Some Remarks on the Realization of the Realis/Irrealis Opposition in Child Language: Towards a Universal Characterization of the Root Infinitive Stage Across Languages

Susana Perales
Language Acquisition Lab, University of Ottawa

Juana Liceras
Language Acquisition Lab, University of Ottawa

Aurora Bel
Univesitat Pompeu Fabra

Abstract: In this paper we adopt Hyams’ (2001) proposal that RIs in child language fall out of the attempt to establish a realis/irrealis opposition. More specifically, we argue that the Root Infinitive stage or, in other words, the way in which children mark the realis/irrealis opposition, is manifested across languages depending on the (un)availability of the features Person [+/-P] and Infinitive [+/-R]. This proposal allows us to incorporate null subject languages into the discussion on RIs, which up to now had been considered not to show a RI stage. Furthermore, we claim that the forms selected by children to express realis and irrealis meaning can be predicted from the presence versus absence of these features in a given language.

Keywords: root infinitives, optionality, realis, irrealis

1. The Root Infinitive (RI) phenomenon

It has long been noticed that children often produce non-finite forms in contexts where the adult grammar requires an inflected verb, as in the examples below from Hoekstra & Hyams (1998).

(1) a. Dutch b. German
   Papa schoenen wassen   Thorsten das haben
   Daddy shoes wash-INF   Thorstn that have-INF

c. French d. English
   Michel dormir          Eve sit(∅) floor
   Michel sleep-INF

There are languages like Dutch, German or French (see (1a), (1b) and (1c)) where non-finite forms are actual infinitives, as evidenced by the presence of the infinitival morpheme on the verb. On the other hand, in languages lacking infinitival morphology, like English, the RI phenomenon is manifested by the
presence of bare forms in which the corresponding tense or agreement morphology has been omitted (see 1d). However, the early literature on RIs did not differentiate between the two types of languages (Harris & Wexler 1996, Rizzi 1993/1994, Wexler 1994).

Another cross-linguistic difference that was soon noticed is that in null subject languages, like Italian, Spanish or Catalan, the percentage of non-finite forms in root contexts is not as high as the one in non-null subject languages like Dutch or German, as shown in table 1 adapted from Hoekstra & Hyams (1998).

Table 1: Percentage of RIs in null subject and non-null subject languages

<table>
<thead>
<tr>
<th>Null subject</th>
<th>Child</th>
<th>% RIs</th>
<th>Non null subject</th>
<th>Child</th>
<th>% RIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian</td>
<td>Paola</td>
<td>0.07</td>
<td>French</td>
<td>Natalie</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Daniel</td>
<td>0.08</td>
<td>Swedish</td>
<td>Freja</td>
<td>0.38</td>
</tr>
<tr>
<td>Spanish</td>
<td>María</td>
<td>0.08</td>
<td></td>
<td>Tor</td>
<td>0.56</td>
</tr>
<tr>
<td>Basque</td>
<td>Mikel</td>
<td>0.13</td>
<td>Dutch</td>
<td>Laura</td>
<td>0.36</td>
</tr>
<tr>
<td>Catalan</td>
<td>Júlia</td>
<td>0.07</td>
<td></td>
<td>Tobias</td>
<td>0.36</td>
</tr>
</tbody>
</table>

This asymmetry in the overall percentages of RIs across languages led several authors to claim that children learning null subject languages did not go through a RI stage (Guasti 1994). Along these lines, numerous proposals attempted to establish a correlation between the fact that a language allowed null subjects and the absence of RIs in child data (Rizzi 1993/1994, Wexler 1994, 1998). However, if the RI phenomenon is considered a stage of child language development, then the question arises as to why some languages, but not others, have a RI stage.¹

¹ There are other accounts of the presence of RIs in child grammars. For instance, Gavruseva (2003, 2004, in press) attributes the presence of RIs in child L1 and child L2 grammars to the underspecification of the aspectual head (her ‘underspecification of the AspP hypothesis’). According to this author, the fact that the telic/atelic distinction is predicate-based in English (it resides not only in the syntax of aspectual semantics but also in the syntax/semantics of DPs as in *he drunk beer* versus *he drunk a glass of beer*) has implications for the acquisition of finiteness. The rationale is as follows: since in English, and in English-like languages, the specification of Q in the DP is linked to the telic value of the predicate and Q and Asp have a comparable value in
2. The RI stage in null subject languages

Some authors have tried to define a RI stage in null subject languages. Hoekstra & Hyams (1995) propose that Tense is a means of connecting the structural temporal meaning into the discourse and that the relation between discourse (CP) and Tense is encoded by different elements across languages: Number morphology for Dutch and English, Person morphology in Spanish and Italian, and Tense morphology in languages like Japanese. They propose that the RI stage derives from the underspecification of the corresponding feature for each language. Thus, in Dutch and English, the underspecification of the feature Number brings about the presence of RIs in child language. For null subject languages, these authors claim that the underspecification of the feature Person triggers the Avoid Plural Phenomenon, which alludes to the fact that in null subject languages children do not produce plural verbal forms.

Tsimpli (1992) and Ezeizabarrena (1997), after analyzing child Spanish and child Basque data, suggest that the 3rd person singular (a form that lacks inflection but for the thematic vowel of the verb) instantiates the form unspecified for agreement features and that this verbal form and RIs lack functional content. More recently, Salustri & Hyams (2003) have claimed that the imperative is the corresponding RI form in null subject languages. What is interesting about these two proposals is that the second person singular of the imperative and the third person singular of the indicative are exactly identical in some null subject languages and these forms are characterized by the lack of a specific morphology, as shown in the following examples in Spanish.

\begin{enumerate}
\item[2. a.] \textit{comer} (to eat)\\
\begin{align*}
\text{como} & \quad \text{¡come}(\emptyset)! \\
\text{comes} & \\
\text{come}(\emptyset) & \\
\text{comemos} & \\
\end{align*}
\end{enumerate}
As shown in (2), the third person singular of the present indicative lacks specific morphology and it only carries the thematic vowel of the verb. The same is the case for the second person singular of the imperative. In fact, other studies have also highlighted the importance that these less marked forms seem to have in the development of verbal morphology in child Spanish (Serrat & Aparici 1999, Aguado-Orea 2004, among others).

3. Aspectual and modal properties of RIs

Despite the attempt to bring together null subject and non-null subject languages, these accounts fail to account for the fact that finite and non-finite forms are not in free alternation in child language, as they have different interpretative properties. Wijnen (1997) showed that RIs in child Dutch are restricted to verbs referring to events and do not occur with stative or auxiliary verbs. Moreover, Boser, Lust & Santelman (1992), Ferdinand (1996), and Ingram & Thomson (1996) showed that in Dutch, German and French, RIs have, in most cases, a modal (irrealis) interpretation. In accordance with this, Hoekstra & Hyams (1998) formulate the Modal Reference Effect and its derivative, the Eventivity Constraint.

(3) Modal Reference Effect (MRE): with overwhelming frequency, RIs have modal interpretations.

(4) Eventivity Constraint (EC): RIs are restricted to eventive verbs.

However, the MRE and the EC could not be applied to English bare forms because it was found that only 13% of English bare forms had modal meaning versus 86% of RIs in child Dutch (Ud Deen 1997). Therefore, English bare forms cannot be compared to the true infinitives of Root Infinitive languages.
4. Hyams (2001): root infinitives are not optional infinitives

In order to account for the different aspectual and modal properties of RIs as well as for their apparent optional nature in child language, Hyams (2001:47) advanced the Semantic Opposition Hypothesis, which states that “the alternation between finite and non-finite forms falls out of the attempt to set up a system of semantic oppositions and map them onto morphological forms”. Therefore, according to this proposal, RIs are not optional because they are not in free alternation with inflected verbs. The Semantic Opposition Hypothesis states that children map meanings onto inflection elements on the basis of a semantic hierarchy in which mood represents the most primitive opposition: Irrealis mood (desire, necessity or futurity of some event) versus realis mood (actual occurrence, whether past or ongoing of some event).

The infinitival markers of the RI languages (Dutch, German, for instance) realize the irrealis mood, whereas finite forms realize realis mood. For English, Hyams proposes that the absence of infinitival morphology means that the child will not select the bare verb as a carrier of irrealis value. Bare forms in child English are pure realis forms with no temporal specification, whereas finite forms (those having the appropriate present or past morphology) mark a tense opposition. As for irrealis meaning, Hyams proposes that semi-auxiliaries of the sort of hafta, wanna, gonna are selected by children to express irrealis meaning.

Hyams goes on to argue that in Greek, a null subject language with no infinitival marker, it is the bare subjunctive form (or the bare perfective) that realizes the irrealis mood, leaving the expression of realis value to finite forms. And finally, for a null subject language like Italian, Salustri & Hyams (2003) argue that the imperative is the selected form to be the carrier of irrealis mood. The following table summarizes the typology with respect to the realization of the Semantic Opposition Hierarchy from which the cross-linguistic manifestation of the RI stage in child language can be inferred.
Table 2: Cross-linguistic expression of the realis/irrealis opposition

<table>
<thead>
<tr>
<th></th>
<th>Realis</th>
<th>Irrealis</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+/-NS]</td>
<td>[+/-R]</td>
<td></td>
</tr>
<tr>
<td>Dutch, German</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>English</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Greek</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Italian</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

NS: Null Subject; R: Infinitival Marker

Despite its theoretical appeal, this proposal does not account for the small number of RIs that are found in the data produced by children learning null subject languages. Even though the percentage of RIs is small when compared to the one obtained in non-null subject languages, it is consistent across languages and across children, as shown in table 1. Therefore, our proposal attempts to identify a null subject RI stage as predicted by the Semantic Opposition Hypothesis.

5. RIs in null subject languages

Although far from being a robust phenomenon, it has been well documented that children learning null subject languages also produce RIs. Liceras, Valenzuela & Díaz (1999) show that in child L1 Spanish RIs occur mainly between 1;7 and 1;8, as shown in table 3.

Table 3: RIs in child Spanish (adapted from Liceras, Valenzuela & Díaz 1999)

<table>
<thead>
<tr>
<th>Inflected</th>
<th>Infinitives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>Stage II</td>
</tr>
<tr>
<td>Magín</td>
<td>173 91.05%</td>
</tr>
<tr>
<td>María</td>
<td>147 67.7%</td>
</tr>
</tbody>
</table>

Stage I: from 1;0 to 1;9  Stage II: from 2;5 to 2;7
The *Rrealis/Irrealis* Opposition in Child Language

Liceras et al. also indicate that these infinitives can have intentional (modal) value sometimes whereas at other times they have an extensional (descriptive or ongoing activity) value. Perales, Spradlin & Liceras (2004) also isolate non-finite temporal infinitives, that is, RIs with a *realis* value, produced by María in the López-Ornat corpus.

(5) a. El otro buscar  (*irrealis* value)
    the other look for-INF
b. Este tapar  (*irrealis* value)
    this one cover-INF
c. Darle culo  (*realis* value)
    to spank-INF him

Ezeizabarrena (1997, 2002) provides data from two bilingual Basque/Spanish children (see table 4 below). As Spanish, Basque is also a null subject language which has specific infinitival forms (King 1994). Ezeizabarrena maintains that the production of some RIs is simultaneous to the action referred to, thus having a temporal or *realis* reading, whereas others have a modal or future reference. Many have a past (therefore *realis*) interpretation (see examples in (6)).

Table 4: RIs in child Basque (adapted from Ezeizabarrena 2002)

<table>
<thead>
<tr>
<th></th>
<th>Mikel</th>
<th>Jurgi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I</td>
<td>31.6%</td>
<td>20.15%</td>
</tr>
<tr>
<td>Phase II</td>
<td>15.2%</td>
<td>10%</td>
</tr>
<tr>
<td>Phase III</td>
<td>2.6%</td>
<td>3.1%</td>
</tr>
<tr>
<td>Phase IV</td>
<td>3.2%</td>
<td>1.9%</td>
</tr>
</tbody>
</table>

(6) a. Ho(r)i amatau  (*irrealis* value)
    that turn off-INF
b. Hartu Ana!  (*irrealis* value)
    Take-INF Ana
Finally, Bel (1998, 2001) also isolates a RI stage in three monolingual Catalan children. In her data, RIs represent 6% of the total sentences produced during the RI stage. In table 5 we represent the percentages obtained for one of the children during the period between 1;10 and 2;5. Bel’s analysis of RIs shows that infinitival forms convey modal as well as temporal meanings in different proportions depending on the child (see examples in (7)).

Table 5: RIs in child Catalan (adapted from Bel 2001)

<table>
<thead>
<tr>
<th>Age</th>
<th>RIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1;10</td>
<td>21%</td>
</tr>
<tr>
<td>2;0</td>
<td>3.2%</td>
</tr>
<tr>
<td>2;1</td>
<td>3.9%</td>
</tr>
<tr>
<td>2;2</td>
<td>9.8%</td>
</tr>
<tr>
<td>2;5</td>
<td>5.6%</td>
</tr>
</tbody>
</table>

(7) a. Sortir (irrealis value)
    come out-INF

    b. Aixó recollir, mama (realis value)
    this pick up-INF, mummy

To summarize, a close scrutiny of the production of RIs in null subject languages shows that (i) the number of RIs is scarce but consistent across children and across languages and (ii) that the RIs of null subject languages, unlike those of non-null subject languages, may encode both realis and irrealis value. Therefore, any account of the production of RIs by children learning null subject languages must account for these two findings.
6. Person and infinitival features and the realis/irrealis opposition

We would like to claim that the manifestation of the realis/irrealis opposition in child language will be determined by the presence/absence of the features Person [+/-P] and Infinitive [+/-R] in a given language. In other words, the presence of [+P] provides strong evidence for the realization of the realis/irrealis semantic opposition in that all person markers (the inflected forms) will realize the realis mood, whereas the bare forms (the third person singular indicative/imperative) may in addition realize the irrealis mood. In languages in which the infinitival morpheme is also available (as in Spanish, Catalan or Basque), this latter form competes with the Person marker to be a carrier for irrealis. Table 6 shows how the combination of [+/-P] and [+/-R] determines how the realis/irrealis opposition will be expressed in a language.

<table>
<thead>
<tr>
<th>Person feature</th>
<th>Infinitive feature</th>
<th>Realis</th>
<th>Irrealis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch, German</td>
<td>-</td>
<td>Inflected</td>
<td>Infinitives</td>
</tr>
<tr>
<td>Greek</td>
<td>+</td>
<td>Inflected</td>
<td>Bare subjunctive/perfective</td>
</tr>
<tr>
<td>Italian, Spanish</td>
<td>+</td>
<td>Inflected</td>
<td>Bare 3rd person Infinitives</td>
</tr>
<tr>
<td>English</td>
<td>-</td>
<td>Inflected</td>
<td>Semi-auxiliaries Bare forms</td>
</tr>
</tbody>
</table>

P: Person feature; R: Infinitive feature

In the absence of the corresponding Person morphology (Dutch, German, French) the infinitive marker will be a salient morpheme and will therefore be selected as an irrealis marker. When an infinitival marker is not available (Greek), the dichotomy person marker/bare form makes the bare form the salient candidate for irrealis. This is also the case for Italian, Spanish, Catalan and Basque. However, the presence of an infinitival marker in these languages also makes it a potential candidate for the realization of irrealis, giving rise to the
observed pattern that both forms may encode both values. Finally, in languages with neither Person nor Infinitive morphology (English), semi-auxiliaries will realize *irrealis*, both because they have a salient morpheme (the final vowel) and because they may serve as the proto-forms for the implementation of actual modals, although bare forms may also occasionally carry *irrealis*. In conclusion, what the data show is that when there is not a transparent bound morphology that can realize the *realis/irrealis* opposition, bare or unmarked forms as well as infinitival forms will encode both values.

References


The *Rrealis/Irrealis* Opposition in Child Language


