A fresh look at root infinitives from a cross-linguistic perspective[1]

Tanja Kupisch and Esther Rinke
University of Calgary / University of Hamburg

Abstract

In this paper we examine the relation between the quantity and quality of the adult input to the child and the intensity of the root-infinitive stage in child language. We compare the languages English, French, German, Italian and Brazilian Portuguese and test whether children produce infinitives more extensively if the verb morphology of their target-languages is ambiguous with respect to the distinction between finite and nonfinite verb forms, or whether the token-frequency of nonfinite verbs in the Input is crucial. We conclude by proposing that the latter is not decisive. Rather, children seem to avoid the use of finite verb forms especially in languages whose verb paradigms are characterized by ambiguities. Root infinitives may thus be viewed as a temporary phenomenon in a phase during which children are learning the inflectional properties of their target language.

1. Introduction

In many languages, children go through a stage during which they produce bare, uninflected infinitives instead of finite, inflected verbs, as in (1).

(1) a. En. Paul eat.

  b. Fr. Paul manger.

  c. Ge. Paul essen.

The question what these infinitives stand for has always been discussed controversially. Possible answers are: finite, inflected verbs, such as (2a), periphrastic constructions with a modal and an infinitival complement, as in (2b), or periphrastic constructions with an auxiliary and an infinitival complement, like in (2c).

(2) a. Paul eats.

  b. Paul wants to eat.

  c. Paul will eat.

In acquisition research in the generative tradition, Wexler (1994) has referred to this phenomenon as the *Optional Infinitive Stage*. This term cap-

---

tures the observation that children produce these infinitives inconsistently, in variation with finite, inflected verbs. Today, such infinitives are commonly referred to as *Root Infinitives* (RIs).

Different models have been proposed to account for this phenomenon. Radford (1988), Alridge (1989), Platzack (1992), Tsimpli (1992) and Vainikka (1993/94) related them to the existence of a prefunctional stage, i.e. a stage during which the child grammar lacks all functional categories. Rizzi (1993/94) suggested that children may sometimes truncate syntactic trees above the VP-level, so that the functional categories crucial for the production of finite verb morphology are lacking from the tree. Hoekstra & Hyams (1995) and Hoekstra, Hyams and Becker (1997) related RIs to the underspecification of the Number Phrase, while Wexler (1994) argued that RIs involve a Tense node that is not interpreted at LF. Clahsen et al. (1996) proposed that in order for functional projections to be instantiated, children have to acquire the morphological features associated with these positions. According to Philips (1995), Hoekstra and Hyams (1995), Hyams (1996), Hoekstra, Hyams and Becker (1997) and Avrutin (1999) children produce RIs because of non-syntactic processing or performance limitations, involving utterance planning and pragmatic skills.

The debate seems to have faded now. Common ground has been reached especially with respect to one empirical fact brought forward by these careful investigations, namely, that the extent to which children produce RIs varies across languages. In fact, RIs are virtually absent in *pro*-drop languages, such as Italian and Spanish, while they are frequent in German and French (Guasti 1993/94, 2002) (see, however, Rus 2007).

In this article, we will not discuss whether RIs are specified for Tense (Phillips 1995), underspecified for Tense (Hoekstra and Hyams 1995), or indicative of the absence of Tense (Rizzi 1993/94). Instead, we would like to elaborate on the aforementioned cross-linguistic differences, as we believe that they may shed light on the question why children produce RIs at all. More specifically, we investigate whether the *intensity* of the RI-stage can be predicted based on the frequency of infinitival constructions and/or existing ambiguities in the verb morphology in the children's input.

Interestingly, RIs have often been used as an argument against language acquisition models that take token frequency to be an important factor in acquisition. The idea would be that adults do not produce RIs since they are ungrammatical and consequently not part of the adult input to children. Nevertheless, some children ‑ depending on the target-language ‑ pass through a stage in which they produce RIs in abundance. There are at least two reasons why this view is too simplistic. First, the argument seems to be grounded on the assumption that children embark on

acquisition not only with an innately specified notion of finiteness but also with the ability to recognize finite verbs as such in the language they are exposed to, which may not be justified as it is not immediately apparent from the input whether a verb form is finite or not. Second, previous studies on adult data have shown that adults do produce RIs (Krämer 1993, Wijnen and Bol 1993, Schlichting 1996, Avrutin 1997, Kempen, Gillis and Wijnen 1997, Lasser 1997, Bohnacker 1999).

The article is structured as follows. Section 2 introduces our data and explains our methods. Section 3 presents an analysis of verb morphology in the speech of adults in Italian and Brazilian Portuguese, French, English and German. Section 4 looks at the amount of RIs in the speech of children acquiring these languages. Section 5 concludes the paper.

2. Data and method

2.1. Data

For the analysis of child language, we selected individual transcripts representing the language of children aged between 2;0 and 2;2. An overview of the corpora is presented in Table 1, indicating the mean age and the absolute number of verbs coded. We selected the transcripts randomly, using transcripts from eight different children per language, if available. We did not include these transcripts when they contained less than 5 lexical verbs. This made us exclude and replace one English and one German transcript of those we originally selected. We are aware of the fact that it cannot be granted that children at comparable ages have reached the same stage of linguistic development, but we assume that minor differences are out-leveled by selecting data from different children (unfortunately only one corpus is available for Brazilian Portuguese).

To examine the distribution of verb forms in the adult data we used individual transcripts of the same database.

Table 1: Corpora

language	corpus	no. of files analyzed	mean age	coded verbs	
				child	adult
Italian (IT)	Tonelli/ Antelmi/ Calambrone	8	2;1,20	692	933
Brazilian Portuguese (PTG)	Florianopolis	1	2;2,8	1254	1306
French	Champaud/ York/Pauline/Leveille/ Geneva/	6	2;1,23	692	1398
English	Manchester	8	2;1,29	753	1133
German	Szagun	8	2;1,12	1227	1099

2.2. Method

In the analysis of adult and child data we classified all verbs in terms of whether they were *unambiguously finite*, *nonfinite* or *ambiguous* between the two. Finiteness was classified by morphological marking only, not by word order. In the following we provide examples for each of these categories.

Within the category of *unambiguously finite verbs*, we distinguished between lexical verbs (e.g. *goes*), auxiliaries and modals (e.g. *has* and *wants*), copulas (e.g. *is*) and imperatives (e.g. It. *guarda*! 'look').

Within the category of *nonfinite verbs*, we distinguished between target-like and non-target-like forms. Target-like nonfinite verbs include complements of modal verbs (*can go, want to go*), complements of auxiliaries (*will go*), and complements of prepositions (It. *per fare* 'in order to make'). We considered nonfinite verbs to be non-target-like if they did not co-occur with finite verbs or prepositions that are normally considered to "license" them. Examples are bare participles (En. *fallen*), bare gerunds (It. *facendo* 'be in the process of making') and bare infinitives (Ge. *schnell machen)*. It is, of course, debatable whether all these forms are justifiably to be considered as non-target-like, given that adults also produce them. For instance, (as native speakers of German) we would consider Ge. *schnell machen* 'make fast' acceptable in an imperative context, but other native speakers might disagree. Since we cannot presuppose that children are sensitive to such subtle distinctions, and since we are mainly interested in finding out how many such forms children actually hear and how many they produce, we took the most "naïve" approach and classified such cases as non-target-like. We will use the terms *root infinitive* (RI) *(root participle, root gerund)* when referring to non-target-like nonfinite forms in child language.

Our third category is that of *ambiguous verb forms*. This category includes finite verb forms, such as e.g. German verbs in the 1st or 3rd person plural, illustrated in (3), and words other than verbs which are homophonous with infinitives, as illustrated in (4).

(3)   gehen / wir gehen / sie gehen
       to go / we go /   they go

(4)   essen (V) / das Essen (N)
       to eat   / the food

In the case of periphrastic verb constructions, we counted each verb separately. For example, in the following French utterance, we coded *peut* as an instance of a finite modal verb and *laisser* and *cuire* as instances of target-like nonfinite verbs.

(5)   on peut la laisser cuire
        we can it let   cook
        we can let it cook

In the child data, we also coded *copula omissions* provided that the utterances contained at least two syntactic arguments. That is, utterances consisting of a single DP, such as *my car*, were ignored but small clauses, such as *that_my car, that car_mine, this car_ green,* or *here_the truck* (see Radford 1988), were counted as copula omissions.[2]

In summary, our classification is based on the following distinctions (recall that copula omissions are only relevant for child data):

- unambiguously finite verbs
- nonfinite target-like verbs (licensed by preceding auxiliaries, modals or prepositions)
- nonfinite non-target-like verbs (gerunds, participles and infinitives which are "bare", i.e. not preceded by auxiliaries, modals or prepositions)
- morphologically ambiguous forms (finite verbs and categories other than verbs homophonous with infinitives)
- copula omissions.

We excluded incomprehensible speech, children's repetitions of adult utterances, adults' repetitions of child utterances, and forms that appeared to be rote-learnt, e.g. parts of songs and rhymes or frequent repetitions of one and the same verb form in the absence of any other verb.

3. The adult systems

3.1. Italian and Brazilian Portuguese

Italian exhibits a high amount of unambiguously finite verbs (72%). Nonfinite verbs (overall 27%) occurred in the following contexts: as infinitives after modal verbs (6a), after prepositions (6b) and after causative verbs (often in the imperative) (6c), as negative imperatives (6d), as participles after auxiliary verbs (6e) and after imperatives with *mettere* 'put' (6f). Contexts like (6c) and (6f) were less frequent. (The relevant part in the examples is marked in bold print).

---

[2] In some German corpora, we also found instances of reduced verb forms, containing less material than the stem, e.g. *pa* for *passen* 'fit' or *ha* for 'haben'. Although these forms might be considered nonfinite, we excluded them from the analysis, because they only occurred in German. Since their number was relatively low (28 tokens), their exclusion does not have a bearing on the overall results.

(6)   a.   no, non **puoi**   **toccare** niente
                no not can-you touch   anything
                no you can't touch anything

      b.   lo puoi   tenere in mano **senza**   **pigiare**
                it can-you hold   in hand without pushing
                you can hold it (in your hands) without pushing (it)

      c.   **fate**mi **vedere** anche a me le figurine!
                let-me see   also   to me the small figures
                let me also see the small figures

      d.   *non* **ti**   **sporcare** più!
                not yourself stain-Inf again
                don't stain yourself again

      e.   **non** ho   **capito**   che cosa volevi?
                not have-I understood what   wanted-you
                I don't understand what did you want?

      f.   vieni, **metti**ti **seduta!**
                come put-you seated
                come sit down!

Note that Italian has different verb endings for each person in the present tense (mang**io** 'I eat', mang**i** 'you eat', mang**ia** 's/he eats', mang**iamo** 'we eat', mang**iate** 'you eat', mang**iano** 'they eat') and in most other tenses.[3]

In Brazilian Portuguese, we found an even higher number of unambiguously finite verbs (81%). Nonfinite verbs (18% of the total) occurred as infinitives after modals (7a) and prepositions (7b), as simple (7c) or inflected infinitives (7d) after causative auxiliaries like *deixar* 'let' (often as imperatives), and as participles after the auxiliary *ter* 'have' (7e). Nonfinite verb forms also appeared as infinitives after *ir* 'go' (referring to the near Future, (7f)) and as gerunds after the auxiliary *estar* 'be' (7g).

(7)   a.   eu **quero ver**
                I want   see-Inf
                I want to see

      b.   é   **pra pôr**?
                is-it for   put
                Is it for putting (there)?

---

[3] For reasons of space, we cannot go into more detail here, and we restrict our remarks to verbs in the present tense (these are also most the frequent ones in adult-child conversation).

    c.   **deixa ver**   o   que você tem aí
        let   see-Inf that   what you have there
        let's see what you have there

    d.   **deixa** eu **ver**[4]
        let   I see-Inf
        let me see

    e.   você **tem brincado** com o   Jonathan?
        you have played   with the Jonathan
        did you play with Jonathan?

    f.   **vamos guardar** esse?
        go-we keep   this
        Are we going to keep this?

    g.   **está falando** muito baixinho
        is-he speaking very   quietly
        he is speaking very quietly

It should be noted that Brazilian Portuguese is in the process of losing its null subject property, although it still allows for null subjects in some contexts (Kato and Negrão 2000). That is, subject pronouns are realized in an increasing number of contexts where they would be omitted in a typical null subject language like Italian. The loss of the null subject property goes hand in hand with the reduction of verb endings. Brazilian Portuguese only has three distinct morphological forms: *(eu) falo* 'I speak', *(você) fala* 'you speak', *(ele/ela) fala* 's/he speaks', *(a gente) fala* 'we speak', *(vocês) falam* 'you (Pl.) speak', *(eles/elas) falam* 'they speak'. In contrast to peninsular Portuguese, the third person singular marker **-a** is also used to refer to the 2nd person singular and the 1st person plural. Thus, the reduction of morphologically distinct forms results in the ambiguity of *-a* with respect to person and number, but it does not imply an increase of ambiguous non-finite/finite forms. Such ambiguous forms are restricted to the 1st and 3rd person singular subjunctive future (*Futuro Imperfeito do Conjunctivo*) of regular verbs. In our data base, we found only 3 such cases (0.2%), see (8).

(8)   quando o papai chegar ele
        when the daddy arrives he
        when daddy comes, he ..

---

[4] Inflected infinitives are a particular type of infinitival construction in Portuguese (Raposo 1987). They take a nominative subject and agree with it in person and number, but they are unspecified for tense and cannot appear in isolation. In (7d), the inflected infinitive is morphologically equivalent to a non-inflected infinitive. It appears in the 1st person singular and agrees with the subject pronoun *eu* 'I'.

3.2. French

French exhibits nonfinite verbs in similar contexts as Italian and Brazilian Portuguese: after modals (9a), prepositions (9b), causative verbs (9c), and auxiliaries (9d). In addition, nonfinite verbs also occur with the *Future Proche* (10a), and with some fixed constructions, such as *venir voir quelque chose* 'come to see something' (10b), and *être en train de faire quelque chose* 'be about to do something' (10c). Finite verbs may also be followed by two infinitives (10d-e):

(9) a. tu **veux jouer** Marie?
         you want play   Mary
         do you want to play Mary?

     b. il ne   fait   pas assez   chaud **pour rester** comme ça.
         it NEG makes not enough warm to   stay   like   that
         it's not warm enough to stay like that

     c. tu   le   **fais   parler**
         you him make talk
         you make him talk

     d. tu   **as**   **mangé** de la viande?
         you have eaten   of the meat
         did you eat (some) meat?

(10) a. tu **vas manger** des   pommes de terre?
         you go eat   of-the potatoes
         will you eat potatoes?

     b. tu   **viens** le **voir**?
         you come it see
         will you come to see it?

     c. c'est là   qu'   il est **en train de manger**
         it's there where he is eating
         it's there where he's eating

     d. il   ne   **faut** rien   **faire tomber**
         one NEG must nothing make fall
         one shouldn't let anything fall

     e. tu   **veux essayer** de **mettre** la culotte de maman?
         you want try-Inf to put   the slip   of mum?
         do you want to try to put mum's slip on?

The participles and plural imperative forms of French verbs in *–er* are homophonous with infinitives. Although plural imperatives did not occur in our data, participles were frequent.

(11) a.   chanter / chanté / chantez!   [ʃãte]
            to sing / sung / sing-IMP

       b.   cacher / caché / cachez!   [kaʃe]
             to hide / hidden / hide-IMP

Moreover, finite verbs in the 1st, 2nd, 3rd person singular and 3rd person plural are homophonous:

(12) a.   je chante / tu chantes / elle/il chante / ils/elles chantent [ʃãt]
            I sing   / you sing / s/he sings   / they sing

    In our data base, French exhibits more contexts with nonfinite verbs than Italian and Brazilian Portuguese, and there is more homophony between finite and nonfinite verbs as well as among finite verbs.

3.3. English and German

Similar to Italian and French, German infinitives occur after modal verbs (13a) and after prepositions (13b). Infinitives also appear in the progressive construction with *am* 'at-the' (13c) and after the verbs *lassen* 'let' and *gehen* 'go' (13d,e). Participles occur after auxiliaries (13f).

(13) a.   der **will** den **fang'n**
            *he wants him **catch***
            he wants to catch him

       b.   und **zum reinmachen** muss man es auf den Tisch legen ne?
            and to   put-in   must one it onto the table lay   ey?
            and to out in inside one has to lay it onto the table, ey?

       c.   jetzt bist 'e   da … **am   wegwischen** hm?
            now are you there   at-the cleaning, hm?
            now you're cleaning, hm?

       d.   guck mal den **lässt** 'e   hier unten **häng'n**
            look   it   let-you here down hang
            look you let it hang down here

       e.   ja   dann *geh* du   mal *gucken* und dann kommst'e wieder
            well then go you just see   and then come-you back
            well then just go and see and then you come back

       f.   den **hast** du **weggelegt**
            it   have you put-away
            you put it away

German infinitives are further homophonous with finite verbs (14a), nomi-nalized verbs (14b) (see also (3-4)), and arguably also with imperatives ending in nasals, such as *komm*! 'come!'[5]

(14)   a.   wir/sie wollen ***trinken***
                we/they want to drink
      b.   das is dein ***Trinken*** ja?
                that's your drink, isn't it'

In addition, there are elliptical sentences, which resemble child RIs, as they lack the subject, sometimes also modals or auxiliaries, and the verb is nonfinite. They may express comments (a-d), advice (e,f), or imperatives (g-i). Given their resemblance with child RIs, we coded them as non target-like infinitives. Similar examples are discussed in Lasser (2002).

(15)   a.   schon **ausgetrunken**
                already drunk-out
                it's already finished

      b.   oh schon wieder **hingefall'n**
                oh once again fallen
                oh it's fallen again

      c.   **gucken** was Gisela macht oder was?
                look   what Gisela does or   what?
                do you want to see what Gisela is doing?

      d.   eh Strümpfe **anzieh'n** ne?
                eh socks   put-on ey?
                you're putting or socks on don't you?

      e.   so jetz hier **kurbeln**
                so now here crank
                now crank here!

      f.   ja   ein bisschen **dreh'n** und dann kriegt ma das raus
                yes a little   turn   and then get   one that out
                yes one has to turn that a little and then one gets it out

      g.   komm her die Nase eben **putzen**!
                come here the nose just clean!
                come here let's clean your nose

---

[5] As Ute Bohnacker correctly pointed out to us, forms like *komm!* 'come,' *sing*! 'sing!', *fang!* 'catch!' are distinguishable from the corresponding infinitives by consonantal lengthening on the latter and its absence on the former. We agree with her, but we are not sure whether this distinction is always perceived in spoken speech.

h.   nich **umkippen**!
        not   dump!
        don't dump it!
i.   aber vorsichtig **sein**!
        but   careful   be!
        but be careful!

Overall, constructions with infinitives are more frequent in German than in the other languages. In addition, some finite verb forms and nouns are homophonous with infinitives, and there are elliptical sentences containing bare participles and infinitives.

In English, infinitives occur with modal verbs (16a), and present and past participles occur after auxiliaries (16b-d).

(16)   a.   **shall we put** them away?
        b.   it is **raining**
        c.   he has **left** us some new bin bags.
        d.   they were **cleaned**

Imperatives (17a), the auxiliaries *have* and *do* (17b) (except in the 3rd person singular) and some nominalized verbs (17c) are homophonous with infinitives. Modal verbs, such as *will, shall, might* etc. are invariable.

(17)   a.   *Make* a fancy tunnel!
        b.   *Have* you heard it? *Do* you hear it?
        c.   That's an alternative *use* of my tunnel.

All verb forms in the Simple Present except the 3rd person singular are homophonous with the infinitive, e.g., I **go**, you **go**, we **go**, you **go**, they **go.**

Except in the Simple Past, English tenses are compositional and contain a nonfinite verb form, e.g., will *go*, have *gone*, had *gone*. Hence, as mentioned in previous work (e.g. Wexler 1994, Philips 1995: 334-335), English verb morphology is an unreliable indicator of finiteness.

In summary, we may observe that the quantity of nonfinite verbs (target- and non target-like) is highest in English (33.5%). French and Italian are similar (27.5% and 29%), as are German and Brazilian Portuguese (19% and 18%). As we shall argue, neither the token-frequency of target-like nonfinite verbs nor that of non-target-like nonfinite verbs can be decisive when it comes to the production of RIs in child language. Rather, the difference in the amount of verb forms which are ambiguous between finite and nonfinite appears to be crucial. Such forms are absent in Italian and Brazilian Portuguese, while constituting 5% in French, 7% in German, and 12.5% in English. Table 2 and Figure 1 summarize our results.

Table 2: Verb forms in the adult data

	categories	IT.	PTG.	FR.	EN.	GE.
unambiguously finite verbs	lexical verbs	304 *(33%)*	503 *(38%)*	435 *(31%)*	143 *(12.5%)*	321 *(29%)*
	auxiliaries/ modals	151 *(16%)*	303 *(23%)*	238 *(17%)*	224 *(20%)*	146 *(13%)*
	copula	113 *(12%)*	154 *(12%)*	200 *(14%)*	179 *(16%)*	210 *(19%)*
	imperatives	107 *(11%)*	105 *(8%)*	76 *(5.5%)*	62 *(5.5%)*	95 *(9%)*
target-like nonfinite verbs	infinitives	129 *(14%)*	174 *(13%)*	271 *(19.5%)*	245 *(21.5%)*	125 *(11%)*
	participles/ gerunds	119 *(13%)*	66 *(5%)*	110 *(8%)*	137 *(12%)*	50 *(5%)*
non-TG nonfinite verbs	infinitives	10 *(1%)*	17 *(1%)*	0 *(0%)*	0 *(0%)*	4 *(0%)*
	participles/ gerunds	10 *(1%)*	5 *(0%)*	2 *(0%)*	0 *(0%)*	33 *(3%)*
morphologically ambiguous forms		0 *(0%)*	3 *(0%)*	66 *(5%)*	143 *(12.5%)*	115 *(11%)*
total		943 *(100%)*	1330 *(100%)*	1398 *(100%)*	1133 *(100%)*	1099 *(100%)*

Figure 1: Distribution of finite and nonfinite verb forms in adult data
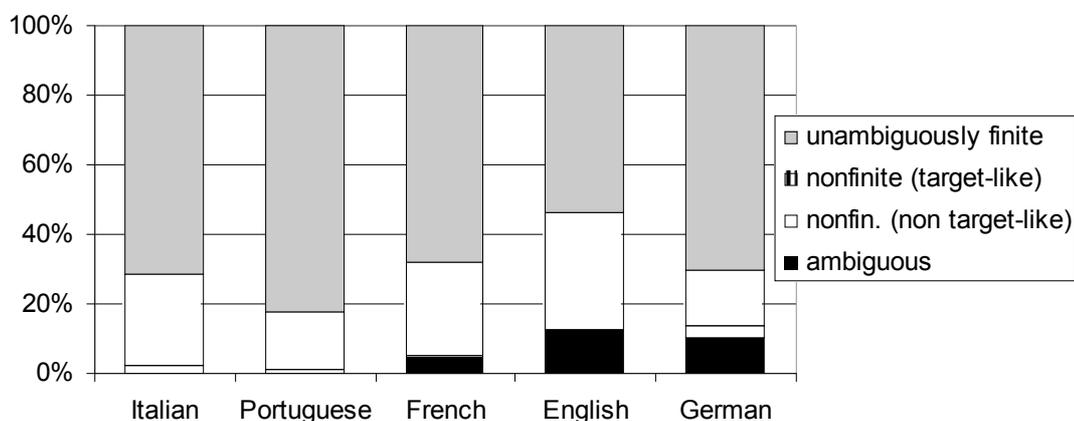
4. Infinitives in child language

4.1. Coding

In coding the child data, we encountered the following problems. Due to morphological ambiguity, it is impossible to determine whether English verbs are finite, unless they are marked for Past Tense or 3rd person

singular Present Tense. One possibility is to take the subject – if realized – as an indicator of the intended verb form. However, there are reasons against such a procedure. First, the data are representative of a phase for which subject drop is typical, as illustrated in (18):[6]

(18)   MOT: tell me what you want!
       GAI: play bricks

In other words, not all utterances actually contain a subject. Second, the realization of finiteness and the presence of a subject do not go hand in hand. Table 3 shows an analysis of all utterances containing a verb and a 3rd person singular subject (pronominal or lexical). If we look at combinations of lexical verbs and 3rd person singular subjects, there are more contexts lacking subject-verb agreement (verbs showing no AGR) than contexts showing subject agreement (verbs showing AGR). It is not always clear whether agreement errors are due to an incorrect pronoun or an incorrect/missing inflection. Moreover, most contexts with a 3rd person singular subject lack a verb altogether (no verb) or contain a copula verb (copula). For these reasons, we decided to consider verb inflection independently of pronoun use, and coded all verb forms which were marked neither as Past nor as 3rd person singular Present Tense as ambiguous regardless of the subject.

Table 3: Third person singular contexts, agreement and verb omission

	copula verbs	Vs showing AGR	Vs showing no AGR	no verb	total
occurrences	50	25	68	206	349
%	14.3%	7.2%	19.5%	59%	100%

Strictly speaking, the German and English child data contain no clear cases of root infinitives, because both languages have formally indistinguishable finite forms.[7] Only if we presuppose that children have mastered subject-verb agreement or if we take the absence of subjects to be indicative of root infinitives can we conclude that forms like *go* or *gehen* are actually infinitives. This is why we represent such forms as ambiguous in our summary further below (Table 4 and Figure 2) (it also allowed us to be consistent in coding adult and child data). We wish to emphasize, however, that these forms are likely to represent RIs. In German, only finite

---

[6] Note that utterances such as (18) are marginally possible as elliptical structures in adult English, e.g. *what do you want to do now? _ read books.*

[7] The only clear cases of non target-like nonfinite forms are participles and gerunds.

plural verbs (*wir/sie gehen* 'they go') are homophonous with infinitives (*gehen* 'to go') and since plural contexts were rare in the child-caretaker interaction we looked at, it is unlikely that the German children's verbs which look like RIs are in fact inflected plural verbs. In French, where the only finite verb forms indistinguishable from nonfinite verbs are plural imperatives (cf. 11), we coded the corresponding forms as non-target-like rather than ambiguous, as contexts requiring plural referents are rare in the data and the respective forms are not used by the adult French speakers either. Furthermore, we did not distinguish between participles and infinitives in French, because these have the same forms as the large class of verbs in *–er*.

Tense marking is also absent in utterances containing no verb at all, as in (19). These were coded as copula omissions.

(19) a.   un petit camion ça   (French, Philippe 2;1,26)
           a small truck   that

      b.   qua l' agnellino   (Italian, Marco 2;1,11)
           there the lamb

      c.   aqui carro.   (Brazilian Portuguese, Paolo 2;2,8)
           here car

      d.   lorry blue   (English, Dominique 2;1,11)

      e.   Puppe d(r)in
           doll   inside   (German, Rahel 2;1)

4.2. Results

4.2.1. Italian and Portuguese

The children acquiring Italian and Brazilian Portuguese produced mostly finite verbs. Only 5% of all nonfinite forms in Italian are not target-like and only 6% of all nonfinite forms in Brazilian Portuguese. In the Italian corpus these 5% constitute bare participles (see also Guasti 1993/94).

In the Brazilian Portuguese data, most nonfinite tokens (N=50) are infinitives (cf. 20). This seemingly high number should not be overvalued as the 50 occurrences include only 26 different verb types and many of them are repetitions (i.e. the child repeats himself), e.g. *pôr aqui* 'put here' (N=10) and *desligar a musiquinha* 'switch off the music' (N=8). Furthermore, there were 8 bare participles and 8 bare gerunds (cf. 21). The Italian- and Portuguese-learning children produced few copula omissions (between 0% and 23% in Italian, and only 2% in Portuguese).

(20) MOT: pega o   garfo +...
          take the fork
      CHI: **comer**
          eat-Inf
      MOT: vai   essa   primeiro
          comes this one first

(21) CHI: veio   correndo e   bateu
          comes-it running and broke
      FAT: dois carros?
          two cars?
      CHI: é.
          is (yes)
      CHI: três   carros **correndo**
          three cars   running

Interestingly, the Brazilian child often used 3rd person singular verbs when referring to the 1st person singular, even in combination with the 1st person pronoun *eu* 'I':

(22) a.   que eu *pintou*?
            what I painted-3.Ps.Sg.?'
      b.   *põe* eu
            'lays-3rd.Ps.Sg. I'
      c.   eu *quer*   aquele
            I want-3.Ps.Sg. this one

Overall, there were 82 utterances where the subject pronoun *eu* was combined with a finite verb. In 44 examples (54%), the child used the correct finite form of the 1st person singular. In 38 cases (46%) he used the non-target-like 3rd person singular form, even with modals like *querer* 'want' and auxiliary verbs like *ir* 'go'. We suspect that this observation is related to the high frequency of forms with this inflection. In other words, while infinitives might be the default forms for children acquiring German or English, the 3rd person singular form may play a similar role in the acquisition of Brazilian Portuguese.

4.2.2. French

The French-learning children produced a noticeably higher number of non-target-like nonfinite verb forms (19%). It is worth pointing out that French-learning children omitted few copulas (between 2% and 8%, mean 4%), possibly because the French copula is learned as a chunk in the presenta-

tional construction *c'est un/le+NOUN* 'that's a/the+NOUN', where the pronoun *ce* 'it/that' is proclitic to the copula *est* 'be'.

4.3.3. German and English

Given that we take the ambiguous forms to be nonfinite, as argued above, the German-speaking children produced a mean of 37% non-target-like nonfinite forms and 21% copula omissions, i.e. considerably more than the "Romance" children.[8] Interestingly, RIs even occurred with the verb *sein* 'to be', although previous studies (e.g. Hoekstra and Hyams 1995) have noted a tendency for RIs to occur with eventive verbs (see however Lasser 1997, Unsworth 2003). An example is illustrated in (23).[9]

(23)   CHI: das Ø keine Katze.
             that   no   cat
       MOT: ja richtig.
             yes right
       CHI: ein Schäferhund   **sein.**   (Falko 2;1,14)
             a   German shepherd be

In English, again assuming that the ambiguous verb forms are non-finite, non-target-like nonfinite forms constitute 58.5%, while copula omissions are as frequent as in German (22%). Many of the children produced no or hardly any lexical verbs which were unambiguously finite.

4.3.4 Summary

An overview of the results is presented in Table 4 and Figure 2. They may be summarized as follows:

The adult data exhibits only few cases of *non target-like nonfinite forms*. The numbers are too infrequent to have any overall effect in acquisition: 2.3% in German, 1.1% in Italian (only participles), 1% in Brazilian Portuguese, 0.1% in French. Moreover, if the token frequency of non-

---

[8] One child produced an extremely high number of copula omissions (85%). We did not exclude these data, as our exclusionary criterion (production of less than 5 lexical verbs) was not met. The exclusion of this corpus would have yielded a mean of 19% instead of 22% copula omissions.

[9] As noted by Ute Bohnacker, the example may also represent an instance of auxiliary omission (*das soll **ein Schäferhund sein** 'that must be a German shepherd'), as discussed previously in the work of Boser, Lust, Santelmann and Whitman (1992), Krämer (1993), and Bohnacker (1999).

target-like nonfinite verbs in the input would play a role, there should be no RIs in child French - contrary to fact.

If the *overall* frequency of *nonfinite verb forms* in the adult input were crucial in determining the children's rate of non target-like infinitives, one would expect the Italian and French children to exhibit similar patterns, as well as the Portuguese and the German children. However, the French children produced noticeably more RIs than the Italian children (and the Brazilian child), and the German children produced more RIs than the Portuguese child (and all other children acquiring a Romance language). Therefore, the number of nonfinite verb forms in the input alone cannot account for the amount of RIs produced by the children.

The *combined percentage of nonfinite and ambiguous verbs* amounts to the following percentages: 46% in English, 32.5% in French, 30% in German, 29% in Italian, 18% in Brazilian Portuguese. If these were crucial in pushing the rate of RIs in child language, a much higher rate of RIs in Italian would be expected.

Table 4: Distribution of finite and nonfinite verb forms in child data (age 2;0-2;2)

	categories	IT.	PTG.	FR.	EN.	GE.
unambiguously finite verbs	lexical verbs	295 (43%)	385 (31%)	256 (37%)	28 (4%)	262 (21%)
	auxiliaries/ modals	38 (5%)	278 (22%)	46 (7%)	30 (4%)	30 (2%)
	copula	168 (24%)	248 (20%)	138 (20%)	47 (6%)	171 (14%)
	imperatives	57 (8%)	29 (2%)	19 (3%)	-	7 (1%)
target-like nonfinite verbs	infinitives	20 (3%)	157 (13%)	57 (8%)	19 (2.5%)	28 (2%)
	participles/ gerunds	37 (5%)	43 (3%)	16 (2%)	22 (3%)	21 (2%)
non-TG nonfinite verbs	infinitives	-	64 (5%)	134 (19%)	-	-
	participles/ gerunds	32 (5%)	16 (1%)		57 (7.5%)	30 (2%)
morphologically ambiguous forms		-	-	-	384 (51%)	425 (35%)
copula omissions		45 (7%)	34 (3%)	28 (4%)	166 (22%)	253 (21%)
total		692 (100%)	1254 (100%)	694 (100%)	753 (100%)	1227 (100%)

Figure 2: Distribution of finite and nonfinite verb forms in child data (age 2;0-2;2)
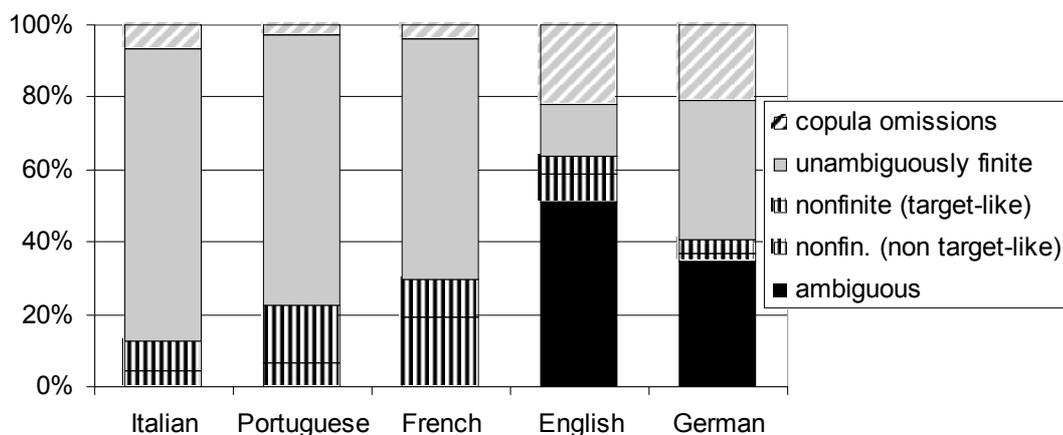
The only comparison resulting in a correlation between adult and child data is that between forms that are ambiguous between finite and nonfinite in the adult data and non-target-like nonfinite forms produced by children (Table 5). The question arises, however, how such fairly low frequencies in the adult data can cause such great effects in the child data. Possibly, the effect we found has been strengthened by ambiguities *within* the paradigm of inflected verbs. Since we did not quantify the respective numbers, however, this remains a plausible but speculative assumption to be investigated systematically in the future. [10]

Table 5: Ambiguous verb forms in input and non-target-like verbs in child speech

	IT.	PTG.	FR.	EN.	GE.
Ambiguous forms in adult speech	-	-	5%	12.5%	11%
Morphologically ambiguous forms	5%	6%	19%	56.5%	37%

---

[10] Note that we have not studied the amount of agreement errors here. Future research could focus on a comparison between agreement errors and RIs, and it may turn out that children in some languages start to produce finite verb morphology early but fail to use correct forms. The Italian data runs counter to this idea (Guasti 1993/94). Recall, however, that the Brazilian Portuguese data shows overgeneralization of 3rd person singular forms. Thus, where German and English-learning children resort to infinitives when they lack the knowledge of Tense morphology, Brazilian children may resort to a finite default form, yet marking Tense.

5. Conclusions

Our results disconfirm that the token frequency of infinitives (no matter whether target-like or non-target-like) in the input makes children produce root infinitives. Rather, the data suggest that syncretism and ambiguous verb morphology cause delays in the production of finite verb, i.e. Tense morphology. English and German show the highest amount of ambiguities and homophonies between finite and nonfinite verbs, and English- and German-learning children produce the highest amount of verbs unmarked for Tense, although German to a more moderate extent than English. The same is true for French, albeit to a more limited degree than in German. Italian and Brazilian Portuguese exhibit the most explicit Tense morphology, and children acquiring these languages produce the lowest number of non-target-like infinitives, participles and gerunds. Our results indicate that the morphological properties of the target languages can be related to the quantity of RIs in child language and are thus in line with recent claims by Legate and Yang (2005) and Blom and Wijnen (2006), according to which the RI-phenomenon in child language is best viewed and explained as a morphological learning problem.

Legate and Yang (2005) propose a selectional model of syntactic acquisition, according to which children are equipped with an innate, universal population of competing grammars. Children converge on their target grammar by means of a probabilistic algorithm which takes parsing success as primary data. Grammars that fit the input are strengthened, while non-fitting grammars are gradually demised, in proportion to the amount of parsing failure. The speed with which a grammar is rejected is proportional to the amount of evidence for Tense in the input, i.e., the frequency of verb forms explicitly and unambiguously marked for Tense. This account correctly explains why the RI-stage is shorter and less extensive in some languages than in others and is perfectly consistent with our findings (although the phase we examined here is too short to make strong claims about the *duration* of the RI-stage).

However, as Blom and Wijnen (2006) remark, Legate and Yang's hypothesis seems to lack an explanation of *how* children will know what the relevant morphemes, i.e. those marking Tense, are. Clearly, it is not immediately evident which input forms are marked for Tense and which are not. Blom and Wijnen (2006), who look at RIs in Dutch from a developmental perspective, argue that productive morphology (inflection) is the result of an extended and gradual learning process. In order to learn which parts of words are morphemic, children have to compare similar, partly overlapping forms. By correlating morphological overlap with differences in meaning and use, children discover Tense morphemes. In

languages with abundant Tense morphology, there are many different verb forms morphologically marked for Tense. Consequently, sufficient paradigmatic and lexical variation in the child's lexicon can be built up fast, which is crucial to attaining productivity in verbal morphology. As the authors stress, these suggestions do not contradict the idea that language-specific syntax is a very early attainment. Yet, productive morphology has to be learnt and is the result of an extended gradual process.

Taken together, these recent studies suggest that differentiated Tense morphology does not cause problems in learning, but facilitates and accelerates the acquisition process. The findings replicate the results and revive ideas put forward in the 1970s and 1980s, indicating that morphologically rich languages show relatively early acquisition of morphology (see, e.g., the articles in Slobin 1985). Within the generative framework, Clahsen and his colleagues (e.g. Clahsen et al. 1996) most strongly argued that full productivity of verbal agreement morphology is a crucial factor in the transition from an underspecified to an adult-like grammar, and that attaining productivity depends on lexical learning.

One aspect of our study seems to contradict our conclusions. Figure 3 shows that German and English children show high percentages of copula omissions (21% and 22%), while the three Romance languages only exhibit between 2% and 7%. However, English and German copula morphology is not as ambiguous as the morphology of lexical verbs and if variation can be correlated with ambiguity/homophony, it is expected that languages vary less with respect to copula omissions than they do with respect to lexical verbs. We can only speculate at this point. Possibly, children's use of RIs indicates that they have a general problem with the acquisition the syntactic category of Tense, which is also the category in which copula verbs and auxiliaries are base generated. But why do Italian/Portuguese children omit hardy any copulas/auxiliaries, while German/English children omit them so frequently?[11] We suspect that studying copula omissions from a developmental perspective may shed more light on this question. Recall that the 7% copula omissions observed in Italian only represent a mean rate, individual rates varying between 0% and 23%, which suggests that some children might already have passed the copula omission stage (cf. also Franchi 2006 for copula omissions in Italian). Hence, it is plausible to assume that children in all five languages pass through a stage of copula

---

[11] The reason why French children produce many RIs, while omitting only few copulas, was already alluded to above: most French copulas appeared in presentationals, such as *c'est un* chat 'that's a cat', where the subject pronoun *ce* is proclitic to the copula *est*. Hence, *c'est* may be produced as a chunk.

omissions, but in some languages children leave this stage earlier, possibly through an indirect influence of finite Tense morphology.[12]

This raises the question whether children also pass through an RI-stage in all languages. In fact, Rus (2007) argues that the division between RI and non-RI languages is too superficial. Children in all languages pass through an RI-stage, but in some languages this period is extremely brief. The idea is certainly plausible, but since our data only covers a small acquisition period, we do not wish to discuss it any further here.

Coming back to the initial question what root infinitives stand for, the results of our cross-linguistic comparison suggest that the use of infinitives in finite contexts are neither the result of performance limitations, nor do they point to a reduced syntax (indications for finiteness already being present in the data). Rather, they reflect a morphological learning problem which is related to ambiguities in the Input. In conclusion, the fact that children proceed along different paths toward the target grammar suggests that language learning is experience dependent. We believe that morphology is an important factor in acquisition. These ideas do not contradict the idea of an innate grammar and very early syntactic knowledge. Rather, they can explain why there is cross-linguistic variation despite an innate and universal grammar.

References

Aldridge, Michelle. 1989. *The acquisition of INFL*, Bloomington: Indiana University Linguistics Club.

Avrutin, Sergey. 1997. 'Events as units of discourse representation in root infinitives', in: Jeanette Schaeffer (ed.), *The interpretation of root infinitives and bare nouns in child language, MIT Occasional Papers in Linguistics* 12, 65-91.

Avrutin, Sergey. 1999. *Development of the syntax-discourse interface*, Dordrecht: Kluwer.

Blom, Elma and Frank Wijnen. 2006. 'Development need not be embarrassing: The demise of the root Infinitive and related changes in Dutch child language', ms., University of Amsterdam/Utrecht University.

---

[12] A related question is why the languages vary so much with respect to auxiliaries. The Brazilian child produces 22% finite auxiliaries, while the children learning the other languages only produce between 2% and 7%. There may be two reasons for these contrasts. First, we classified Ptg. *eu vou* 'I will' consistently as an auxiliary, although it can also be used to mean 'I come' (the distinction is not always clear). The same is true for *foi embora* 'was away' or 'went away'. Second, the child used many periphrastic verb forms, such as *vou*+infinitive (*eu vou olhar* 'I go to see') and *vou*+gerund (*vou brincando* 'go playing') or *estou*+gerund (*eu estou servindo* 'I am serving'). These constructions are typical of Brazilian Portuguese.

Bohnacker, Ute. 1999. *Icelandic plus English: Language differentiation and functional categories in a successively bilingual child*, Ph.D. dissertation, University of Durham.

Boser, Katharina, Barbara Lust, Lynn Santelmann and John Whitman. 1992. 'The syntax of CP and V2 in early child German (ECG): The strong continuity hypothesis', in: Kimberley Broderick (ed.), *Proceedings of NELS* 22, Amherst, MA: Graduate Linguistic Students Organization, University of Massachusetts, 51-65.

Clahsen, Harald, Sonja Eisenbeiss and Martina Penke. 1996. 'Lexical learning in early syntactic development', in Harald Clahsen (ed.), *Generative Perspectives on Language Acquisition,* John Benjamins. Amsterdam/Philadelphia, 129-159.

Franchi, Elisa. 2006. 'Patterns of copula omission in Italian child language'. In Vincent Torrens and Linda Escobar (eds), *The Acquisition of Syntax in Romance Languages*, Amsterdam/Philadelphia: John Benjamins, 35–158.

Guasti, Maria Teresa. 1993/94. 'Verb syntax in Italian child grammar: finite and nonfinite verbs', *Language Acquisition* 3, 1-40

Guasti, Maria Teresa. 2002. *Language Acquisition. The Growth of Grammar*. Cambridge, MA: MIT Press.

Hoekstra, Teun and Nina Hyams. 1995. 'The syntax and interpretation of dropped categories in child language: A unified account', in *Proceedings WCCFL XIV, CSIL*, Stanford University.

Hoekstra, Teun, Nina Hyams, and Misha Becker. 1997. 'The underspecification of number and the licensing of root infinitives', in: Elizabeth Hughes, Mary Hughes, and Annabel Greenhill (eds), *Proceedings of the 24th Annual Boston University Conference on Child Language Development*, Somerville, MA: Cascadilla Press, 293-306.

Hyams, Nina. 1996. 'Underspecification of functional categories in early grammar', in Harald Clahsen (ed.), *Generative Perspectives on Language Acquisition.* Amsterdam/Philadelphia: John Benjamins, 92-127.

Kato, Mary A. and Esmeralda Negrão (eds). 2000. *Brazilian Portuguese and the Null Subject Parameter*, Frankfurt: Vervuert.

Kempen, Masja, Steven Gillis and Frank Wijnen. 1997. 'Dutch children and their mothers' infinitives', in: Antonella Sorace, Caroline Heycock and Richard Shillock (eds.), *Proceedings of the GALA '97 conference on language acquisition*, Edinburgh, 4-6 April 1997, 85-89.

Krämer, Irene. 1993. 'The licensing of subjects in early child language', *MIT Working Papers in Linguistics* 19, 197-212.

Lasser, Ingeborg. 1997. *Finiteness in adult and child language*, Ph.D. dissertation, New York: CUNY Graduate Center.

Lasser, Ingeborg. 2002. 'The roots of root infinitives. Remarks on infinitival main clauses in adult and child language', *Linguistics* 40, 767-796.

Legate, Julie and Charles Yang. 2005. 'Morphosyntactic learning and the development of tense', ms., University of Delaware/Yale University.

Phillips, Colin. 1995. 'Syntax at the age of two: Cross-linguistic differences', *MIT Working Papers in Linguistics* 26, 325-382.

Platzack, Christer. 1992. 'Functional categories and early Swedish', in Jürgen Meisel (ed.), *The Acquisition of Verb Placement*, Dordrecht: Kluwer, 63-82.

Radford, Andrew. 1988. 'Small children's small clauses', *Transactions of the Philological Society* 86, 1-43.

Raposo, Eduardo. 1987. 'Case theory and Infl-to-Comp: the inflected infinitive in European Portuguese', *Linguistic Inquiry* 18. 1, 85-109.

Rizzi, Luigi. 1993/94. 'Some notes on linguistic theory and language development: The case of root infinitives', *Language Acquisition* 3, 371-391.

Rus, Dominik. 2007. The acquisition of tense and agreement in early grammars and early root nonfinites (in child Slovenian) revisited', in: Heather Caunt-Nulton, Samantha Kulatilake and I-hao Woo (eds), *Proceedings of the 31st annual Boston University Conference on Language Development*, Somerville, MA: Cascadilla Press. 543-554

Schlichting, Liesbeth. 1996. *Discovering syntax: An empirical study in Dutch language acquisition*, Ph.D. dissertation, Nijmegen: Nijmegen University Press.

Slobin, D. (ed.) 1985. *The crosslinguistic study of language acquisition*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Tsimpli, Ianthi-Maria. 1992. *Functional categories and maturation: The prefunctional stage of language acquisition*, Ph.D. dissertation, University College London.

Unsworth, Sharon. 2003. 'Testing Hulk & Müller (2000) on crosslinguistic influence: Root Infinitives in a bilingual German/English child.' *Bilingualism: Language and Cognition* 6.2**,** 143-158.

Vainikka, Anne. 1993/94. Case in the development of English syntax, *Language Acquisition* 3.3, 257-325.

Wexler, Kenneth. 1994. 'Optional infinitives, head movement and the economy of derivation', in David Lightfoot and Norbert Hornstein (eds), *Verb Movement,* Cambridge: CUP, 305-350.

Wijnen, Frank and Gerard Bol. 1993. 'The escape from the optional infinitive stage', in: Ale de Boer, Jelly de Jong and Rita Landeweerd (eds.), *Language and Cognition* Vol. 3., University of Groningen, 239-248.