

---

# The optional infinitive stage and child L2 English<sup>1</sup>

Corinne Helland & Esther Álvarez  
Universitat de Barcelona

---

## 1 Introduction

The Optional Infinitive (OI) Stage of child L1 speech is one of the major discoveries of the last decade in early linguistic development (Wexler 1990, 1992, 1994). The stage is characterized by the appearance of non-finite verbs, most typically root infinitives and present participles, as possible grammatical sentences that may co-exist with finite forms. Wexler (2002) cites the concomitant characteristics of the stage for English to include among the following: tense morphemes appear only in correct syntactic positions; the semantic properties of tense morphemes are used correctly; auxiliaries as tense morphemes are omitted; accusative pronouns are often substituted for nominative pronouns as subjects of root clauses.

The purpose of this study is to examine a longitudinal corpus of data elicited from five bilingual (Catalan and Spanish) children by means of a narrative and a personal interview to show that an apparent OI stage of their L2 English is rather the result of the interaction of rules of the L1 and L2 grammars. The article is organized as follows. Section 2 includes a summary of the OI stage and its characteristics. In section 3 the methodology is outlined and in section 4 the data in the form of results are presented. Section 5 contains discussion and conclusions.

## 2 The Optional Infinitive Stage

When performing at the OI stage, in addition to producing root forms which lack subjects altogether, English speaking children produce forms such as the following:

- (1) a. he cries

- b. him cry
- c. him cried
- d. he cry

With respect to the variation in the forms produced, Schütze and Wexler (1996) and Schütze (2003) claim that these patterns of data can be explained by means of their proposed Agreement/Tense Omission Model (ATOM). The model assumes the separation of T(ense) and A(greement) (Chomsky 1993) and a strong separation between morphological case marking and structural licensing.

A consequence of the latter assumption is that whatever allows NPs to surface in particular positions is formally independent of whatever assigns morphological case features to NPs. A further assumption is that A, not T, checks nominative case. A and T can thus be underspecified in children's root clauses, the result of which yields a pre-determined set of feature combinations and utterance types, as exemplified below:

- (2) a. [+tns, +agr] = NOM assigned (e.g., *he cries*)
- b. [+tns, -agr] = NOM unassignable, default ACC (e.g., *him cry*, *him cried*)
- c. [-tns, +agr] = NOM assigned, agreement invisible (e.g., *he cry*)

When the configuration includes [+agr], as in (2a) and (2c), NOM case is necessarily assigned due to the presence of [+agr]. But the two configurations differ with respect to the verb form. In (2a) the verb form *cries* surfaces because the present tense suffix *-s* indicates the presence of both agreement and tense, whereas in configuration (2c) given the negative value of the tense feature, the present tense suffix *-s* cannot surface. When [-agr] appears in the configuration, as in (2b), NOM case is unassignable and thus default ACC case surfaces.

The verb can take two possible forms, with or without tense marked, as a consequence of the fact that the suffix *-ed* is not associated with agreement features such as person and number. In conclusion, the [tns, agr] configurations described account for a highly systematic distribution of nominative and accusative case assigned subjects and correlated verb forms. The fact that many of the subjects of root verbs appear in accusative case is another significant characteristic of the OI stage, exemplified in the structures in (3):

- (3) a. her going
- b. me here
- c. me like candy

As Wexler (2002) points out, nominative case forms rather than accusative case forms should be more likely to appear in the input, so a standard frequency argument does not suffice to explain the appearance of accusative case-marked subjects in child language. Wexler posits instead that accusative is the default case form in English.<sup>2</sup> The default case form is used when there is no structural case position. For example, in English speakers say *It's him*, not *It's he* and the answer to the question *Who wants candy?* is *Me*, not *I*. Children learn the specific default case form of their language based on input.

Based on the corpus of data studied herein, a stage similar to the OI stage of L1 development appears to exist in child L2. The child L2 stage, however, differs in significant ways from the L1 stage. First, finite and non-finite forms do not alternate freely, the former being almost non-existent. Second, errors occur in the semantic content of tense morphemes. Third, accusative pronouns are not substituted for nominative pronouns in any of the L2 data. Finally, copular and auxiliary forms of -be appear in their correct forms.

### 3 Method

#### 3.1 Participants

The child participants, who number five, are students of English as a foreign language in the Barcelona public school system from different schools in one of the city's middle class neighborhoods. Their first languages are Catalan and Spanish, so English is a foreign or second language for them. The participants were screened to factor out those who study or have studied English outside of school, who have English speaking parents or close relatives, or who have spent time in English speaking countries in order to equalize the input factor.

#### 3.2 Instrument and procedure

All child participants began instruction in English at an average age of 8;9 years and the narrative and personal interview tests were administered at three different time periods: data collection time one was after 200 instructional hours and at an average age of 10;9 years, data collection time two was after 416 instructional hours and at an average age of 12;9 years; data collection time three was after 726 instructional hours and at an average age of 16;9 years. The data are thus longitudinal and the result of a personal interview and an oral narrative based on a series of

six drawings.<sup>3</sup> The tests were audio-recorded and later transcribed into document form. The data collectors were instructed to intervene as little as possible.

## 4 Summary of the Results

At the first data collection time relevant examples are scarce but give the appearance of an OI stage (*look the map*). The second data collection time includes numerous root and present participles forms with (*I go to the excursion, dog eating, your ... he ... look the map*) and without nominal or pronominal subjects (*go to school, staying in the mountain*). The third collection time includes similar structures with more elaborate phrase structure (*The dog eat all of the food*). In the totality of the narrative data there appear, however, a total of two finite lexical verbs (*stayed: the dog stayed in the basket* and *came: the dog came into the basket*) in appropriate contexts. In the personal interview data there appears a single instance of a finite lexical verb form in the appropriate context (*went: Last weekend I went to the disco*) and a total of two finite lexical forms in inappropriate contexts. Thus, finite forms cannot be said to alternate freely with non-finite forms. Furthermore, in the narrative the use of copular *be* is consistently target-like by the third data collection time (*the dog is in the basket, the two children are surprised*) and auxiliary *be* as well (*they are preparing the breakfast, this son and daughter are going to an excursion*). Use of *do*-support in negative structures and questions (*I don't like some teachers, Do you like your job?*) and use of the future marker *will* (*I will go to the town.*), as well as *wh*-questions (*Where was (were) you born?*) also appear in target-like form in the personal interview. The L2 stage is thus characterized by virtual lack of alternation between finite and non-finite forms, the presence of nominative case-marked subjects with non-finite forms and adult target-like use of tense marked auxiliaries and copular *be*, properties inconsistent with the L1 OI stage. The specific characteristics of each of the data collection times are outlined in the sections below.

### 4.1 Data collection time one

Verb forms are scarce at data collection time one and thus relevant merely as the starting point. The examples in (3a) and (3b) contain the two forms which appear in the totality of the narrative data:

- (4) a. look the map
- b. cooking

Participant	Verb forms
A	0
B	1 root infinitive = 50%
C	1 present participle = 50%
D	0
E	0

Table 1: 10;9 years old after 200 hours of instruction

These results are summarized in Table 1.

Only a total of two verbs appear, a root form for subject B and a present participle form for subject C. Thus, in the entire corpus of narrative data at collection time one, there is a total of two verb forms, both non-finite. The results from the personal interview are similar:

(5) play football

(6) look tv

Table 2 contains the summary of the data.

It must be added, however, that in participant A's personal interview there occur two instances of auxiliary do, one in answer to a question (7), and the other (8) in a question:

(7) Yes, I do.

(8) Do you like (name of school).

In both sets of data verb forms are scarce.

Participant	Verb forms
A	0
B	1 root infinitive
C	0
D	0
E	1 root infinitive

Table 2: 10;9 years old after 200 hours of instruction

## 4.2 Data collection time two

At data collection time two the total number of verb forms is higher and a greater variety of forms appear. In the narrative there are root infinitive forms, as in (9), and present participle forms, as in (10):

- (9) a. cut bread  
b. go to school
- (10) a. dog eating  
b. staying in the mountain

No auxiliaries, tensed or untensed, appear in any of the narratives at this data collection time. There are examples with pronominal subjects in three non-finite verb sentences, which appear in (11):

- (11) a. I go to the excursion  
b. I go in the mountain  
c. Your ... he ... look the map

These three structures contain errors with respect to matching the story to the correct subject pronoun. In all three examples the correct subject pronoun would be they. The point illustrated, however, is that the attempt to produce a subject for these root forms generates a subject pronoun in nominative case. Table 3 summarizes the data from the narrative at data collection time two.

At the second data collection time a total of sixteen root infinitive forms appear and a total of six present participles in obligatory present progressive context but all lacking any form of the auxiliary be, tensed or untensed. The subjects that occur with these verb forms are noun phrases. There are several examples of other verb forms. Two participants, B and D, the latter twice, use the form have as an auxiliary:

Participant	Root : -tns	Prpart
	76%	24%
A	7	1
B	2	1
C	0	0
D	4	1
E	3	2

Table 3: 12;9 years old after 416 hours of instruction

- (12) have got  
(13) a. this picture have got a family  
b. have got dog

Participants A and E use copular be, the latter with incorrect agreement:

- (14) a. the dog is in the sandwich  
b. the dog are hungry

Participant D uses be as an auxiliary in the present progressive tense in an appropriate context:

- (15) this son and daughter are going to an excursion

With respect to the personal interview, the results are similar at data collection time two and the forms below representative:

- (16) a. play football  
b. playing football  
(17) a. watch tv  
b. watching tv  
(18) a. go mountains  
b. I go beach

Table 4 summarizes the results from the personal interview at data collection time two.

Again, there is a greater number of verb forms at time two but with the same degree of structural variance as at time one. Almost the same distribution between root infinitive forms without tense and present participles occurs in both the narrative and the personal interview. With respect to subjects, the narratives contain

Participant	Root : -tns	Prpart
	77%	23%
A	11	2
B	2	0
C	4	0
D	4	4
E	3	1

Table 4: 12;9 years old after 416 hours of instruction

noun phrase subjects such as *the dog*, *the boy and the girl* or *the mum*, whereas in the personal interview the majority of both types of verb forms appear without subjects, although in each interview a single instance of the nominative case-marked first person singular pronoun I appears.

With respect to other categories of structures, the use of main verb be both in questions and in affirmative statements with agreement is invariably correct, as exemplified in (19):

- (19) a. My name is (name).  
b. What's your name?  
c. How old are you?  
d. This school is very big.  
e. I am eleven years old.

There are also examples of questions with do-support:

- (20) a. Where do you live?  
b. Do you like sports?

Thus, the use of correctly inflected be as a main verb appears fairly frequently in the data and do-support also appears.

### 4.3 Data collection time three

At the third collection time there are still many root forms, although a greater number of obligatory constituents appear:

- (21) The dog eat all of the food.  
(22) The dog look the sandwich.

Furthermore, root forms may also appear with pronominal subjects in nominative case, as exemplified in (23):

- (23) a. He make the food.  
b. They stay in the mountain.  
c. We see your eat.

Present participles in isolation no longer appear. The participants who construct present progressive forms for obligatory contexts include both the tensed auxiliary form and a subject, either a pronoun in nominative case as in (24a) or a noun phrase as in (24b,c).



- (24) a. They are preparing the breakfast.  
 b. The mother is telling the street.  
 c. The boy is going to a tree.

Table 5 summarizes the results from the narrative at data collection time three.

As indicated in Table 5, root infinitive forms without tense constitute the majority of forms in the narrative data at collection time three. In these forms, like the correctly formed present progressive constructions, subjects alternate between nominative case-marked pronouns, most typically the third person plural form *they*, and noun phrase subjects such as *the boy and the girl* or *the dog*. Of the two root infinitive forms which appear with tense, one is regular, *stayed*, and the other irregular, *came*, but both forms appear with the noun phrase subject *the dog*. All of the participants use at least one tensed copular form.

With respect to the personal interview the results are similar. Root infinitive forms prevail across the corpus of data and are used for present tense contexts as well as past and future contexts. Nominative case-marked subjects, almost always in the first person singular pronoun, appear in the majority of examples which include a root infinitive form, although there are occasional null subjects:

- (25) a. I like the school.  
 b. This is easier.  
 c. I study.  
 d. I go to play football.  
 e. I have many friends.  
 f. I listening music.  
 g. I arrive at eight o'clock.  
 h. I study very much.

Participant	Root : -tns 70%	Mainv+tns6% 6%	S+Aux+tns+pp23% 23%
A	4	0	3
B	5	0	1
C	6	0	0
D	6	1	1
E	4	1	2

Table 5: 16;9 years old after 726 hours of instruction

- (26) a. — have a sister.  
 b. — is older.  
 c. — is in Tarragona street.

Table 6 summarizes the results from the personal interview at time three.

In the personal interview at data collection time three, the majority of verb forms are root infinitives and the overall majority of these forms appear with the first person singular nominative case-marked subject. Only a total of three tensed main verbs appear, one in the appropriate context, (*Last weekend I went to the disco*), and two in an inappropriate context (*I went my house, I had the homework*).<sup>4</sup> For usage of the present participle not in the full present progressive form, participant A uses a nominative case-marked pronoun as the subject of one participle (*I listening music*) and no subject at all for the other two present participles. The single present participle used by participant C also contains a nominative case-marked pronoun (*I listening music*).

With respect to other structures used, main verb be and do-support are for the most part target-like and there are correctly formed wh-questions as well.

## 5 Discussion and conclusions

The results outlined in section 4 show that child L2 English replicates to a certain extent the child L1 OI stage. Root infinitives and present participles constitute the majority of verb forms present across all three data collection periods. What is missing for a more consistent replication of the OI stage, however, is alternation or co-existence with finite forms. The very few forms finite forms that appear tend not to occur in appropriate contexts. For example, at data collection time three in the narrative an overall total of 30 verb forms appear and only two of them are tensed main verbs, although both used in the correct context. In the personal interview

Participant	Root : -tns 86%	Mainv+tns6%	S+Aux+tns+pp23%
A	1	0	3
B	11	1	0
C	6	2	1
D	11	0	0
E	14	0	0

Table 6: 16;9 years old after 726 hours of instruction

there is a greater total number of verb forms, 50, but only a single form is used in the correct context and two in an incorrect context. It is therefore difficult to posit that the same sort of alternation between non-finite and finite forms characteristic of the L1 OI stage exists for the L2 data. Furthermore, from a theoretical perspective there is an important difference. Underlying the OI stage is the fundamental tenet that children do not distinguish tense values because they do not understand tense. The claim leads to the assumption that the OI stage can only conclude once past tense has developed and is understood (Wexler 1994). The tenet does not hold true, however, of the L2 participants of this study who begin instruction in the L2 at an average age of 8;9 years and who thus already can be assumed to dominate the notion of tense in their first languages, so there is a fundamental difference in the starting point of the proposed L2 stage as opposed to the starting point of the OI stage of L1 child language.

Lack of alternation with finite forms may have another explanation from the L2 literature along the lines of the Missing Inflection Hypothesis (Haznedar and Schwartz 1997). The participants' knowledge of English verbal inflectional morphology is limited, so target-like verbal morphology just may not be mapped onto the syntax in the spirit of Haznedar and Schwartz (1997). Finally, characteristic of the L1 child OI stage is that tense morphemes appear only in correct syntactic positions, a characteristic also found in the L2 data but to a very limited extent since so few tense morphemes appear in the data making this an unreliable diagnostic of a possible L2 child OI stage.

In the English child L1 OI stage, there exists alternation among the following structures: *he cry*, *he cries*, *him cry* or *him cried* (Wexler 2002). The appearance of one or another of the forms is based on the combination of the values of tense and agreement. In the child L2 data the first type of structure, *he cry*, prevails, and constitutes always above 70% of the total verb forms present, except in the case of narrative at data collection time one, where only two verb forms appear in the totality of the data, one root infinitive and one present participle. Thus, the form with the inflectional feature matrix of [-tns, +agr] appears but other combinations of tense and agreement features seem not to occur. Following Chomsky (1995) functional categories and their feature specifications are the locus of all cross-linguistic differences. For first language development Wexler (1994) and Hyams (1996) propose that children begin the development process with a full set of functional categories, including INFL (tense and agreement). For second language development, assuming the Full Transfer / Full Access hypothesis (Schwartz, B. and R. Sprouse 1996), the L2 initial state also includes a full set of functional categories for which second language speakers must develop the features and their relevant values. Thus, in L2, until the relevant features and values are established, both agreement and tense

should be omitted or underspecified. In other words, the category is present, as in the L1 initial state, but due to its lack of specification it is omitted in speech, and as mentioned above, the L2 children may have difficulties mapping the morphology as well.

Children performing at the L1 OI stage correctly use the semantic properties of tense morphemes. In contrast, in the L2 child data examined here, tense morphemes are for the most part non-existent or if used may appear in an inappropriate context. For example, participant C at data collection time three replying to a question about the future, employs the past tense:

(27) I went my house and I had the homework and I study.

Another characteristic of the L1 OI stage is that auxiliaries as tense morphemes are omitted. Nevertheless, in the L2 data examined here, by data collection time three -be as an auxiliary is used correctly and consistently for the present progressive tense by a total of four out of the five participants in the narrative data and representing a total of 23% of verb forms for the narrative but only 8% for the personal interview. This difference in percentages between the two data collection instruments should be expected given the context of the exercises, narration of a story viewed in drawings versus the participant answering questions about his or her everyday life.

Furthermore, at data collection time two, nominative case-marked pronominal subjects appear consistently and there is not a single instance of an accusative case-marked pronominal subject in the corpus of the data. This divergence is confirmed at data collection time three where more pronominal subjects, all in nominative case, appear with non-finite verbs. In contrast, at the L1 OI stage, when children supply a subject, the subject may appear in accusative case depending upon the value of the agr feature. Wexler (2002) points out that accusative is the default case form in English and for this reason accusative case-marked subjects appear in the OI stage. In contrast, the default case in the L1s of the participants is nominative, as the examples in (27), in Catalan, and (28), in Spanish, demonstrate:

- (28) a. Qui vol més? Jo/\*mi.  
Who wants more? I/me.  
b. Qui hi ha? Jo/\*mi.  
Who is (there)?I/me.
- (29) a. ¿Quién quiere más?Yo/\*mí  
Who wants more?I/me.  
bg. ¿Quién hay? Yo/\*mí.

Who's there? I/me.

Children learn the default case form of their language based on input. Thus, Catalan and Spanish speaking children learn the default case to be nominative, unlike English speaking children who learn accusative case as the default case form. Given the instructional context of the participants, it seems unlikely that they would have an opportunity to learn the default case of English based on input. Instead, they transfer the default case of their first languages to their production in English.

Finally, it should be noted that untensed auxiliaries are unattested in the corpus of data. In other words, in obligatory present progressive contexts, the present participle appears without any form of *be* and there is no transition phase where a form of the auxiliary appears in an untensed form. At data collection time three, however, present progressive forms appear correctly structured with a tensed form of *be*. Yet in the child L1 OI stage auxiliaries as tense morphemes are omitted.

In conclusion, at first glance the data give the appearance of a child L2 developmental stage similar to the OI stage of L1 child speech as posited in Wexler (1990, 1992, 1994), but on closer inspection it becomes clear that the stage lacks many of the characteristics of the L1 OI stage. At issue may be that incremental stages of L2 development in which the characteristics overlap to a greater extent occur in between the established data collection periods. Nevertheless, the five participants do not seem to be at exactly the same points in their L2 development at any of the data collection times, so the data may already show that such characteristics do not exist at any incremental stages of development.

## Notes

1. The data analyzed belong to the BAF (Barcelona Age Factor) Project currently in progress at the University of Barcelona. The authors gratefully acknowledge DGICYT for financial support in the form of grants PB94-0944, PB97-0901 and BFF2001-3384.
2. The default case form varies cross-linguistically (Wexler 2002).
3. Alvarez and Helland (to appear) apply a system of stages to account transversally for the narrative data.
4. The question asked was 'What will you do later today?' and a comprehension error cannot of course be ruled out as the source of the tense error.

## **Bibliographical references**

- Álvarez, E. A. & Helland C. (to appear) "Route and rate of acquisition in EFL narrative development at different ages." In Muñoz C. (ed.) *Age and foreign language learning rate*. Clevedon: Multilingual Matters Ltd.
- Chomsky, N. 1995. *The Minimalist Program*, Cambridge, MA: MIT Press.
- Chomsky, N. 1993. "A minimalist program for linguistic theory," in K. Hale & S. J. Keyser (eds.), *The view from Building 20: Essays in honor of Sylvain Bromberger*, MIT Press, Cambridge, MA, 1-52.
- Ionin, T. & Wexler, K. (2000). L1-Russian Children Learning English: Tense and Overgeneration of "Be". To appear in the *Proceedings of the Second Language Research Forum*, Madison, WI.
- Schuetze, C. 2003. Why non-finite be is not omitted and finite be is. Handout for the Boston University Conference on Language Development 28.
- Schütze, C. T. & Wexler, K. (1996). Subject case licensing and English root infinitives. In A. In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (Eds.). *BUCLD 20* (pp.670-681). Somerville MA: Cascadilla Press.
- Schwartz, B. & R. Sprouse (1996) "L2 cognitive states and the 'full transfer/full access' model". *Second Language Research* 12, 40-72.
- Wexler, K. (1990) Optional infinitives, head movement and the economy of derivations in child grammar. Paper presented at the Annual Meeting of the Society of Cognitive Science, MIT, Cambridge, Mass.
- Wexler, K. (1992). Optional infinitives, head movement and the economy of derivation in child grammar. Occasional paper #45. Center for Cognitive Science, MIT, Cambridge, Mass.
- Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot and N. Hornstein (Eds.). *Verb Movement*. (pp. 305-350). Cambridge University Press: Cambridge, U.K.,
- Wexler, K. (2002). "Lenneberg's Dream: Learning, Normal Language Development and Specific Language Impairment". In J. Schaffer and Y. Levy (Eds.), *Language competence across populations: Towards a definition of Specific Language Impairment*. Earlbaum, 2002.