Rohrbacher and a reanalysis. *The Linguistic Review* 19:225-269.
- Hyams, Nina 2004. Child non-finite clauses and the mood-aspect connection: evidence from child Greek. In Roumyana Slabakova and Paula Kempchinsky, eds., *Aspectual Inquiries*. Dordrecht: Kluwer.
- Ingram, David, and William Thompson. 1996. Early syntactic acquisition in German: evidence for the modal hypothesis. *Language* 72 (1):97-120.
- Johnson, Kyle. 1990. The syntax of inflectional paradigms. University of Wisconsin unpublished manuscript.
- Josefsson, Gunlog. 2002. The Use and Function of Nonfinite Root Clauses in Swedish Child Language. *Language Acquisition*, 10, 4, 273-320.
- LaBelle, Marie. 2000. Les infinitifs racines en langage enfantin. Unpublished manuscript.
- Lasser, Ingeborg. 1997. *Finiteness in Adult and Child German*. MPI Series in Psycholinguistics.
- Leonini, Chiara. 2001. Assessing the acquisition of Italian as an L1: A longitudinal study of a child of preschool age. Unpublished University of Siena manuscript.

- MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk. Third Edition.*  
Mahwah, NJ: Lawrence Erlbaum Associates.
- Madsen, Shannon, and Gilkerson, Jill. 1999. A pattern of differences between bare forms and root infinitives. UCLA unpublished manuscript.
- Newport, Elissa, Gleitman, Lila, and Gleitman, Henry. 1972. Mother please, I'd rather do it myself. In C. Snow and C. Ferguson, eds., *Talking to Children: Language input and acquisition.* Cambridge: Cambridge University Press.
- Ngonyani, Deogratias. 1996. *The Morphosyntax of Applicatives.* UCLA: PhD dissertation.
- Peters, Ann (2001). Filler syllables: what is their status in emerging grammars? *Journal of Child Language*, 28, 1, Feb, 229-289.
- Pine, Julian, Lieven, Elena & Rowland, Caroline. 1998. Comparing Different Models of the Development of the English Verb Category. *Linguistics*, 36, 4(356), 807-830.
- Plunkett, Kim, and Sven Stromqvist. 1990. The acquisition of Scandinavian languages. *Gothenberg papers in theoretical linguistics* 59.
- Pollock, Jean-Yves. 1989. Verb movement, UG and the structure of IP. *Linguistic Inquiry* 20:365-424.
- Portner, Paul. 1994. The semantics of mood, complementation and conversational force. Georgetown University unpublished manuscript.
- Pullum, Geoff and Wilson, Dierdre, 1977. Autonomous syntax and the analysis of auxiliaries. *Language* 53, 741-788.
- Rizzi, Luigi. 1994. Some notes on linguistic theory and language development: the case of root infinitives. *Language Acquisition* 3:371-393.
- Roberts, Ian. 1993. *Verbs and Diachronic Syntax.* Dordrecht: Kluwer.

- Sachs, Jacqueline. 1983. Talking about the there and then: The emergence of displaced reference in parent-child discourse. In K. Nelson, ed., *Children's Language* Hillsdale, NJ: Lawrence Erlbaum.
- Sano, Tetsuya, and Hyams, Nina. 1994. Agreement, finiteness, and development of null arguments. *Proceedings of NELS 24*:543-558.
- Schütze, Carson. 1997. INFL in Child and Adult Language: Agreement, Case and Licensing. MIT: PhD dissertation.
- Schütze, Carson, and Kenneth Wexler. 1996. Subject case licensing and English Root infinitives. *Proceedings of BUCLD 20*. Somerville, MA; Cascadilla Press.
- Salustri, Manola and Hyams, Nina. 2003. Is there an analogue to the RI stage in the null subject languages? *Proceedings of the BUCLD 26*. Somerville, MA: Cascadilla Press.
- Smith, Carlotta. 1980. The acquisition of time talk: relation between child and adult grammars. *Journal of Child Language* 7, 263-278.
- Stephany, Ursula. 1981. Verbal grammar in modern Greek early child language. In Phillip Dale and David Ingram, eds. *Child Language: An International Perspective*, Baltimore: University Park Press.
- Stephany, Ursula. 1986. Modality. In Paul Fletcher and Michael Garman, eds. *Language Acquisition*. Cambridge: Cambridge University Press.
- Stephany, Ursula. 1995. The acquisition of Greek. In Dan Slobin, ed. *The Cross Linguistic Study of Language Acquisition*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Stowell, Tim. 1981. *Origins of Phrase Structure*. Doctoral dissertation, MIT.
- Stowell, Tim. 1982. The tense of infinitives. *Linguistic Inquiry* 13:561-570.
- Suppes, Patrick. 1973. The semantics of children's language. *American Psychologist* 88:103-114.



- Tomasello, Michael. 2000. Do Young Children Have Adult Syntactic Competence? *Cognition*, 74, 3, Mar, 209-253
- Torrence, Harold, and Hyams, Nina. 2003. The role of aspect in determining finiteness and temporal interpretation in early grammar. *Proceedings of GALA 2003*. University of Utrecht.
- Tsimpli, Ianthi. 1992. *Functional Categories and Maturation: The Prefunctional Stage of Language Acquisition*. University College: PhD dissertation.
- Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40.
- Varlokosta, Spyridoula, Vainikka, Anne, and Rohrbacher, Bernhard. 1998. Functional projections, markedness, and 'root infinitives' in early child Greek. *The Linguistic Review* 15:187-207.
- Vitale, Anthony. 1985. Aspects of Kiswahili stress. *Kiswahili: Journal of the Institute of Kiswahili Research* 52:106-113.
- Wexler, Ken. 1994. Optional infinitives, head movement and the economy of derivations. In D. Lightfoot and N. Hornstein, eds. *Verb Movement*, Cambridge:Cambridge university Press.
- White, Lydia. 1982. *Grammatical Theory and Language Acquisition*. McGill University PhD dissertation [distributed by Indiana University Linguistics Club].
- Wijnen, Frank. 1997. Temporal reference and eventivity in root infinitives. *MIT Occasional Papers in Linguistics* (12):1-25.
- Zanuttini, Raffaella. 1997. *A Comparative Study of Romance Languages*. New York: Oxford University Press.

## Tables and Figures

Table 1a. Temporal/modal reference of RIs in 4 Dutch children

Child	Present	Future/Modal	Past	Total
Jos	49 (17%)	212 (74%)	25 (8%)	286
Mat	83 (11%)	603 (86%)	15 (2%)	701
Niek	15 (4%)	343 (94%)	5 (10%)	363
Pet	47 (8%)	467 (87%)	19 (3%)	533
	194 (10%)	1625 ( <b>86%</b> )	64 (3%)	1883

Table 1b. Temporal/modal reference finite verbs in 4 Dutch children

Child	Present	Future/Modal	Past	Total
Jos	105 (99%)	0	1 (1%)	106
Mat	274 (94%)	7 (2%)	9 (3%)	290
Niek	201 (93%)	13 (6%)	0	214
Pet	77 (87%)	1 (1%)	11 (12%)	89
	657 (93%)	21 ( <b>3%</b> )	21 (3%)	699

Table 2. Temporal /modal reference of bare forms in English

Child	Present	Future/Modal	Past	Total
Nina	54 (62%)	6 (7%)	27 (31%)	87
Naomi	16 (59%)	0	9 (33%)	27
Adam	142 (75%)	15 (7%)	32 (17%)	189
Eve	30 (39%)	19 (25%)	27 (36%)	76
Total	242 (64%)	40 (11%)	95 (25%)	377

Table 3a. Modal reference of perfective and imperfective verbs in 2 Greek children

Child	Perfective verbs	Imperfective verbs
Spiros	93% (69/74)	4% (3/73)
Janna	72% (143/200)	1% (2/136)
Total	<b>77%</b> (212/274)	2% (5/209)

Table 3b. Proportion of RI/finite verbs (Dutch) and perfective/imperfective verbs (Greek) showing modal reference (MRE).

<i>Dutch</i>		<i>Greek</i>	
RIs	Finite	Perfective	Imperfective
<b>86%</b> (1625/1883)	<b>3%</b> (21/694)	<b>77%</b> (212/274)	<b>2%</b> (5/209)

Table 4. Tense/Aspect prefixes in Swahili

Morpheme	Meaning	Morpheme	Meaning
Li	past perfective	Me	present perfect
Na	present on-going/habitual	Ki	habitual, conditional
Ta	Future	ng/nga	Hypothetical
Ka	Narrative	Ku	Infinitival

Table 5. Subject information

Child	Age range	No. of recordings	MLU
Haw	2;2 – 2;6	7	1.54–2.46
Mus	2;0 – 2;11	23	1.52–3.57
Fau	1;8 – 2;2	10	2.97–3.93
Has	2;10 – 3;1	5	3.15–4.23

Table 6. Criteria for assigning data to stages:

	MLU	Verbs per utterance	Proportion of PSDs
Stage 1	<2.5	<0.15	>25%
Stage 2	2.5-3.0	0.15-0.20	15%-24%
Stage 3	3.0-3.5	0.20-0.25	5%-14%
Stage 4	>3.5	>0.25	<5%

Table 7. Number of indicative, subjunctive, negative verbs in early Swahili.

Stage	Indicative	Subjunctive	Negative	Total
1	210	9	19	238
2	295	7	11	313
3	460	50	76	586
4	377	37	22	436

Table 8. Proportion of different clause types in stages 1 though 4

Stage	Full clause SA-T-V	[-SA] clause Ø-T-V	[-T] clause SA-Ø-V	Bare stem Ø-Ø-V	RI INF-V	Total
1	18% (39)	29% (60)	20% (42)	32% (67)	0.9% (2)	210
2	20% (58)	52% (154)	8% (25)	19% (55)	1% (3)	295
3	51% (235)	36% (166)	5% (21)	7% (34)	0.9% (4)	460
4	60% (225)	28% (104)	7% (26)	4% (15)	1.8% (7)	377

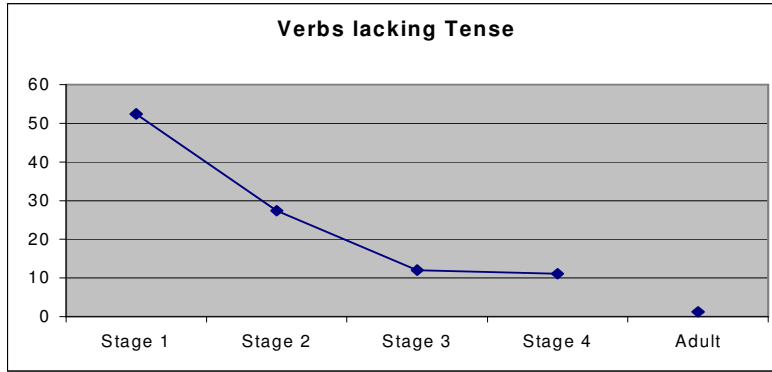


Figure 1. Proportion of tenseless clauses by stage

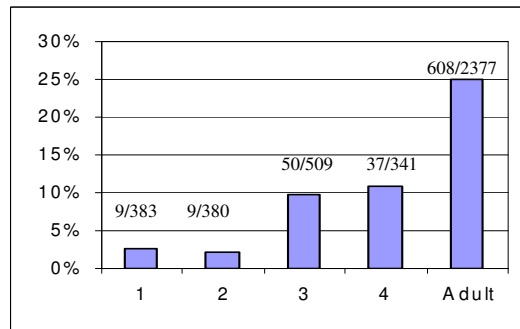


Figure 2. Subjunctives as a proportion of non-imperative verbs

Table 9. Types/tokens of verbs expressing irrealis mood and particular meanings expressed

	Stage 1	Stage 2	Stage 3	Stage 4
Types*	8 (4)	7	29 (23)	21
Tokens	9 (5)	9	50 (26)	37
Irrealis meanings	Desire Request Suggestion	Desire Request	Desire Possibility Request Permission Suggestion	Desire Possibility Request Permission Suggestion

\* Note: data in stage 1 comes from Mus and Haw, and data in 3 come from Mus and Fau. Mus' data cover stages 1 through 3, and are indicated in parentheses in stages 1 and 3.

Table 10. Productive, likely productive and unproductive verbs in stage 3

<b>Stage 3</b>		
<b>Productive</b>	<b>Likely</b>	<b>Unproductive</b>
Andika 'write'	Amuke 'awaken'	Mwashie 'disobey'
Angalia 'look'	Weke 'place'	Tengeze 'make proper'
Angusha 'throw down'	Pike 'cook'	Tengeneze 'make'
Chapa 'slap'	Ulize 'ask'	Ninamaze 'be quiet'
Chukua 'pick up'	Ruke 'jump'	Nipe 'give me'
Enda 'go'	Bebe 'carry'	
Finya 'pinch'		
Goja 'wait'		
Kata cut'		
Lala 'sleep'		
Leta 'bring'		
Mwambia 'tell'		
Ona 'see'		
Onesha 'show'		
Ongea 'talk'		
Pasulia 'tear for'		
Patia 'give to'		
Piga 'hit'		
Simama 'stand'		
Soma 'read'		
Towa 'remove'		
Vaa 'wear'		
Wacha 'leave'		
(total = 23)	(total = 6)	(total = 5)

Table 11. Productive, likely productive and unproductive verbs in stage 4

Stage 4		
Productive	Likely	Unproductive
Fungua 'open'	Andike 'write'	-none-
Piga 'hit'	Chukulie 'carry for'	
Towa 'remove'	Pengue 'blow the nose'	
Enda 'go'	Nyamaze 'be quiet'	
Sikiza 'listen'		
Kuja 'come'		
Chukua 'pick up'		
Sema 'say'		
Amuka 'awaken'		
Imbia 'sing for'		
Kimbia 'run'		
Weka 'place'		
Towa 'remove'		
Kula 'eat'		
Bandika 'attach'		
Angalia 'look'		
Peleka 'deliver'		
(total = 17)	(total = 4)	(total = 0)

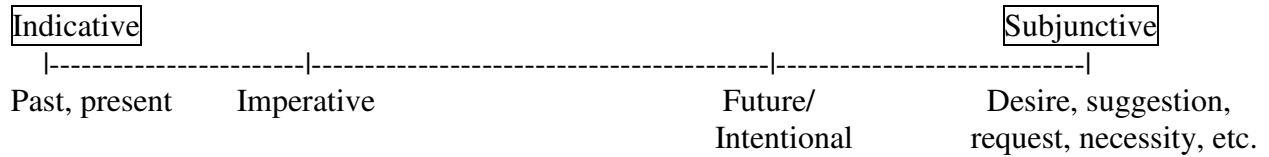


Figure 3. The unmarked 'gradient' of markedness.

Table 12. Summary of Swahili mood marking

Indicative/realis	Subjunctive/irrealis
Past	Desire
Present	Suggestion
Imperative	Necessity
<b>Future</b>	Request
<b>Intentional</b>	Permission
	Possibility

Table 13. Interpretation of child indicative verbs

	MEANING	
	Present, on-going	737
	Present result	160
Adult Indicative	Past	266
	Future	89
	Intentional	161
	Desire	<b>17</b>
	Possibility	0
Adult Subjunctive	Necessity	0
	Request	0
	Suggestion	<b>6</b>
	Total	1436

Table 14. Interpretation of child bare indicative stems

	MEANING	
	Present, on-going	92
	Present result	18
Adult Indicative	Past	25
	Future	6
	Intentional	18
	Desire	<b>3</b>
	Possibility	0
Adult Subjunctive	Necessity	0
	Request	<b>1</b>
	Suggestion	<b>1</b>
	Total	164

Table 15. Interpretation of child subjunctive verbs

	MEANING	
	Present, on-going	<b>2</b>
	Present result	<b>3</b>
Adult Indicative	Past	<b>3</b>
	Future	<b>1*</b>
	Intentional	<b>9*</b>
	Desire	18
	Possibility	2
Adult Subjunctive	Necessity	0
	Request	46
	Suggestion	21
	Total	105

Table 16. Form-meaning contingencies

	Indicative Marking	Subjunctive marking
Target Irrealis	5 (3.1%)	87 (83%)
Target Realis	159 (96.9%)	8 (7.6%)
	164	95

Notes:

<sup>1</sup> Here and throughout the paper we adopt the convention of referring to the semantic categories of mood and tense and mood and tense features using lower case letters, and the syntactic (functional) categories, Mood and Tense, by capitalizing the initial letter.

<sup>2</sup> As reported by Blom and Wijnen (2000), most of the non-modal RIs occur during a very early stage during which Dutch children typically produce only RIs and have no finite forms. At this stage children have no basis for analyzing the infinitive as such because it does not contrast with any finite forms and so they also have no basis for identifying the irrealis feature. Blom and Wijnen refer to the change in the two periods as the “modal shift”. In Blom’s (2003) reanalysis of the Dutch corpus data (based on 6 children including those in Wijnen’s study) she reports that the percentage of RIs with modal meaning after the modal shift is 80%, somewhat lower than what is originally reported by Wijnen (1997) and Blom and Wijnen (2000). This is in part due to the fact that several of Blom’s children are younger at the first recording, hence more of their data fall into the “pre-modal-shift stage.” Also, Blom excludes Niek from her analysis, the child with the highest percentage of modals RIs. This modal shift is discussed further in section 4.0, where we identify a similar phenomenon in child Swahili.

<sup>3</sup> There are various instances in which infinitives do not seem to have an irrealis meaning. For example, infinitives under raising verbs do not typically have a future/modal reading (e.g. *John seems to be happy/to have left early/to be dancing*), but show simultaneity with the matrix tense. We assume that raising verbs are functional verbs and that the [seem to V] functions as a complex verb whose tense and aspectual specification is determined by the matrix tense/aspect (Stowell 2001 class notes). Thus the infinitival morpheme (*to* in English) is not “visible” and there is no modal reading in this case. We assume along with Stowell and others that it is infinitival *to* that carries the irrealis feature.

A reviewer noted that non-finite verbs under perception verbs also fail to denote irrealis events, as in *I saw John walk* or the Dutch equivalent *Ik zie hem lopen*. Following the spirit of Stowell’s analysis for raising verbs, we propose that the small clause under perception verbs is reanalyzed as a complex predicate making the infinitive morpheme (in Dutch) invisible. In the English small clause (under perception verbs) there is no infinitival *to*, hence no source for the irrealis feature in any case. Similar considerations apply to Dutch and German examples (pointed out by a reviewer) such as *Guck mal, Fuesse washen* ‘Look now we wash your feet’, or *Hij toen snikkend naar zijn moeder lopen* ‘then he ran to his mother, sobbing’ in which the embedded infinitive has the properties of a small clause rather than a true infinitival clause.

<sup>4</sup> A reviewer notes that the English semi-auxiliaries *have to/has/had to* and *be able to/am able to/was able to* etc. express modality and are inflected for tense, in apparent contradiction to the proposal in the text. It should be noted that semi-auxiliaries are raising predicates (cf. *It has to rain (or my garden will die)*, *Tabs were able to be kept on the terrorists*) and hence have a bi-clausal structure: [It<sub>i</sub> has [t<sub>i</sub> to rain]] (cf. Pullum & Wilson 1977). Thus, we may assume the higher verb checks Tense while the embedded *to* infinitive checks Mood.

<sup>5</sup> Our proposal is reminiscent of Clahsen et al’s lexical learning hypothesis (LLH) mentioned in section 2 in that we assume that the child begins with a flatter structure than the adult (in some languages). It differs from the LLH in important respects: (i) we do not assume that children lack Agr at any point in development; (ii) structure is not added on the basis of the learning of the inflectional *paradigms*, a claim that is crucial to the LLH; (iii) the LLH hypothesis makes no predictions concerning the complementarity of tense and modality, and finally; (iv) our analysis predicts the optionality of RIs and finite verbs while Clahsen et al’s hypothesis predicts that children will exit the RI stage once agreeing forms are productive, contrary to fact.

<sup>6</sup> This is the structure that we assume for subjunctive and indicative clauses when no other intervening material is present. However, when Object Agreement is projected, it is projected in its universally determined structural position between MoodP and TP (see Giorgi & Pianesi, 1997; Cinque, 1999). The presence of OA causes the unitary option to fail, and MoodP and TP must therefore split, an option provided for by Giorgi and Pianesi who claim that “Scattering (i.e. splitting)...is allowed only when an extra head is needed to provide a *spec* position for



---

some feature (or feature bundle) which must merge there. If no such additional position is needed, scattering...cannot take place” (p. 231-232). In Swahili therefore MoodP and TP are split when OA features are specified, but unitary in all other cases. Positive evidence for this option occurs in the form of differential placement of mood and tense morphemes with respect to OA, and thus should not pose a problem for the child (see section 4.3). The little evidence that there is in the Swahili corpus points to the conclusion that OA begins to be used with subjunctive mood beginning in stage 3 (see Deen 2004 for more on the acquisition of OA in Swahili), suggesting that the children have acquired the unitary/split properties of TP and MoodP. Further discussion of this issue is beyond the scope of this paper.

<sup>7</sup> Note that by the same reasoning the English bare verb is not an RI-analogue in the sense intended in this paper (a non-finite form with irrealis mood). This departs from the assumptions of Wexler (1994), who takes the English bare verb (e.g. *Eve sit floor*) to be an RI. As far as we can tell, this issue is purely terminological, since we agree with Wexler that the English bare verb is non-finite. Hyams (2001) suggests that the English RI analogue may be the semi-auxiliaries such as *hafta*, *wanna* and *gonna*, which incorporate infinitival *to*, the source of the modal meaning.. However, this is speculative at this point. It is also possible that the English RI analogue is the imperative. Imperatives have not been studied in the context of early non-finiteness since there is no morphology to distinguish the imperative from the 2<sup>nd</sup> person indicative. This remains an area for future work.

<sup>8</sup> Given that the imperative is the RI-analogue in Italian, a reviewer asks if it could also be a plausible candidate in Swahili. In Swahili the imperative is morphologically a bare stem, that is, it does not take subject agreement or tense. It occurs with the default [a] final vowel, which is homophonous with the indicative. In early stages of child Swahili the imperative occurs at high rates: approximately 18%-20%. However, by stage 4, that rate diminishes to approximately 7%. The imperative may thus be an early RI analogue in Swahili, with the subjunctive partially replacing the imperative at later stages as the primary means of the expression of irrealis modality. We have more to say on the subjunctive in this paper because of the additional facts described in sections 4.1 and 4.2. A fuller analysis of the trade-off between imperatives and subjunctives is necessary, but space limitations prevent us from pursuing this here.

<sup>9</sup> In a similar vein, Blom (2003) argues that a lack of significant lexical overlap between finite and non-finite forms (e.g. *doe/doen* ‘does/to do’) in the early stages of Dutch verb acquisition supports her hypothesis that RIs are a default morphological form used before the child knows finite morphology. A detailed review of Blom’s data is beyond the scope of this paper. We note, however, that the calculation of lexical overlap is not independent of other factors that were not considered by Blom. In particular, the expected rate of overlap is a function of the number of finite forms that occur (if there are few finite forms, there can not be much overlap) and the rate of overlap in the further developed grammar (there must be some baseline overlap rate to compare the early data to).

<sup>10</sup> Thanks to Ginny Gathercole for pointing this out to us.