What are Root Infinitives Good For?

Michael Clauss

December 2013

Here I examine the question of whether or not Root Infinitives (RIs) in children's speech have a semantic-functional utility. I will introduce evidence that has been proposed to support such an account, discuss how this account divides languages as RI or non-RI languages in a different way than other accounts, and examine experimental extensions of this account.

1 Infinitives and Modality

A central area of inquiry among acquisition researchers of various theoretical backgrounds is how to explain patterns of the acquisition of tense and agreement morphology, particularly for languages (such as English, Dutch, and French) where these are somewhat delayed. Children commonly go through stages where matrix clauses show the morphology of embedded infinitival clauses, called Optional or Root Infinitives. Notable about these stages is that they often display multiple correlates of lack of tense: no agreement, no tense, no nominative case marking (in English), verb final word order (in Dutch), and lack of post verbal negation (in French).

(1)   RIs cross-linguistically.

    a English (Wexler 1998)
       Him like ice cream

    b French (Pierce 1992; cf. *Il (le) casse pas*)

       pas   cass-er
       NEG   break-INF
       'Didn't break'

    c Dutch (Hoekstra and Hyams 1998, cf. *Papa wast (de) schoenen)*

       Papa   schoenen   wass-en
       Papa   shoes   wash-INF
       'Papa wash the shoes'

One way of looking at these, which a number of researchers have adopted, has been to focus on semantic properties, particularly those which are found in the adult grammar, which children may be encoding into these non-adultlike utterances. One which arises frequently in particular is that infinitives in adult grammars often encode irrealis modality. Modality in general is divided into realis and irrealis, with the latter having further subtypes. Such a typology is described in (2).

(2)   Types of modality

       a Realis: Denotes an assertion that holds in the real world

       b Irrealis, Deontic: Denotes likelihood of a world where an utterance holds as true, based on external forces (laws, etc.)

       c Irrealis, Epistemic: Denotes likelihood of a world where an utterance holds as true, based on speaker's commitment to it

As children typically develop the use of Epistemic semantics later than RIs tend to last (Deen 2011), in the following discussion "Irrealis" will refer primarily to deontic modality, though particular types of irrealis semantics will not be closely discussed in general.

The following examples from Stowell (1982) demonstrate the modality of infinitive clauses in English; (3a) has an infinitival complement, and the sentence does not require either that the event described in the complement is true, nor that anyone believes it to be true. Conversely, (b) has a participial complement, and this requires that either the complement is true or that it is true to the best of the subject's knowledge.

(3)   a Jenny remembered to bring the wine

         Holds if Jenny did not bring the wine, but remembered that she was expected to do so

       b Jenny remembered bringing the wine

         Only holds if Jenny brought the wine (or at least does not doubt that she did)

Further, Stowell develops the idea that infinitives are not simply clause which lack a specification or projection for Tense in the syntax, but rather have the full structure associated with tensed clauses, and are simply semantically differentiated, and in a sense underspecified for temporal interpretation. This sort assumption, that infinitives are not structurally deficient but rather functionally distinct, plays an important role in work on Root Infinitives that looks to find a semantic function for them related to the irrealis interpretations associated with adult infinitives. In the following, I will examine such hypotheses. The next section looks at data, mostly from naturalistic studies of child languages, that supports such an approach. The third section describes empirical and theoretical challenges for this approach. Finally, the fourth section describes how experimental approaches could better guide and either confirm or disconfirm this sort of hypothesis.

2   The Modal Reference Effect in RIs

In looking at RI effects in Dutch and other languages, Hoekstra and Hyams (1998) show that many languages show what they term the Modal Reference Effect (MRE) on Root Infinitives:

(4)   *Modal Reference Effect*:

      With overwhelming frequency, RIs have modal interpretations.

Later work by Deen and Hyams (2006) reformulates this generalization as a hypothesis on functional projections in children's syntactic representations:

(5)   *Semantic Opposition Hypothesis* (SOH):

      The expression of irrealis mood in the early grammar excludes a tense specification.

This is further developed by Deen and Hyams as a language-sepcific morphological rule: in Child Dutch, where children's RI utterances have future or modal interpretations roughly 86% of the time, there is a rule that the morphological correlate of [-T] (or [-Fin]) is interpreted as irrealis; in English, where only about 11% of RI utterances have future or modal interpretations, there is no interpretive rule concerning [±Fin]. What is predicted by this hypothesis is an acquisition path for finiteness like the one shown in (6)

(6)   An acquisition path with the MRE

	I	II	III
Realis	-Fin	+Fin	+Fin
Irrealis	-Fin	-Fin	+Fin
Infinitive	-Fin	-Fin	-Fin

This schematizes an acquisition path with three stages: one where finiteness morphology has not been acquired, and thus non-finite forms appear in all contexts (I); one where finiteness morphology has been acquired, but the morphological correlate of the Semantic Opposition Hypothesis is in effect, and thus non-finite forms appear in both adult Infinitival contexts and more generally in semantically irrealis contexts (II), and a final stage where non-finite forms appear only when syntactically licensed in the adult grammar (III). Stages I and III seem fairly evident (particularly III, it being the adult grammar, which children typically do reach eventually). The crucial stage for this account is II, where the morpho-syntactic realization of [-Fin] is used particularly in contexts where it is licensed by irrealis semantics. In this section I examine apparent evidence of such stages. In doing so I will adopt a formalization of the morphological generalization of Deen and Hyams:

(7)   *Morphological extension of SOH*:

Children will assign a unique morphological exponent to irrealis, and this form will be a non-tensed form in the adult grammar.

2.1   Dutch

The study by Hoekstra and Hyams examining the MRE focuses on data from Dutch, where has long been observed to exhibit the MRE; they reproduce the following figure from Wijnen (1996) showing this.

		Present	Future/Modal	Past	Total
(8)	-Fin	194 (10%)	1625 (86%)	64 (3%)	1883
	+Fin	657 (93%)	21 (3%)	21 (3%)	699

Here, the number of utterances without finiteness marking seem to be concentrated heavily among utterances with a modal intent, and present tense finiteness marking is much more common than irrealis (or, for that matter, past). These data are based on the apparent intent in context of utterances like (9), where the matrix infinitive seems to be used with deontic sorts of intent.

(9)   *Niekje   buiten   spiel-en*
       Niek   outside   play-INF
       'Niek (the speaker) wants to play outside'

Thus Dutch seems to be an example of Stage II, in that children's RI utterances seem to be overwhelmingly restricted to irrealis conditions.
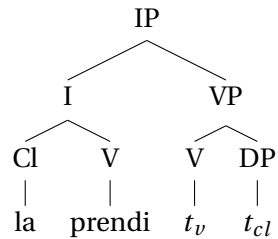
2.2 Italian

A widely observed generalization about RIs is that they are more likely to occur in languages with mandatory overt subjects (English, German, French) than in languages with optional subject drop (Italian, Spanish). However, Salustri and Hyams (2003, 2006) explore a possibility of child Italian exhibiting an equivalent phenomenon. Their studies focus on instances of the imperative. The imperative is a semantically irrealis form, and typically doesn't show tense (though in some languages it shows a form of agreement). Salustri and Hyams's claim is based around two observations: one concerning the syntactic form of the imperative, the other concerning its use in child Italian.

Regarding form, Salustri and Hyams claim imperatives, like French and German infinitives, do not undergo verb movement, as evidenced by relative position of verb and clitic.
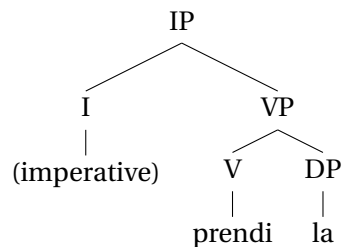
(10)   a   Indicative

      la   prendi
      it.Cl take.2Sg
      'You take it'

```
                    IP
                  /    \
                 I      VP
               /  \    /  \
              Cl   V  V    DP
              |    |  |    |
              la prendi t_v  t_cl
```

    b   Imperative

      prendi   la!
      take.2Sg it.Cl
      'Take it!'

```
                    IP
                  /    \
                 I      VP
                 |     /  \
           (imperative) V   DP
                        |    |
                      prendi la
```

The claim regarding children's behavior is that children produce a number of imperative-looking utterances that does not match their input. Specifically, they point out that while children learning German (in a study of children ages 2;0 to 2;4) produce imperatives at a *lower* rate than what they hear (while producing many RIs, which show an MRE), children learning Italian produce imperatives more than their input and almost no RIs. Thus, it looks like, in a language where RIs per se are blocked, imperatives do their job.

This is claimed to fall out from economy constraints, as imperatives don't move, but it would also fall out from the morphological extension of the SOH discussed above: regardless of conditions on RIs per se, grammatical constraints pressure children into assigning a specific morphological exponent to irrealis, and in this case choose the imperative.

2.3 Tamil

Acquisition studies of Tamil are somewhat scarce, but it has been claimed previously that the language does not exhibit RI patterns (Sarma 1999). However, a close examination of corpus data from the Vanitha corpus on CHILDES (MacWhinney 2000) indicates that something like an RI may exist, though it may more resemble the imperative pattern seen in Italian.

Adult Tamil marks tensed verbs with a tense marker followed by agreement, modals by an infinitive marker followed by a modal suffix, and imperatives by a bare stem.

(11) a *pō-v-ān*
         go-FUT-3SGM
         'He will go'

      b *avan pō-kka-ṇum*
         He   go-INF-must
         'He must go'

      c *pō!*
         go
         'Go!'

A previous study of child Tamil showed modal morphology emerging around 2;2 to 2;6 (Raghavendra and Leonard 1989), interestingly a similar age range to where children learning languages like English and French end RI stages.

An examination of the Vanitha corpus shows fairly few instances of verbs without expected tense and agreement morphemes. However, there are a number of instances of bare stems where inflected verbs would be expected, and a number of these seem to be in irrealis contexts, both modal (12) and future (13). Note also in (13) that Vanitha is allowing other morphology on the bare stem (the question particle -*ā*), showing that this is not purely a resistance to morphologically complex verbs. There is also exactly one example of something that looks like an RI proper in the corpus (14).

(12) Bare stem as modal (1;7)

      tūkki   enju
      away.DAT   throw
      'I want to throw (this) away!'

(13) Bare stem as future (1;3)

      vari-yā?
      come-Q
      'Will you come?'

(14) RI (2;4)

      nān   vari-kka-vā?
      I   come-INF-Q
      'Shall I come?'

These data are summarized in (15), which shows, for three semantic categories of inflection (past/present/realis, future, modal), the earliest age where adultlike morphology appear in her speech, and the percentage of non-adultlike morphology (ie. bare stems and the single RI) seen for each function.

(15)   Vanitha's errors[1]

	Realis	Future	Modal
Onset of morphology	1;5	1;8	2;1
Error rate	1.2%	15.1%	7.6%
	(*3/254*)	(*8/53*)	(*5/65*)

It thus appears that Vanitha goes through something like the expected Stage II described above: adultlike realis forms are used consistently from a fairly early age while unexpected bare stems remain, with the latter apparently reserved for irrealis contexts. This also seems to be a positive result for the Morphological extension of the SOH, with the bare stem being a morphological reflex of Irrealis, rather than Irrealis being marked by a general avoidance of functional projection in RI-like contexts (as demonstrated by the presence of polar question morphology on non-adultlike modals). It seems then that, at least for Vanitha's data, the modal account of RIs is on the right track.

3   Challenges to this account

If we adopt this view of RIs, we are forced to say that English, which shows no Modal Reference Effect, is not a Root Infinitive language in the same sense that German and French (and, putatively, Italian and Tamil) are. There is nothing wrong in principle with this terminological shift, though it does require a separate description of what goes on in "Root Infinitives" in English. And this introduces a problem: RIs in English and German share qualities which *are* predicted by purely morpho-syntactic accounts, accounts that liken all languages which go through similar looking stages.

Wexler (1998, 2010) gives an account of RIs as an interaction between an overly stringent constraint on Agree relationships between DPs and functional heads which is peculiar to children and the conditions on the morphological realization of tense and agreement. What is crucial about this account is that, as the licensing of tense and agreement are linked with other syntactic phenomena, such as Nominative case in English and Verb Second in German. The importance of this feature of the account is apparent because RI patterns, in English as in German, show a sensitivity to these licensing conditions. The figures in (16-17) show this.

Four permutations of tense and its secondary correlates are possible: An adultlike utterance (tense and its correlate represented), a true, prototypical RI (neither tense nor its correlate), a tensed verb without finite syntax (tense but not its correlate), and finite syntax without a tensed verb (no tense, but the correlate of tense). All of these are seen in child grammars except the one in which the secondary correlates of tense are represented but tense is not.

(16)   English RIs

Tense	Subject Case	
	Nominative	Accusative
+Fin	She walks	Her walks
	(Adultlike)	(No secondary correlate)
-Fin	She walk	Her walk
	(No tense)	(Prototypical RI)

---

[1]If we combine Future and Modal to make an Irrealis category, its error rate is 11.1%.

(17) German RIs

Tense	V Movement	
	V2	V Final
+Fin	John plays outside (Adultlike)	John outside plays (No secondary correlate)
-Fin	John play outside (No tense)	John outside play (Prototypical RI)

The important pattern here is that while children may produce adultlike utterances, never produce tense without its secondary correlate (case in English, movement in German). Again, this is predicted if RIs are a unified cross-linguistic phenomenon driven by constraints on morphosyntactic derivation, but not if they are distinct phenomena for German versus for English.

Some aspects of this are quite amiable with the modality hypothesis; In Deen and Hyams (2006, inter alia) the explanation for using infinitives to express Mood being that there is a limit in how many functional heads children will project in the syntax. This thus becomes a story of Economy of Representation, rather than Economy of Derivation; sufficient functional material to license adultlike modal constructions is not available, and this deficiency correlates with deficient morphosyntactic realization of Tense; in fact, according to Hyams, the Tense projection specifically is not licensed when room must be made for the Mood projection.

The account then runs into the question of what the difference is between English and German syntax, such that tense must be sacrificed only when Mood must be licensed in the latter, but in the former Tense may at any time be barred from projection. One such difference might be the unpredictability of morphology in English modal constructions, especially regarding finiteness; there is no bound morpheme expressing [-T] (other than an empty morpheme) in English, and the free morpheme 'to' appears in some but not all modal constructions (as in 18). It is thus difficult for the child to learn the mapping of infinitive morphology to the Mood projection in English.[2]

(18)   a She wants *(to) go

         b She must (*to) go

It's not immediately clear that this account may in fact be saved in this way, but this is the path of argumentation which it would require. The next section examine begins down this path by looking at potential avenues of experimental study of English-learning children's knowledge of the connection between (non-)finiteness and modality.

4   How to investigate this account

The fact that English RIs don't seem to show the Modal Reference Effect raises two distinct questions: (1) are English RIs the same phenomenon as other languages' RIs? and (2) are there other circumstances where children display knowledge of the modal quality of adult English modals? Here I investigate the former.

In an experiment looking at the relationship between tense in complement clauses and children's answers to syntactically complex questions, de Villiers et al (2012) found that, in fact, the scope of interpretation given to embedded Wh words varies based on whether they are attached to finite or non-finite clauses; that is, in (19a), where the embedded clause is infinitival, children will interpret 'what' as having scope over 'say', but in (b), where the embedded clause is finite, 'what' scopes only over the embedded clause.

---

[2]Deen and Hyams (2006) assert specifically this, but do not give a detailed explanation for *why* it ought to be the case.

(19) a Did Mary say what to feed the rabbit? → What is such that Mary said to feed it to the rabbit

b Did Mary say what Jimmy fed the rabbit? → What is such that Jimmy fed it to the rabbit

De Villiers et al explain this in terms of syntactic constraints on Wh movement related to tense features on I^0.

(20) a Infinitive complement → Wh must move and receive wide scope
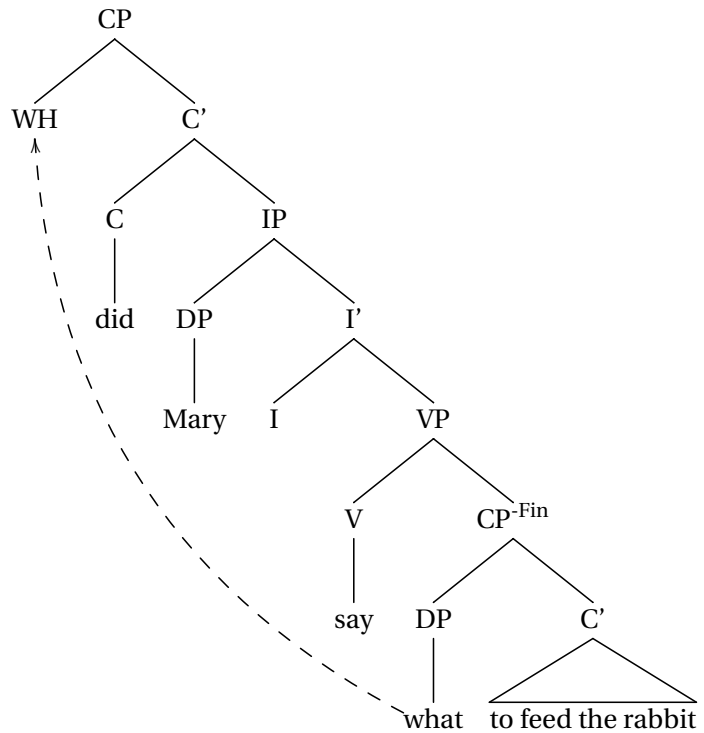
b Finite complement → Wh stays and receives narrow scope

CP
                          ┌──────────────┴──────────────┐
                         WH   C'
                                          ┌──────────────┴──────────────┐
                                          C   IP
                                          │   ┌─────────────┴─────────────┐
                                         did   DP   I'
                                                           │   ┌─────────────┴─────────────┐
                                                         Mary   I   VP
                                                                                      ┌───────────────┴───────────────┐
                                                                                      V   CP^{+Fin}
                                                                                      │   ┌──────────────┴──────────────┐
                                                                                     say   DP   C'
                                                                                                      │   ╱‾‾‾‾‾‾‾‾‾‾‾‾╲
                                                                                              what Jimmy fed the rabbit
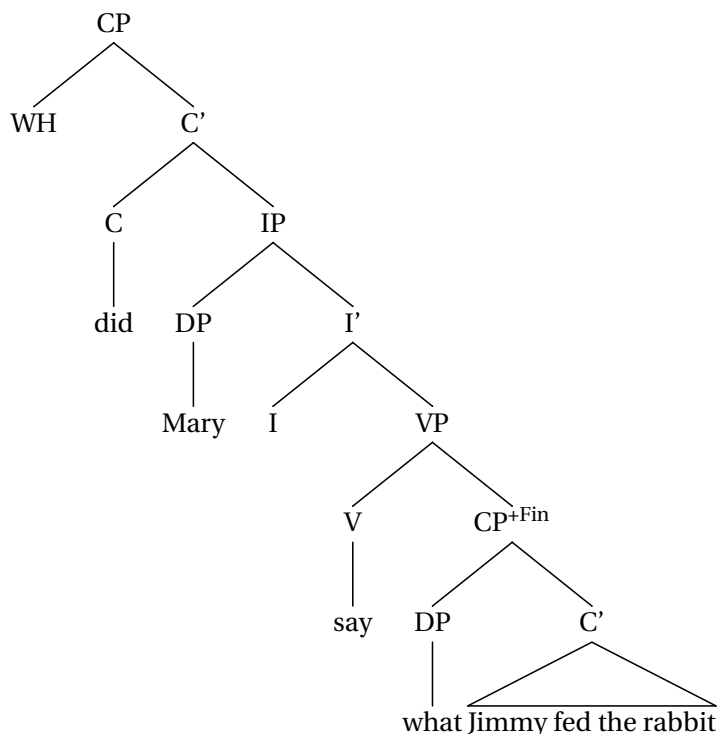
An alternative analysis, however, would be to say that the interpretation of the Wh word is not what changes across these two contexts, but rather the interpretation of the complement. Specifically, that the finite clause with past tense marked requires that the lower verb be interpreted as realis, whereas the non-finite clause with a *to*-infinitive requires an irrealis interpretation. Certainly this seems to be the case for adult English, as indicated by the truth conditions for declarative versions of these sentences, given a context like (21).

(21)   Context: Mary says to Bill, "You should drink water."

       a Mary said what to drink - true no matter what Bill does.

       b Mary said what Bill drank - true only if Bill drinks water.

The question then becomes: do children know this contrast? A simple truth value judgment task would elicit this answer, modulo any problems children have giving adultlike responses in false belief tasks. However, both this and the de Villiers et al (2012) are imperfect studies of children's uses of RIs while acquiring English as they are studies of children's interpretations of embedded infinitival clauses licensed in the adult grammar, not of the conditions under which children will produce tenseless verbs in root clauses. Further, the test by de Villiers et al was performed with children ages 3 to 4, well above the age at which RIs typically subside, and it is not clear that either their task or my proposed task would be suitable for 2 year olds.

    An alternative would be to attempt to make an elicitation experiment where modality is somehow controlled for, so as to extract any Modal Reference Effect that might be lurking behind the nebulous interpretations of naturalistic utterances. In this case, it is presumably better not to have contrasts like in the context in (21), where there is both something drunk and something that ought to be drunk; rather, it should be clear the event of interest is either realis or irrealis. This is because, if the goal is to elicit either a realis or irrealis utterance, encoded under the above assumptions as either a tensed or tenseless

sentence, it should be absolutely clear which of the two is expected. Given the assumptions of, for example, Deen (2011), that the modal distinction is a primal one conceptually, this should be easily doable for 2 year olds, and the contrast between tensed/realis and tenseless/irrealis should come out in their utterances. (22) provides a couple of possible items for such a test.

(22)   a Irrealis

      – **Context**: Here's a ball, on the high shelf. Dora wants to play with the ball, but it's so high up! She can't grab it!

      – **Prompt**: Tell me about Dora and the ball

      – **Expected response**: Dora grab the ball

    b Realis

      – **Context**: Here's a ball, on the ground. Dora is playing with the ball. Look, she kicked it!

      – **Prompt**: Tell me about Dora and the ball

      – **Expected response**: Dora kicked the ball

This tests directly if there is in fact some sort of Modal Reference Effect in English, and could be done just as well for languages which have been purported based on naturalistic data to have such an effect. If children in English and other languages equally show MRE in this experiment, then it is a more reliable test than analysis of naturalistic data for modality in RI stages. If it happens that the conclusions from naturalistic data hold, *and* that English-learning children are sensitive to the modality of *to* infinitives, then we can conclude that the morphological idiosyncracies of irrealis in English are what blocks MRE in English RI stages. If, however, no MRE is shown from the elicitation test in English specifically, but children learning English understand modality in the *to* experiments, a better explanation will be necessary for why English do not show an MRE and while languages like Dutch and Tamil do.

5 Conclusion

This discussion has presented the prospects and challenges for an account of Root Infinitives that seems them as a way of encoding generalized irrealis semantics with slightly non-adultlike constraints on the Economy of Representation, rather than simply being a result of much more stringent constraints on Derivational syntax. This account is clearly imperfect as it stands, but the extensions described above could easily put it in much more solid standing, and provide a clearer relationship between the modality of non-finite clauses both cross-linguistically and between adult and child language.

References

de Villiers, J., T. Roeper, E. Harrington, and E. Gadilauskas. 2012. Tense and truth in children's question answering. *Proceedings of BUCLD 36.*

Deen, K. 2011. Mood Alternations. Manuscript: University of Hawai'i at Mānoa.

Deen, K., and N. Hyams. 2006. The Morphosyntax of Mood in early child grammar. *First Language* 26:67-102.

Hoekstra, T., and N. Hyams. 1998. Aspects of Root Infinitives. *Lingua* 106:81-112.

MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk. 3rd Edition.* Mahwah, NJ: Larwence Erlbaum Associates.

Pierce, A. 1992. *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars.* Dodrecht: Kluwer.

Raghavendra, P., and L. Leonard. 1989. The acquisition of agglutinating languages: Converging evidence from Tamil. *Journal of Child Language* 16.313-322.

Salustri, M., and N. Hyams. 2003. Is there an analogue to the RI stage in null subject languages? PROCEEDINGS OF BUCLD 26.

Salustri, M., and N. Hyams. 2006. Looking for the universal core of the RI stage. In Torrens and Escobar (eds.) *The Acquisition of Syntax in Romance Languages* 159-182. Philadelphia: John Benjamins.

Sarma, V. 1999. *Case, Agreement, and Word Order: Issues in the Syntax and Acquisition of Tamil.* Unpublished MIT Dissertation.

Wexler, K. 1998. Very early parameter setting and the Unique Checking Constraint: A new explanation of the Optional Infinitive stage. LINGUA 106:23-79.

Wexler, K. 2010. Grammatical computation in the Optional Infinitive Stage. In Roeper and de Villiers (eds.) *The Handbook of Generative Approaches to Language Acquisition* 53-118. Dodrecht: Springer.

Wijnen, F. 1996. Temporal reference and eventivity in Root Infinitives. *MIT Occasional Papers in Linguistics* 12.