From Root Infinitive to Finite Sentence

The acquisition of verbal inflections and auxiliaries

From Root Infinitive to Finite Sentence

The acquisition of verbal inflections and auxiliaries

Van Infinitiefconstructie tot Finiete Zin
De verwerving van werkwoordelijke inflecties en hulpwerkwoorden
(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. Dr. W.H.Gispen,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op woensdag 29 januari 2003
des middags te 14:30 uur

door
Wilhelmina Bernardina Theodora Blom
geboren op 4 december 1972 te Warnsveld

ACKNOWLEDGEMENTS

Where to begin? I'd like to think that it all started a long time ago when I was doing my high school exams and studied some novels from the Dutch author W.F. Hermans. Hermans' novels and essays were nice objects to analyse and they sketched a fascinating picture of academic life. They stimulated my urge to understand and hence, learn the methods of understanding. The battle of study advisors was won by Riekje Doeksen, who motivated me to take up the study of Dutch Language and Literature (and not psychology) at Utrecht University. Especially the teachers from the linguistics department excelled in sharing their own enthusiasm about linguistics with their students. Thus, I decided to concentrate on linguistics (and not medieval literature).

I was very lucky to have Frank Wijnen as my daily supervisor and guide through the vast fields of language acquisition research. His work has had great influence on this thesis, and in many cases his insights and theories are the point of departure for my investigations. I want to thank him for his suggestions and comments, patience and confidence. Paul van Geert was a supervisor at distance, both in kilometres as well as in research topic. Nevertheless, his influence shows in the statistical parts of this thesis. I hope they reflect Paul's enthusiasm and his clear way of dealing with statistical problems. Henriette de Swart joined the team of supervisors when I had already written the first versions of this book. At that time, I did not know how many versions would follow and she was wise enough not to let me know. Child language is not Henriette's main scientific concern, which made her into a perfect judge of the relevance of some questions I was wrestling with. She always managed to be very critical and motivating at the same time, so that I, after a meeting, jumped on my bicycle to rush to my computer and rewrite the entire manuscript.

I would like to thank the members of my 'leescommittee' - Sergey Avrutin, Peter Coopmans, Steven Gillis, Ger de Haan and Nina Hyams. A special thanks goes to Nina Hyams for her quick and detailed comments. The problems she pointed out did not make my life easier at certain moments. However, in the long run, her criticism led to many improvements. I am also very grateful to her

TABLE OF CONTENTS

"Next: how shall we define the whale, by his
obvious externals, so as conspicuously to label him
for all time to come? To be short, then, a whale is
a spouting fish with a horizontal tail."

H. Melville, Moby Dick

INTRODUCTION

Across languages, the first sentences of children are remarkably similar.
Children systematically omit finite verbs and use infinitives instead. In (1) -
(4), some examples are given from early child Dutch, French, Russian and
Hebrew:[1]

(1)   jij de walvis **maken**   Daan 2;04.28
       you the whale make-*inf*

(2)   pas **tomber** la poupée   Nathalie 2;2.2
       not fall-*inf* the doll
       'The doll does not fall'

(3)   papa peèku **topit'**   Zhenya 1;9
       daddy stove keep-going-*inf*
       'Daddy keeps the stove going'

(4)   tapuax **lishtot**   Lior 1;8.08
       apple drink-*inf*
       'I want to drink an apple'

---

[1] The examples (2) to (4) are taken from Ferdinand (1996), Brun et al. (1999) and Armon-
Lotem (1995), respectively.

The same characteristic that unifies many child languages all over the world reveals a difference between child and adult language.[2] Just to illustrate the contrast between children and adults: Daan, one of the children I examine in this thesis, used (1) where his parents would use sentences like (5) or (6):

(5)   jij **moet** de walvis **maken**
       you must-*aux* the whale make-*inf*
       'You have to make the whale'

(6)   Jij **maakt** de walvis
       you make-*fin* the whale
       'You are making the whale'

A comparison between (1), on the one hand, and (5) and (6), on the other, shows that Daan either leaves out a finite auxiliary or fails to inflect and move the verb. On the basis of this comparison, we could hypothesise that Daan does not know the forms to express finiteness (i.e. inflected main verbs and auxiliaries), and that he therefore is not able to use finite sentences.

In the last paragraph, I emphasised the difference between child and adult language: Daan uses sentences that his parents do not use. This conclusion is not entirely correct, though. In telling a story to the investigator who audio-taped his son Daan's father said:

(7)   Oeh hij **krijsen**
       oeh he scream-*inf*
       'And then he started to scream'

The sentence in (7) is not directed to Daan, but addressed to another adult. Thus, Dutch adults can use non-finite main clauses (commonly named *root infinitives* and abbreviated as *RIs*) that apparently have the same structural properties as Daan's utterance in (1). The adult RIs are restricted by certain

---

[2] In many different child languages, non-finite sentences are found. There are considerable differences between languages with regard to the length of the period in which children omit finiteness. This gave rise to the idea that there are 'RI-languages' and 'non-RI-languages' (Hoekstra & Hyams, 1995, amongst others). Romance languages like Italian or Spanish are examples of non-RI-languages.

discourse conditions, like embedding in a narrative context as in (7). It seems that children violate these rules and overuse RIs. Examples as in (7) give rise to the hypothesis that young children lack the knowledge of *where* to use finiteness, rather than of finite forms. Before we can draw the conclusion that children overuse RIs as the effect of pragmatic 'errors' (maybe due to processing limitations), we first have to know whether children are in command of the finite alternatives for RIs that would be felicitous. If these alternatives are not yet acquired, children are forced to violate discourse conditions by using RIs. Then, their use of RIs is not due to pragmatic deficits, but to the unavailability of finite forms.

For these reasons, the goal of this thesis is to learn more about children's early *grammatical* development. The aim is to explain the early absence of a finiteness system and its successive acquisition. The most appropriate empirical material for such a study comes from the developmental phase that begins when children combine their first words into two-word utterances and ends when children use multi-word sentences. I focus on two topics: (i) semantic properties of RIs (denotation and verb type) and (ii) the development of child language from RI-stage to finite-stage. Chapters 3 and 4 deal with temporal, modal and aspectual denotation and verb type, respectively, whereas Chapter 5 presents a detailed study of developmental changes in which special attention is paid to the aforementioned semantic aspects. Chapter 1, *Perspectives on Root Infinitives*, gives an overview of the main ideas and running themes in the RI-jungle. Chapter 2 describes the analysed data and the statistical methods applied to them.

In the literature, the semantic properties of RIs are taken as indicators of children's grammatical knowledge. The first relevant semantic aspect is the denotation of RIs, more specifically, their temporal, modal and aspectual denotation. The predominance of RIs that express desires or commands and the sentence-final placement of the infinitive could indicate that RIs are 'normal' adult sentences that contain a silent modal auxiliary (cf. (5)). Hence, it could be argued that children are grammatically fully competent. The observation that RIs differ in meaning in languages like Dutch and German, on the one hand, and English, on the other hand, motivates a second hypothesis: semantic properties of RIs relate to infinitival morphology. RIs

in Dutch and German contain 'true' infinitives with infinitival morphology, whereas English RIs contain bare stem forms:[3]

(8)   Dis **go** right there   Adam 3;3

I evaluate these two claims on the basis of analyses of spontaneous speech data from six Dutch-speaking children and experimental data from Dutch and English child language. We can find evidence showing that Dutch RIs are aspectually restricted, but there is no support for the claim that RIs have a pre-determined modal or temporal meaning. Moreover, it is confirmed that Dutch and English RIs differ in meaning.

Building on these observations, I explain the semantics of RIs by the initial absence and gradual acquisition of a system of finiteness marking, and the *Heterogeneous Set Effect*. The temporal and modal denotation of RIs is intrinsically free. As long as children do not know any other verb forms than the unspecified infinitive, RIs are frequent and used everywhere. When specified verb forms are learned, RIs become infrequent and are restricted to those contexts for which no specified alternatives are available (*Elsewhere Hypothesis*). The acquisition of specified alternatives for RIs is a grammatical process, since it results in the acquisition of auxiliaries, inflection and verb movement. Cross-linguistic differences are argued to follow from the *Heterogeneous Set Effect*: as an effect of inflection drop, English RIs comprise non-finite and finite sentences. This effect does not occur in Dutch.

The second semantic property of RIs, their restriction to certain verb types, received much attention in the literature. Abel - another child whose language I study in this thesis - uses his RIs for describing events that refer to something dynamic, such as *pakken* 'get', *doen* 'do' or *bouwen* 'build':

(9)   Koppie thee **pakken**   Abel 2;03.02
        cup tea get-*inf*
        'I am going to get a cup of tea'

---

[3] The examples are taken from Harris & Wexler (1996).

His finite sentences also contain state-denoting predicates:

(10) Ik **wil** jou   Abel 1;11.26
      I want-*fin* you
      'I want to sit by you'

It seems that RIs obey an *Eventivity Constraint*. In the literature, this constraint is approached in three different ways: as an effect of the absence of finiteness in RIs, as the result of the modal usage of RIs, or as an input effect. I argue that there is support for the second and third claim, but not for the first, even though the finiteness-contrast between (9) and (10) may suggest the opposite. The difference between verb type selection in Dutch and English RIs follows directly from my proposal that English RIs represent a heterogeneous set of sentences.

In Chapter 3 and 4, I supply empirical data and single out the hypotheses that are successful with respect to empirical coverage. Methodological questions regarding cross-linguistic comparisons will be discussed. I show how to deal with small datasets (an inherent problem of studies on early child language) and stress the importance of fine-grained analyses of empirical data.

Notwithstanding the relevance of these matters, it is Chapter 5 that constitutes the vital heart of this thesis. Despite the extensive amount of literature on children's early omission of finiteness, only little attention has been paid to children's *unlearning* of RIs. In Chapter 5, the incremental acquisition of grammatical finiteness in early child Dutch is described and interpreted. I formulate the *Grammaticalisation of Finiteness Hypothesis* (*GoF*), which states that the rise of finite sentences passes through three successive stages: (i) the appearance of simple lexical-finiteness markers, (ii) the emergence of combinations of lexical-finiteness markers and infinitives, and (iii) the acquisition of grammatical-finiteness markers in the form of inflected, moved verbs.[4] Though the GoF holds for Dutch, I am hesitant to draw any conclusions about other languages. Since it is argued that the developmental path of Dutch children is influenced by the way in which

---

[4] The first step is characterised by the appearance of sentences as in (10), the second step by the appearance of sentences as in (5), while the final step is marked by sentences as in (6).

grammatical information is packaged in the input, we first have to know more about the input patterns of other target languages. My hypothesis that the rise of finite sentences is directly linked to the disappearance of RIs is supported by the observation that RIs change their meaning at a certain point on the developmental path: the modal shift in RIs can be explained as an effect of the increase of finite equivalents for RIs. Patterns of subject use in RIs and finite sentences let us determine the point of time at which inflection comes in and children move the verb.

Elma Blom
Utrecht, Fall 2002

CHAPTER 1

Perspectives on Root Infinitives

In this chapter, the literature about RIs will be divided into three main categories. According to one view, RIs are one of the symptoms that characterise children's deficient grammatical knowledge. Studies that take this viewpoint, advocate the *reduced competence hypothesis*; the underlying idea is that children are born with reduced grammatical competence, as compared to adults. Other studies - most studies about RIs - defend the *full competence hypothesis*. According to this view, RIs are either a symptom of non-grammatical deficits or result from very small grammatical differences between child and adult grammar. A third type of approach concentrates on the language children hear, rather than on the linguistic competence of children. In these studies, it is argued that properties of RIs are *input-driven*. These three perspectives do not necessarily exclude each other. The first two are fundamentally different but the third input-oriented approach is, in principle, consistent with the other two views. I start the overview by introducing the two opposite perspectives in section 1.1. The difference between the perspectives hinges upon the question whether or not children move the finite verb already at a very early age. More background on verb movement in Dutch is given in section 1.2, while the sections 1.3 - 1.6 continue with discussions on the reduced competence analysis and full competence analyses of verb movement in early child Dutch (amongst other languages). In section 1.7, input-oriented studies will be discussed. Section 1.8 gives a summary, followed by an outline of the thesis in 1.9.

1.1   Two opposite perspectives

Several studies relate the omission of finite verbs to the omission of certain other words. Lebeaux (1988) and Radford (1988, 1990) argued that children systematically omit functional words. They use telegraphic speech, where finite verbs, determiners and complementisers are left out (the examples in (1) are from Radford, 1990):

(1)   a.   bear in chair
           the bear *is* in *the* chair
      b.   bow-wow go?
           where did the bow-wow go?

Such systematic omissions lead to the conclusion that the child's grammar lacks functional categories across the board. According to Radford, adult knowledge of functional categories has to mature. He claims that this happens all at once. Before this happens, the child uses lexical words that are connected to each other by thematic relations. Because functional structure is absent from this early representation, there are less syntactic positions than in the adult equivalent. As a consequence, a syntactic operation such as verb movement cannot take place. A verb, generated in V, cannot move to I to pick up inflection as there is no I; absence of the functional category I results in children's use of RIs.

In a nutshell, this is the maturational reduced competence hypothesis about language acquisition, and, in particular, verb movement. Around the same time as the studies of Lebeaux and Radford appeared - in the 1980's of last century - De Haan (1987) reported data from a Dutch boy, Tim, that indicated that the acquisition of verb movement involves more steps than was suggested by Lebeaux and Radford. Children do not jump from an RI stage without verb movement to an adult finite stage with verb movement; rather, they use RIs and finite sentences alongside each other for a while. In Wexler's (1992, 1994) terms, children go through an Optional Infinitive Stage, i.e. a stage in which finiteness is optional and RIs are used alongside finite sentences. A similar observation to the one which De Haan reported for Dutch, is reported by Poeppel & Wexler (1993) for the German-speaking boy, Andreas. Andreas (2;1) used the following sentences in one and the same session (examples are from Poeppel & Wexler, 1993). The

sentences in (2) contain finite *hab* and *mach* while (3) contains the infinitival form *haben*:

(2) a. Ich hab ein dossen Ball
         I have-*fin* a big ball
     b. Ich mach das nich
         I do-*fin* that not

(3) a. Thorsten Caesar haben
         Thorsten Caesar (= doll) have-*inf*
     b. du das haben
         you that have-*inf*

Observations such as these have puzzled researchers for many years as they seem contradictory: children use finite sentences, suggesting that knowledge of functional categories and verb movement is available to them, yet at the same time, they do not use finite sentences as frequently and systematically as adults do.

De Haan and Poeppel & Wexler dealt with the Optional Infinitive (or OI) stage in two entirely different ways. De Haan advocated a more or less reduced competence analysis, but explained language development not via maturation, as in the original version of the RCH that was explained before, but as a process of re-analysis that children apply. For this reason, I take De Haan's proposal as an exponent of the RCH that takes an *incremental* position with regard to the acquisition of grammatical knowledge rather than a maturational position. According to the maturational version of the RCH, functional categories mature all at once. According to the incremental version of the RCH, functional categories are accessed through an incremental growth of knowledge. In studies taking the second viewpoint, the existence of stages is emphasised; children's productions over time are described as successive steps in the acquisition of grammatical knowledge. De Haan argued that the early appearance of finite verbs and the early correct placement of these verbs do not necessarily imply that children move the verb. According to De Haan, two observations make it hard to believe that Tim knows the morphological and syntactic rules of finiteness: Tim's first finite verbs are not overtly marked for finiteness, hence there is no systematic morphological marking of finiteness yet. Moreover, the small set of lexical items that are placed in the position for finite verbs can be

analysed as *lexical* markers of finiteness. Thus, according to this account, the final stage in which finiteness is a grammatical property is preceded in early child language by a stage in which finiteness is lexical. De Haan's proposal contrasts with Poeppel & Wexler's full competence hypothesis (FCH). Poeppel & Wexler argued that children move the verb from early on. Andreas' data are presented in support of this claim. From a RCH-perspective, it is obvious why child language differs from adult language, and, hence, why children use RIs. From the FCH-perspective, children's use of RIs asks for an explanation. In the literature on RIs, three kinds of answers can be distinguished. These can be grouped under the headings 'Underspecification Hypotheses', 'AUX-drop Hypotheses' and 'Truncation Hypotheses'.

1.2   Verb movement

The controversy between De Haan and Poeppel & Wexler is about the question of whether young children move the finite verb form early on. Before I turn to the proposals to discuss them in more detail, I will first go into some details of verb movement in Dutch necessary for understanding the rest of this chapter. Consider the following difference between the placement of the finite and non-finite verb in adult Dutch. In Dutch declarative main clauses the finite verb is in second position, while the infinitive is placed sentence-finally:

(4)   Zij **dansten** de hele nacht
        they danced-*fin* the whole night
        'They danced all night long'

(5)   Ooit zullen Laura's dromen waarheid **worden**
        once will-*fin* Laura's dreams truth become-*inf*
        'One day Laura's dreams will become true'

The syntactic analysis of the sentences in (4) and (5) can be unified if both are derived from the same underlying structure. Verb movement is a way to unify the two structures. The finite verb moves out of its sentence-final base

position, while the infinitive remains in this position. Structurally, this can be represented as is illustrated below:

(6)   Jan danst de tango
        John dances the tango

(7)
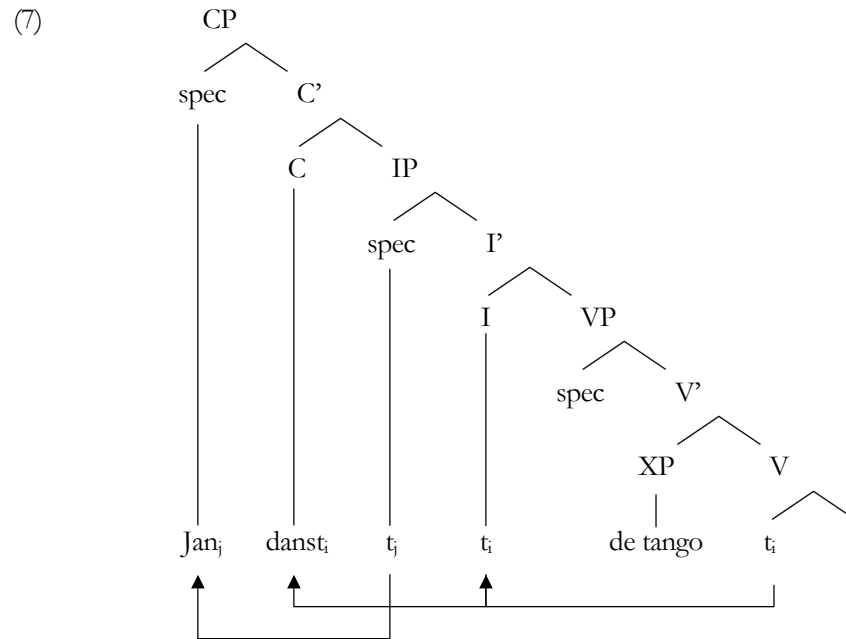
In the standard analysis of Dutch (Koster, 1975; Den Besten, 1983), the verb moves from V-to-I-to-C, as exemplified in (7). The bare lexical stem of the verb is inserted in V. The morphological marking of finiteness, inflection, is generated in I. Thus, the verb moves to I to pick up inflection. As inflection encodes tense, the inflected verb moves subsequently up to C to be in a proper tense position. Tense has scope over the entire sentence and is, therefore, related to the highest projection (CP).[1] Dutch does not only show verb second, but also has obligatory inverted order in, for instance, yes-no questions. In this case, the verb is in first position. In more

---

[1] This story holds for main clauses. In embedded clauses, C is filled with a complementiser. This blocks movement and the verb stays in its base position i.e. V.

recent analyses, the idea of invariant movement from V-to-I-to-C is replaced
with less rigid analyses (Zwart, 1993; Koeneman, 2000). Zwart (1993),
following Kayne (1994), furthermore argued that the Dutch base order is
SVO instead of SOV. Irrespective of the specific analysis, though, the finite
verb moves or is raised in overt syntax in Dutch main clauses.[2] The result of
verb movement is that the finite verb surfaces in the first or the second
position of a sentence, depending on the sentence type: the inverted VSO-
order is used in yes-no questions or conditionals, whereas the SVO-order is
used in the other sentence types. Both Zwart (1997) as well as Koeneman
(2000) describe the traditional analyses and give new analyses of verb
placement in Dutch.

1.3   Why children do not move the verb and use RIs

De Haan (1987) examined Tim's spontaneous speech data to investigate
whether or not verb placement in Dutch child language resembles verb
placement in early child German. This investigation was motivated by
Clahsen's work on child German (Clahsen, 1982; Clahsen & Smolka, 1986).
Clahsen, partially in collaboration with others, found evidence for the
following stages in the development of verb placement in child German:

I	25-29 months	There is no fixed order between the constituents of a sentence; all verbal elements occur in first/second and final position with a preference for final position; this also holds for verbal complexes

---

[2] In the discussion on RIs in child language, Dutch and German are often collapsed.
Though there are various differences between both languages, this seems appropriate for the
early use of verb forms by children that learn either language. As regards the analysis of adult
German, the analysis given for Dutch applies to German as well.

II	31-33 months	Verbal elements containing non-finite parts show up systematically in final position; finite verbs occur in both final and first/second position
III	37-39 months	All and only finite verbs are in first/second position; verbal complexes with finite and non-finite parts appear to be discontinuous
IV	40-42 months	In embedded sentences the finite verb is in final position; this occurs from the moment such sentences are acquired

Clahsen interpreted the finding that all verbal elements, including verbal complexes, can appear in first/second position as an indicator of early verb movement. To explain the initial random distribution of finite and non-finite verb forms, Clahsen proposed the Lexical Learning hypothesis (Clahsen, 1990; Clahsen & Penke, 1992; Clahsen, Eissenbeiss & Penke, 1994), which is another example of the incremental version of the RCH. Children do not have access to the entire set of functional projections; rather, they only have one FP (i.e. Functional Projection). When functional words are acquired, this FP is split into various specific projections. If children start out with only one general FP, there is no specific trigger for finite verbs and both finite and non-finite verbs can appear in sentence-initial position. However, Jordens (1990) has pointed out that the sentence-final finite verbs Clahsen mentions are, in fact, past participles with an omitted prefix. These forms are superficially indistinguishable from simple present tense forms. In general, the finding that children place the verb in an adultlike fashion (i.e. finite verb in initial position and non-finite verb in final position) is robust for German (Boser, Lust, Santelmann & Whitman, 1992; Poeppel & Wexler, 1993) as well as for Dutch (Verhulst-Schlichting, 1985; De Haan, 1987; Jordens, 1990; Wijnen, 1993). In stage II, children acquire the knowledge that only simplex verbs move. In stage III, an adult-like use of verb movement is attained. The fact that children do not move the finite verb in embedded sentences reinforces this conclusion.

Dutch is very similar to German, and therefore appropriate to test Clahsen's claims. Tim's data did not confirm the claims, however. When Tim (data collected at the ages of 2;01.15 and 2;02.15) was in stage I and II, he treated verbs syntactically already in a very restricted manner: finite forms are in sentence-initial position while non-finite forms are sentence-final. Thus, *only* finite verbs are fronted. Systematic marking of tense or agreement is lacking; only the finite-nonfinite distinction is visible in morphology on the verb. De Haan writes:

> "This boils down to the postulation of two verbal categories that are morphologically, semantically and syntactically distinct." (p. 21)

These two categories are AUX and V, given in (8) and (9) respectively:

(8)   a.   gaat niet
           goes-*fin* not
           AUX - neg
      b.   moet daar in
           must-*fin* there in
           AUX - adv - prep

(9)   a.   ik een keertje doen
           I one time do-*inf*
           subj - obj - **V**
      b.   even buiten kijken
           just outside look-*inf*
           adv - adv - **V**

AUX does not carry infinitival morphology and is used to express tense and modality. V, on the other hand, captures notions like 'act' or 'change'. V is morphologically distinct from AUX because of its infinitival morphology: [*-en*]. Syntactically, AUX and V are generated in different positions.

In summary, De Haan says that children do not classify verbs in the same way as adults. What guides children towards a *non*-adult classification? The child may be mislead by the input because the input does not contain explicit information about grammatical classes - the child is not told what a verb is, or a noun - and the syntactic classification must be derived from cues in the input data. Pinker (1984) hypothesised that children apply

semantic bootstrapping to detect syntactic classes. His idea is that children have early access to a number of semantic *observable* primitives. 'Action' or 'motion' are such primitives. Through linking rules, this primitive is linked to the syntactic category V. In this system, predicates that do not denote activities will initially not be classified as V. From this perspective, it is understandable that children end up with two different categories: the predicates that denote tense and modality (non-action predicates) are assigned a different grammatical category than action-denoting predicates. To arrive at one, adultlike, class V that contains all verbal predicates, the children have to merge the two categories. According to De Haan, finite morphology helps the child to generalise. Only if children have access to the finite paradigm, can they deduce that not all elements of the class V denote activities. Knowledge of finite morphology merges the categories AUX and V. This unification is a prerequisite for verb movement.

1.4   Why children *know* verb movement ...

Poeppel & Wexler (1993) made a strong claim against the *No-Overlap hypothesis*, as they call De Haan's approach. Recall that De Haan claimed that different types of verbs appear in RIs and early finite sentences and hence, predicts that there is no overlap between the verb types used in RIs and finite sentences. From their study of Andreas' (2;01) data, Poeppel & Wexler report that out of the 28 verbs that are used twice or more, 8 verbs appear in finite as well as non-finite form. Poeppel & Wexler do not give the list of verbs. The 20 remaining verbs that do not overlap do not fit De Haan's classification, but is is clear from the text that the finite verbs are not only modals such as *können* 'can' and *mögen* 'want' or the copula *sein* 'be', but they denote also activities like *fliegen* 'fly' or *umkippen* 'fall over'. The non-finite forms are nearly all verbs that denote actions. The predicate *stehenbleiben* 'stand still' that appears in non-finite form in Andreas' data, is ambiguous: it can denote an action as well as a state or a property.

Structural properties of Andreas' finite sentences, and differences in structural properties between RIs and finite sentences, are taken by Poeppel & Wexler as evidence for early verb movement. First of all, finite verbs and non-finite verbs appear in different positions. The former appear in

sentence-initial position while the latter are sentence-final. There are 197 sentences with the finite verb in second position; 180 contain an overt subject and in 50 examples, the finite verb is preceded by a nonsubject i.e. object or adverb. Thus, Andreas does not stick to the canonical subject-verb order and it seems that he is able to topicalise constituents. He does this correctly, that is, in accordance with adult German: in all cases, the verb is located in second position. As regards the morphological properties of the early finite forms, Poeppel & Wexler's study remains unclear. Tense is restricted to present tense, which is unmarked. Hence, there is no indication that Andreas has acquired any morphological tense distinctions. As regards agreement morphology, Poeppel & Wexler reported that:

> "Andreas used predominantly singular subjects with correct agreement morphology on the verb." (p. 6)

Based on this, Poeppel & Wexler concluded that Andreas knows agreement because he hardly makes any errors. All in all, their conclusion is that Andreas, at the age of 2;01, has an adult-like sentence representation and full grammatical competence: he moves the finite verb and uses the I and the C domains.

1.5   ...but still use RIs

Poeppel & Wexler's conclusion that Andreas had full grammatical competence was based on Andreas' finite utterances. However, Andreas' German still differed from adult German, because he used RIs. Why do children use RIs even though they know verb movement? In this section, three traditional full competence explanations for RIs that deal with this question will be summarised.

1.5.1   Underspecification Hypotheses

The common FCH explanation is that the child grammar is complete but *underspecified*. This idea relies on the minimalist program as proposed by Chomsky (1993). The minimalist account of verb movement differs from

the Government and Binding analysis (§ 1.2) in several respects. In the GB model, the lexical stem in V moves to I to pick up inflection, as inflection is generated in I. In the minimalist analysis, V is not occupied by a bare stem but by a fully inflected form. This form contains tense and agreement features, that project functional structure: Agr(eement)P and a T(ense)P. The functional heads are occupied by inflectional features and not by inflection itself (as in the GB framework). The features of the verb in V must match the features in the functional domain; the verb moves to the functional positions in order to check its features. Movement takes place in overt or covert syntax, depending on the strength of the feature. Strong features require movement in overt syntax (prior to spell-out), while weak features lead to procrastination of movement (until LF). In an underspecified grammar, V projects, even though the inflectional features are not specified. As an effect, FPs are projected (instead of specified projections such as AgrP or TP). The underspecified FP does not attract the verb and hence, movement is not triggered.

UNDERSPECIFICATION OF TENSE   Recall that one of the problems of the RCH proposals of Radford (1988) and Lebeaux (1990) was the co-occurrence of RIs and finite sentences. Poeppel & Wexler (1993) did not solve this problem. Wexler (1992, 1994) concentrated on children's optional use of finiteness and argued that children pass through an *Optional Infinitive Stage*. According to him, the characteristics of OI stage are that both finite and non-finite verbs appear in the root clause. The finite and non-finite verb forms appear with a different word order: raised and not raised, respectively. This pattern is in accordance with the adult grammar. Wexler claimed that in this stage, Tense is underspecified. Thus, verbs do not move for tense; the only trigger for verbs is agreement. In finite sentences, the verb is raised, resulting in a form that is agreeing. In RIs, the verb is lowered, resulting in a form that is not agreeing.[3] Wexler illustrated his claims with data from French taken from (Pierce, 1989, 1992). Data from French child language show that young French-speaking children have moved or raised the finite verb, while the verb in RIs is unmoved (or lowered, in terms of Wexler). The French examples below illustrate that a verb that precedes the negation

---

[3] Wexler needed these two options because, following the minimalist framework, he assumed that optionality only occurs if two derivations are equally costly. Therefore, in finite sentences, as well as RIs, the verb must move.

pas (hence moved) is finite, while a verb that follows *pas* (hence unmoved) is an infinitive (examples are from Pierce, 1992).

(10) a. marche **pas**
          walk-*fin* not
      b. ça tourne **pas**
          this turn-*fin* not
(11) a. **pas** casser
          not break-*inf*
      b. **pas** rouler en vélo
          not roll-*inf* on bike

The child escapes from the OI stage by specifying Tense. RIs do not occur anymore as the acquisition of forms to mark temporal distinctions leads to the obligatoriness of Tense marking.

Besides French, Wexler discussed the optional infinitive phenomenon in a wide range of languages. He claimed that, although English has no distinct infinitival form (unlike the other languages Wexler mentions), English children do go through an optional infinitive stage, just like Dutch, German, French, Danish, Norwegian and Swedish children. He argued that uninflected verb forms, illustrated in (12), are actually infinitives. In English, these infinitives happen to be indistinguishable from stems and all persons in the simple present tense paradigm except third person singular[4]:

(12) a. Eve sit floor   Eve 1;7
      b. Where penny go?   Adam 2;4.30
      c. That truck fall down   Nina 2;0.24

---

[4] An elaborate discussion of the OI stage in child English can be found in Harris & Wexler (1996). One of the results presented in support of Wexler's claim is that English children do use (ia) while (ib) is unattested:

(i)   a.   She not go
      b.   * She not goes

If the verb form in (ia) is nonfinite, it is predicted that it does not move as there are no finiteness feature to check. The fact that *go* follows negation indicates that this prediction is borne out. English children distinguish untensed verb forms syntactically from tensed verbs; they move the tensed verb to check its features. As expected, (ib) does not appear.

UNDERSPECIFICATION OF NUMBER   Hoekstra & Hyams (1995) and Hyams (1996, 1999) proposed an alternative underspecification hypothesis: instead of Tense, Number is underspecified. Underspecification of Number leads to absence of Tense, according to them. Cross-linguistic patterns are presented in support of this claim. In Germanic languages, such as Dutch, German and English, RIs are frequent although they hardly appear in a Romance language such as Italian.[5] The two language types differ in the richness of their inflectional paradigms. Italian has person as well as number agreement in the present tense paradigm, whereas the Germanic languages only mark Number i.e. the difference between singular and plural.[6] Hyams (1999) states that in languages of the Germanic type 'number is the morpho-syntactic reflex of tense binding' (p. 402). If in these languages Number is underspecified, Tense cannot be bound: a verb form that has no morphological tense features surfaces, namely the infinitive.[7]

---

[5] Data on Italian can be found in Schaeffer (1990) and Guasti (1992).

[6] This generalisation does not follow straightforwardly from the inflectional paradigms in Germanic languages. In order to make this claim, Hoekstra & Hyams (1995) have to start with a set of assumptions. First, they assume that the first person singular may be unmarked. This yields the conclusion that the [-t] extension found with second and third person singular in Dutch (and not with first person singular) marks Number rather than Person. Second, in order to explain early child German, Hoekstra & Hyams argue that German is initially misanalysed by German-learning children as being Dutch-like (the German children fail to detect the Person marking initially). Third, Hoekstra & Hyams follow Kayne (1989) with regard to the assumption that *you* in English is grammatically plural. This yields the conclusion that the [-s] extension found with third person singular in English marks Number.

[7] Tense is thought of as being anaphoric. It relies on an antecedent that is given in the preceding discourse for its interpretation. Partee (1984) observes similarities between temporal and nominal anaphora. In particular, she points out that in both cases the antecedent can remain implicit. In (ia) the temporal reference is not specified while in (ib) it is the nominal reference which is not specified:

> (i)   a.   I didn't turn off the stove
>   b.   She left me

The distinction in anaphoric and pronominal Tense is used in Gueron & Hoekstra's (1989, 1996) analysis. If Tense is anaphoric, it is bound to a tense operator which has the unmarked value of 'here-and-now' or speech time S (co-indexed). If Tense is pronominal, it is free and has the value 'past tense' (contra-indexed). The temporal index is realised by morphology e.g. a past tense suffix. If Tense is underspecified, it is not indexed, so it is neither co-indexed nor

Underspecification of Number prevents children (acquiring languages that have Number agreement) from using finite sentences. As long as Number is underspecified, RIs surface. To explain the OI stage itself and to account for the shift to a finite stage, Hyams (1999) gives an additional explanation. In the OI stage, children have two ways to bind Tense. In RIs, they bind tense deictically, that is to speech time S. In finite sentences, Tense is bound grammatically through a tense chain. Children escape from the OI stage when they learn that Tense, being anaphoric, must be bound grammatically. As long as the general pragmatic principle that prescribes grammatical binding is not acquired, finiteness is optional for children.[8] Avrutin (1999) connects to Hyams' more recent work by locating the RI-phenomenon at the syntax-discourse interface. According to Avrutin, children know the grammatical properties of Tense. They lack, however, the processing capacities required to obey the pragmatic conditions on Tense. Consequently, children fail to use Tense for introducing events in the discourse.

AGREEMENT AND TENSE OMISSION MODEL (ATOM)   A third underspecification account is proposed by Schütze & Wexler (1996) and Schütze (1997). The difference between this and the other two underspecification accounts (the Underspecification of Tense and Underspecification of Number approach) is that it is underspecification of two functional features, Tense and/or Agreement, instead of one that leads to RIs. Support for this proposal comes from subject Case errors in English child RIs. In adult English, the subject must have nominative Case. Children, however, use RIs with accusative subjects. The examples in (13) are from Schütze (1997):

(13)   a.   **him** fall down   Nina 2;3.14
       b.   **her** have a big mouth   Nina 2;2.6

It is argued that Tense and Agreement both bear a relation to the subject: Tense licenses overt subjects, while the subject's Case is assigned by

---

contra-indexed. There are no morphological tense features on the verb. In this analysis, Tense forms a chain between the event in V, the finite verb in I and the tense operator in C (Evers, 1981; Enç, 1987; Gueron & Hoekstra, 1989, 1996).

[8] Hyams (1999) refers to the work of Grodzinsky & Reinhardt (1993) for an example of such a principle stating that the deictic binding option is blocked when grammatical binding yields the same interpretation as deictic binding.

Agreement. As a result, subject drop is only allowed when Tense is underspecified, because the subject is not licensed. The subject can have a wrong Case if agreement is underspecified. This surfaces only in languages where the default Case is non-nominative. As accusative is the default Case in English, it is correctly predicted that English children use RIs with accusative subjects.[9]

1.5.2   The AUX-drop Hypothesis

The idea behind the AUX-drop hypothesis, initially proposed by Boser, Lust, Santelmann & Whitman (1992), is simple. In principle, RIs are equivalent to adult sentences with periphrastic verbs. The only difference between children's RIs and adult sentences with periphrastic verbs is that children do not realise the auxiliary phonetically. The auxiliary is dropped and an RI remains. The null auxiliary is recognisable through cues from morphology, syntax and semantics. The covert auxiliary corresponds to a modal or dummy verb that does not have any lexical content. A more recent interpretation of the AUX-drop hypothesis, Ferdinand (1996) argued that the null auxiliary is specifically a modal or future auxiliary as this would offer an explanation for the observation that RIs are restricted to eventive predicates. In § 1.6.1, Ferdinand's proposal will be discussed in more depth. Note that final placement of the verb in RIs follows automatically from the fact that the null auxiliary fills I. The main verb cannot move from V to I and remains in base position.

The AUX-drop hypothesis has been criticised for several reasons. First of all, there is a methodological problem: to drop the auxiliary in periphrastic constructions children must be able to use overt periphrastic constructions. However, in the sequence of developmental stages, periphrastic verbs are acquired after the RI stage (Jordens, 1990; Wijnen, 2000). Secondly, it is unclear why children should drop the auxiliary in the first place, assuming that they can, in principle, produce it. Thirdly, there are various empirical problems. If RIs were like 'normal' finite sentences, then it is expected that

---

[9]   In a language like Dutch, in which nominative Case is the default Case, no such errors are predicted.

RIs pattern like finite sentences. However, both topicalisation as well as WH movement are unattested in RIs, whereas they are found in children's early finite utterances. Poeppel & Wexler (1993) pointed to topicalisation facts in the finite sentences of Andreas (see: § 1.4). WH asymmetry between RIs and finite sentences is observed in Dutch, German, Swedish and French child language (see Phillips, 1995 for an overview).[10, 11] Another point of criticism concerns the licensing of the null auxiliary. As an empty category, the null auxiliary needs licensing. Boser et al. claimed that an overt subject licenses it. One of the characteristics of RIs, however, is that the presence of a subject is optional (Hyams & Wexler, 1993). Thus, it remains unexplained how the null auxiliary is licensed.

Some objections can be raised against the empirical criticism. The observation that children do not topicalise constituents and that WH words are absent in RIs does not necessarily imply that children cannot use them. It is a common research strategy to make this assumption though. A more serious problem of the empirical criticism is that both the absence of topicalisation facts and WH words in RIs can be seen as effects of development: topicalisation as well as WH words may surface in RIs and not in finite sentences, because RIs are acquired before finite sentences come in. Thus, to examine whether or not this criticism holds, longitudinal data need to be analysed, so more can be said about the patterns that emerge over time.

1.5.3   The Truncation Hypothesis

Although the truncation hypothesis can be seen as less full competence than the other two hypotheses (AUX-drop and underspecification), in the sense that the child grammar allows for representations that are not allowed by the adult grammar, it still claims that children do have access to the entire set of functional projections and know verb movement. In this way, the truncation hypothesis assumes full grammatical competence from early on.

---

[10] Phillips (1995) for English, Haegeman (1994) for Dutch, Kursawe (1994) for German, Santelmann (1994) for Swedish and Crisma (1992) for French

[11] Interestingly, English child language does not show both of these asymmetries, whereas the other child languages do (Roeper & Rohrbacher, 1994; Phillips, 1995).

The central tenet of the Truncation Hypothesis is that children lack the knowledge that C heads every sentence. According to Rizzi (1992, 1994), this knowledge matures. According to Weissenborn (1994), however, truncated trees follow from reduced processing resources. Children do not have to project the entire adult structure, though they have to obey the canonical order of projections that is given by UG. Weissenborn (1994:216) formulated this as the Local Well-Formedness Constraint, which requires that the representation of any utterance of the child is locally wellformed with respect to a representation of the adult grammar. Thus, children cannot omit material from the middle of a syntactic tree; only 'top' nodes can be dropped. The resulting trees are called 'truncated trees' (Rizzi, 1992, 1994; Weissenborn, 1994; Haegeman, 1995) and the root of a sentence can be CP or any projection below CP. If the child has only one grammar, the root principle must look like (14):[12]

(14)   VP or TP or CP is root

The optional use of infinitives follows from the optional projection of TP and CP. In the case of RIs, children take the VP as root:

(15)

```
                VP
              /    \
          spec      V'
           …       /  \
                  …    V
```
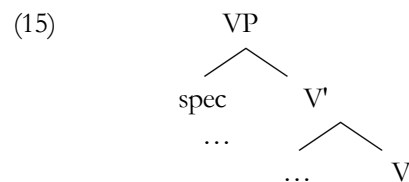
In the case of finite utterances, either TP or CP is root. Rizzi (1994) argued that RIs in Italian are rare (Guasti, 1993) and that this is an effect of feature strength. In both the tensed and untensed paradigm, Italian agreement is strong. Thus, following the ideas of the minimalist program (Chomsky, 1993), Rizzi (1994) proposed that in Italian tensed as well as untensed verbs move as high as AgrP, because they are both triggered by strong agreement features.

---

[12] One could also hypothesise that the child has different grammars: one with CP = root, one with TP = root and one with CP = root. The difference between this idea and Rizzi's formulation is purely conceptual. Formally, there is no difference between a single grammar that contains the unspecified rule in (14) or three co-existing grammars with specified rules.

1.5.4   Where hypotheses diverge: subject drop

The Underspecification, AUX-drop and Truncation Hypothesis differ most clearly with respect to their predictions concerning patterns of subject use in RIs and finite sentences. I already mentioned that the AUX-drop hypothesis predicts that RIs must contain overt subjects. Wexler & Hyams (1993) pointed out that this prediction is not borne out. Observations from Dutch and German child language, reported by Krämer (1993), and from Danish child language, reported by Hamann & Plunkett (1998), indicate that RIs show a clear preference for subject drop and do not appear with overt subjects most of the time. Both the Underspecification and the Truncation hypothesis account for this observation and for the observation that finite sentences used alongside RIs do not share the preference for subject drop. Both hypotheses take asymmetries between RIs and finite sentences as their point of departure, as opposed to the AUX-drop hypothesis (according to which RIs and finite sentences are structurally identical). According to both hypotheses, RIs are untensed clauses that contain a non-agreeing verb form. In this respect, RIs differ fundamentally from finite clauses. With regard to their specific ideas about subject drop in the OI stage, the Underspecification and Truncation hypothesis differ, however.

Underspecification accounts transfer the analysis of adult infinitival clauses to child RIs (Sano & Hyams, 1994; Wexler, 1994; Schütze & Wexler, 1996; Schütze, 1997). It is argued that RIs contain a null subject, and that this null subject is identified as PRO ('big pro'). Why is the subject PRO and is the child not allowed to simply drop the subject? Chomsky (1986) argued that UG contains a principle stating that all sentences must have subjects. This principle, known as the Extended Projection Principle or EPP, explains (among other things) the insertion of the expletive subjects in sentences like (16):

(16)   a.   **There** is a man in the garden
         b.   **It** is fun for Jane to play badminton

The EPP makes direct predictions for the analysis of subjectless sentences. To obey the EPP, it must be assumed that (17) contains a null subject: a subject that is structurally present but does not have a phonetic representation.

(17)   It would be wrong [NULL SUBJECT to leave now]

As the null subject in (17) has a syntactic function but is not phonetically realised, it is an empty category. In generative theory, there are different types of empty categories, each with their own licensing conditions. According to Chomsky's theory, there is only one empty category that must remain ungoverned, namely PRO.[13] Infinitival clauses lack a finite verb, hence I is empty and spec, IP (the canonical subject position) is ungoverned. Thus, PRO is licensed in this position. In child RIs, I - Tense, Agreement or both - is/are underspecified. Therefore, I cannot license an overt subject. As in adult infinitival clauses, the only subject that is allowed is PRO. In sum, Underspecification accounts concentrate on the I domain. They predict that RIs contain null subjects, while finite sentences appear with overt subjects.

Truncation hypotheses focus on the absence of a C domain in child language. According to these accounts, the dropped subject in child language is not PRO, but it is a 'null constant', which is allowed to be antecedentless by virtue of the absence of an antecedent position.[14] Spec CP is an antecedent position, thus absence of CP leads to absence of an antecedent position. The null constant is allowed in the canonical subject position, i.e., spec IP. Truncation accounts predict that RIs have a null subject, while finite sentences have null subjects and overt subjects. Note that the Underspecification and Truncation accounts make different predictions for the patterning over time. According to Underspecification, RIs will show a constant longitudinal pattern of subject drop while finite sentences show a constant pattern of subject use. According to Truncation, a stage with optional subject use in finite sentences is followed by a stage with obligatory use of subjects in finite sentences.

The Agreement and Tense Omission Model, or ATOM, (Schütze & Wexler, 1996; Schütze, 1997) differs from the other Underspecification proposals in

---

[13] The distribution of PRO has puzzled many researchers over the years and has lead to different proposals. In the original PRO theorem, PRO is a pronominal anaphor (Chomsky, 1981). Alternative accounts have been proposed. It is argued that PRO is a pure anaphor, a pure pronoun or sometimes anaphor/sometimes pronoun. I will not go into the details of this discussion but a recent overview can be found in Petter (1998), for example.

[14] Note that this is a stipulation; no further support is given for such a null category.

that it allows for overt subjects in RIs. It is predicted that RIs show optional subject use while finite sentences only contain overt subjects. The ATOM predicts furthermore that subject Case errors in Dutch child RIs will not appear: overt subjects in RIs do not appear with non-nominative Case. In the available literature, the picture of subject Case in Dutch child RIs is heterogeneous. According to Schaerlakens & Gillis (1987), Dutch-speaking children go through a short phase in which they use the accusative form *mij* 'me' instead of nominative *ik* 'I'. Van Ginneken (1917) reported a similar overuse of subjects with accusative Case:

(18)   Mij moet et hebbe   Keesje 2;7
       me must it have-*inf*

(18) is not a RI, though. Boezewinkel (1995) found accusative subjects as well; in her data, the accusative subjects only appear in RIs:

(19)   a.   mij doen   Laura 2;3
             me do-*inf*
       b.   mij zitten   Sarah 2;0
             me sit-*inf*

Kaper (1976), Powers (1994) and Bol & Kuiken (1986) found hardly any examples of accusative subjects in RIs, if any at all. It seems that the use of RIs with accusative subjects is idiosyncratic. In general, children make hardly any Case errors, as predicted by the ATOM.

1.6   The No Overlap Hypothesis revisited

Recall that I started this chapter with the observation from De Haan (1987) that children use different verb types in their early finite sentences and RIs: there was no lexical overlap between the two types of sentences. De Haan argued that this no-overlap pattern indicates that Tim is not able to move the verb. Contrasting this analysis with full competence proposals, leads to the question how full competence accounts deal with the no-overlap patterns during the OI stage. In this section, four full competence explanations for De Haan's observation will be discussed. These four

explanations differ from each other in two respects. Firstly, the explanations differ with regard to whether or not they allow for overlap. Ferdinand (1996) argues, like De Haan, that there is no overlap: there is a class of verbs types that appears in finite forms and there is a different class of verbs types that appears as infinitives in RIs. A verb type that belongs to one class cannot belong to the other class. Wijnen (1997), Hoekstra & Hyams (1998) and Gavruseva (2001, 2002) argue for an *asymmetry* between finite sentences and RIs: verb types in RIs are restricted to a particular class, but the verbs types that appear in simple finite sentences are not. Thus, the prediction is that all verbs that appear as infinitives in RIs may occur as finite verbs, but not the other way around. The full competence studies formulate the difference between verbs in RIs and finite sentences in terms of aspectual classes (Vendler, 1967; Dowty, 1979) and not in terms of grammatical classes as De Haan did.[15] Explanations are based on the incompatibility of certain aspectual properties of verbs and properties of RIs. A second difference between the four full competence explanations will be discussed in this section. It concerns the types of aspectual classes. Two types of hypotheses can be distinguished. According to the Dynamicity Hypotheses, RIs are restricted to [+dynamic] or eventive predicates while finite sentences (also) contain [-dynamic] predicates or states (Ferdinand, 1996; Wijnen, 1997; Hoekstra & Hyams, 1998). According to the Telicity Hypothesis, RIs contain predicates that are not inherently specified for telicity, i.e. [αtelic], while finite sentences contain predicates in which a [+telic] or [-telic] meaning is part of the verb's lexical content (Gavruseva, 2001, 2002).

1.6.1   Dynamicity Hypotheses

FERDINAND (1996)   Ferdinand's (1996) work was briefly mentioned when the AUX-drop hypothesis was portrayed. Since Ferdinand used the AUX-drop hypothesis, more specifically, the MODAL-drop hypothesis, in

---

[15] Note that the lexical items in De Haan's two grammatical classes (AUX and V) overlap for a great deal with the lexical items in the different aspectual classes. The main difference is that De Haan predicts a smaller set of different verbs to appear as early finite verbs that any of the other studies does. Details on this issue can be found in Chapter 4.

order to explain no-overlap patterns in child French, I have chosen to discuss her proposal here. Ferdinand found that children in the age-group she studied use overt auxiliaries that select for infinitival predicates with a future ('inchoative' as Ferdinand calls it[16]) or deontic meaning.[17] She illustrated this with the following examples:

(20) a.   'tacha va met   Nathalie 2;2.2
            Natacha go-*fin* put-*inf*
            'Natacha is going to put it there'

       b.   va chercher maman   Gregoire 1;11.20
            go-*fin* look-for-*inf* mummy
            'I am going to look for mummy'

(21) a.   veux fermer   Daniel 1;8.1
            want-*fin* close-*inf*
            'I want to close it'

       b.   poupee doit faire dodo   Nathalie 2;2.2
            doll have-*fin* to do sleep-sleep-*inf*
            'The doll has to sleep'

Ferdinand claimed that these auxiliaries are dropped in RIs. Since these auxiliaries are incompatible with states, this explains why RIs are restricted

---

[16] Ferdinand chooses the label 'inchoative' to make clear that *va* does not have a temporal meaning. However, future is modal rather than temporal. As future is a better way to describe the meaning of *va* (and it is the standard label in the literature) I refer to *va* as a future auxiliary.

[17] Ferdinand includes volitionary *vouloir* 'want' and dynamic *pouvoir* 'be able to' in the set of deontic modals.

to eventive predicates, according to Ferdinand.[18] Examples from adult French were used to illustrate this claim. When the future auxiliary selects a stative infinitive, the predicate becomes eventive: according to Ferdinand, (22a) denotes the beginning or inception of a state and is, therefore, eventive rather than stative. If a modal selects for a stative infinitive, the sentence receives an epistemic reading. Ferdinand's example in (22b), taken from a French dialect, is used to illustrate this.[19]

(22)   a.   Jean va être malade
                Jean go-*fin* be-*inf* ill
                'Jean is going to be ill'

        b.   il veut faire beau
                it want-*fin* make-*inf* beautiful
                'It is probably going to be good weather'

As children in the RI stage do not use epistemic modals, stative predicates do not appear in their RIs. Eventive predicates are absent from children's early finite sentences because tense is still underspecified: the early finite forms are just [+tense] and not further specified as [-past] or [+past]. The value of tense is determined by speech time S, which yields a here-and-now or present tense reading. Ferdinand assumed that only stative

---

[18] There are more convincing examples that illustrate the relation between modal interpretations and predicate types than the French examples Ferdinand gives, for instance, the 'minimal pair' *weten* 'know'/*leren* 'learn' given by Hoekstra & Hyams (1998) to illustrate that stative predicates force an epistemic reading of the modal auxiliary (i), while with eventive predicates a deontic reading is preferred (ii). The lexical meaning of the two predicates is quite similar, but their dynamicity values (or rather their agentivity values, as I will argue in Chapter 4) differ. :

        (i)   Jan moet het antwoord weten
                  John must the answer know
                  'I conclude that John knows the answer'
        (ii)   Jan moet Frans leren
                  John must French learn
                  'John is required to learn French'

[19] The observation that a relation exists between epistemic modals and stative predicates should not be attributed to Ferdinand but is mentioned in earlier studies, for instance, Steedman (1977), McDowell (1987) and Barbiers (1995).

predicates can survive with an underspecified tense form. Ferdinand (1996: 88) says about this assumption:

> "Non-eventive verbs lack internal temporal structure, they denote states. This entails that they can be [+tense] without being linked to a specific part of the time axis. Eventive verbs, on the contrary, have internal temporal structure; they denote changes taking place in time. For this reason, whenever an eventive verb is marked for [+tense] it must be linked to a specific moment in time."

In sum, Ferdinand provides two explanations: one for the verbs that appear as early finite verbs - these can survive the underspecified tense - and another explanation for the verbs that appear in RIs - these are compatible with the underlying modal auxiliaries in RIs -. The result is that a strict no overlap pattern is predicted. In this respect, Ferdinand's proposal differs from the other three full competence proposals that I will discuss in the remainder of this section.

WIJNEN (1997)   Wijnen (1997) applied Carlson's (1979) distinction between stage and individual-level predicates to explain why stative predicates are absent in RIs. He used Kratzer's (1989) analysis of stage and individual-level predicates to account for patterns in early child language. Stage-level predicates denote temporal properties while individual-level predicates denote permanent properties. Thus, eventive verbs are a subset of stage-level predicates while stative verbs are a subset of individual-level predicates. Semantically, Kratzer (1989) describes the difference between stage and individual-level predicates in terms of an event variable: stage-level predicates (i.e. events) contain this variable while individual-level predicates (i.e. states) do not. There is also a syntactic difference between these two types: subjects of stage-level predicates (i.e. events) originate in the specifier of VP while subjects of individual-level predicates (i.e. states) are generated in spec IP.

Wijnen argued that the absence of stative predicates in RIs shows that Tense is absent in RIs. In tensed sentences, the event variable is bound by Tense. In RIs, an alternative strategy is applied. As Tense is absent, the event variable is bound deictically. Stative predicates do not have an event

variable.[20] Therefore, a temporal interpretation cannot be assigned deictically. As a result, stative predicates are (necessarily) overtly marked for Tense while eventive predicates can remain untensed.[21] Under the assumption that the event variable is bound to speech time by a deictic operator, it is predicted that RIs are restricted to present tense interpretations. But Wijnen found that the RIs of four Dutch children are most often modal or future (Wijnen reported an average of 86% RIs that were modally used), just like the RIs produced by German children (Ingram & Thompson, 1996). To account for this finding, Wijnen hypothesised that a modal or *irrealis* interpretation surfaces as the default. This is the meaning that is assigned in absence of Tense. According to Wijnen, a lack of Tense provides furthermore a syntactic way to account for the absence of stative predicates in RIs. Kratzer's syntactic analysis of stage and individual-level predicates implies that stative predicates need an IP layer to host the tensed verb. In RIs, however, this layer is absent and stative predicates cannot appear. This analysis is compatible with Truncation as well as Underspecification Hypothesis discussed earlier, as the claim of these two theories is that Tense is optional in child language and Tense is absent in RIs.

HOEKSTRA & HYAMS (1998), HYAMS (2001A, B)   In their underspecification account, Hoekstra & Hyams (1995) and Hyams (1996, 1999) argued, like Wijnen (1997), that children bind tense deictically in RIs. Findings reported in various studies suggest that RIs have a preference for modal interpretations, however (Krämer, 1993; Ingram & Thompson, 1996; Wijnen, 1997). To account for the modal interpretation of RIs, Hoekstra & Hyams

---

[20] Most individual-level predicates are stative and most stage-level predicates are eventive, but there are stage-level stative predicates. Dowty (1979) gives the following examples:

(i)   a.   New Orleans lies at the mouth of the Mississippi River
          b.   ?? New Orleans is lying at the mouth of the Mississippi River

(ii)   a.   My socks are lying under the bed
          b.   ?? My socks lie under the bed

[21] The claim that the event variable is bound by a Tense Operator does not come from Kratzer (1989), as Kratzer assumes that Tense is a predicate. In Wijnen's proposal it not really explained why stative predicates appear as tensed forms: if states do not have an event variable, grammatical binding of tense is not possible either. Thanks to Henriette de Swart for pointing this out.

(1998) pursue the idea that the infinitive *itself* has a meaning, instantiated in the infinitival morpheme, and denotes [-realised]. According to Hoekstra & Hyams, [-realised] gives rise to boulemaic and deontic interpretations of RIs.[22] With this claim, an attempt is made to capture differences in meaning between Dutch and German RIs, on the one hand, and English RIs, on the other.[23] Recall that Wijnen found over 80 % modal RIs in the data of four Dutch boys. Deen (1997) and Madsen & Gilkerson (1999) reported the reverse for English: only ca. 20 % of the RIs in English child language are modal. Hoekstra & Hyams argued that absence of a modal meaning follows from absence of infinitival morphology in English. English RIs contain a bare verb form, so there is no suffix to carry the [-realised] meaning. Referring to the work Giorgi & Pianesi (1996), Hoekstra & Hyams assumed that the English bare verb form has an inherent meaning which differs from the meaning of the infinitive, namely [+perfective], an assumption which is not covered by the English RIs as these denote most frequently events and states that take place in the here-and-now and hence, that are ongoing and not completed (which would be predicted by the specification [-perfective].[24] Moreover, the assumption that

---

[22] Boulemaic RIs are used to express a wish, while deontic RIs are used to give a command (see Chapter 3).

[23] This is not only predicted for Dutch child language, but also for German, French and all other languages in which the infinitive is morphologically marked. The generalisation is that RIs in languages with a morphologically marked infinitive are modal while RIs in languages with a base infinitival form are non-modal (or have a free temporal and modal reference).

[24] Hoekstra & Hyams gave the following contrast in support of the claim that the English bare verb form is inherently perfective:

(i) a. * I see him cross the street
     b. I saw him cross the street

In English, the bare verb form cannot be embedded under a present tense matrix verb. The Dutch infinitive does not show this restriction:

(ii) a. Ik zie hem de straat oversteken
         I see him the street cross
      b. Ik zag hem de straat oversteken
         I saw him the street cross

It is unclear, though, why exactly this difference should indicate that the English bare verb form denotes [+perfective] and the Dutch infinitive [-realised]; only the English example displays a meaning restriction. It can be questioned whether this English restriction must

the English bare form denotes a specific meaning, does not really follow from Hoekstra & Hyams' own proposal: if presence of an infinitival morpheme restricts the meaning of infinitives, then it is expected that absence of this morpheme, resulting in a bare stem, leads to absence of these restrictions. This is the position taken by Hyams (2001a, b): English RIs have a free temporal use as they do not have Tense and there is no verbal morphology.[25]

Hoekstra & Hyams and Hyams related differences in use of Dutch and English RIs to differences in predicate restrictions. Deen (1997) reported that 75% of the RIs in the English data he studied are eventive, Madsen & Gilkerson (1999) found that only 60 % of the English RIs they examined are eventive (data reported by Hyams, 2001a, b). This contrasts strongly with the Dutch results, reported by Wijnen (1997). Based on this comparison, Hoekstra & Hyams and Hyams argued that the Eventivity Constraint - *RIs are restricted to event-denoting predicates* - that applies to Dutch (German and French) does not apply to English. The finding that *both*, meaning and predicate type, differ, leads to an analysis in the spirit of Ferdinand (1996): the modal meaning of RIs restricts the predicate type. For English, the reverse argumentation holds: because English RIs are not modal, they are also not restricted to eventive predicates.

1.6.2   The Telicity Hypothesis

GAVRUSEVA (2001, 2002)   Gavruseva (2001, 2002) argued that the eventivity of RIs follows from *aspectual* underspecification, in particular, underspecification of the functional head Telicity. Verbal predicates differ

---

really be attributed to properties of the bare verb form and not to properties of the present tense in English (which is well-known for its severely restricted use).

[25] Hyams (2001a, b) compared the use of nonfinite verbs to the use of other verb forms. Hyams assumes that children initially make a mood distinction: they use different forms in realis and irrealis contexts. English RIs appear in the realis contexts - because of a lack of infinitival morphology they are appropriate in these contexts - while at the same time English children use modal catenatives like *gonna*, *wanna*, *needta* and *hafta* in the irrealis contexts. Dutch and German children use RIs in irrealis contexts and finite verb forms for realis meanings. The finite forms of English children are used for habitual meanings (Madsen & Gilkerson, 1999).

with respect to their telicity value. Some predicates are telic and have an inherent endpoint e.g. *close, die, arrive*. Other predicates are nontelic and do not have an inherent endpoint e.g. *dance, walk, eat*. There are also verbs that are transient: they can be telic as well as nontelic. For instance, the event denoted by *eat an apple* has finished when the apple is eaten. However, *eat* by itself is not telic. As the addition of an object influences the telicity value, telicity is argued to be compositionality derived (Verkuyl, 1972). Gavruseva argued that children lack the ability to compute the telicity value. Thus, if a transient predicate is used the telicity value of the VP remains undetermined in child language. As the VP must be specified for telicity to establish a tense chain, the chain is blocked.[26] The effect is that an untensed form surfaces. Stative predicates are inherently specified for telicity: they are nontelic. Punctual events are telic. In both cases, a tense chain can be established and a finite sentence is used. Gavruseva added a cross-linguistic dimension and compared Germanic and Slavic languages. In Germanic languages, telicity is syntactically derived, while in Slavic languages it is a morphological feature. According to Gavruseva, the aspectual encoding is present in the grammar as a parameter and languages differ with regard to the value of the parameter. She claims that children leaning Germanic languages need more knowledge before they can determine the value of the parameter than children learning Slavic languages.[27] This delay leads to a relatively long RI-period in the Germanic languages, as compared to the RI-period in Slavic languages.

Although the cross-linguistic claim made by the Telicity Hypothesis is appealing, Gavruseva's proposal faces a problem if we consider an observation reported by Bar-Shalom & Snyder (1998). Bar-Shalom & Snyder compared spontaneous speech data from Italian, Polish and Russian children to test whether or not these children use RIs and go through an OI Stage. The Telicity Hypothesis predicts no difference between Polish and Russian: both are Slavic languages with a morphological aspectual system, and hence, it is expected that they are 'no-RI-languages'. Bar-Shalom & Snyder observed a difference between Polish and Russian child language,

---

[26] The syntactic licensing condition on tense chains is not part of Gueron & Hoekstra's original proposal and is added by Gavruseva (2001).

[27] Gavruseva (2001) argues that knowledge of the DP is required as the presence or absence of a determiner can also affect the telicity value.

however. The Polish children used virtually no RIs. The Russian children produced far more RIs than any of the Polish children did. Under the Telicity Hypothesis, this difference is not expected and it is unclear how the Telicity Hypothesis can account for it.

1.7 The role of the input

The RCH and FCH concentrate on children's linguistic capacities, but do not consider the input seriously. Nevertheless, Poeppel & Wexler (1993: 30) say, "the input data are of course relevant". There are various studies about RIs in which children's use of RIs and the properties of RIs have been ascribed to properties of the input. In this section, the role of the input with respect to the meaning assigned to RIs, the no-overlap pattern in the OI stage and the early acquisition of the infinitive by Dutch children.

Ingram & Thompson (1996) claimed that the predominant modal inter-pretation of RIs of German-speaking children is an effect of what the children hear in the input. They argue that German children associate infinitives with modality based on the input. Blom, Wijnen & Gillis (1998) tested whether or not this claim holds for Dutch child language and compared the modal use of the RIs of two Dutch-speaking boys with the frequency of modally used infinitives in the input. Similar correlations between modality and sentence-final infinitives were observed in the children's input and in the children's output. However, though the proportions of modal infinitives were alike, the expressions that children and adults used differed: children used RIs, whereas adults used periphrastic verbs, consisting of a modal auxiliary and a sentence-final infinitive. The similar proportions, on the one hand, and the observed difference in surface form, on the other, gave rise to a proposal along the lines of Boser et al. (1992) and, more specifically, Ferdinand's (1996) proposal that child RIs contain an underlying null modal. However, as has been pointed out, the AUX-drop hypothesis is problematic from a methodological as well as from an empirical point of view.

Pine, Lieven & Rowland (1998) studied data from English children and their mothers in order to test the claim made by Wexler (1992, 1994) and to find

out if the English children really go through an OI stage. They replicated De Haan's (1987) no-overlap observation: in early child English, which they investigated, finite sentences and RIs produced during the OI stage contain systematically different verbs. The early finite verbs are mainly auxiliary-like items, whereas the infinitives are thematic verbs. This observation suggests that finiteness is not truly optional, as finite sentence and sentences that lack finiteness express different meanings. On the contrary, these two forms correspond to different LF representations and denote different meanings. To explain this dichotomy between the verbs children use in finite sentences and RIs, Pine et al. looked at the input that the children received. They concluded that the no-overlap in child English reflected no-overlap in English child-directed speech:

> "…These results suggest that the major determinant of individual mother's use of tensed [finite verbs, EB] and untensed forms [infinitives, EB] is the lexical frequency of those forms in the language, and that it is possible to explain the patterning of children's use of tensed and untensed main verbs fairly straightforwardly in terms of the frequency with which children are exposed to particular lexical forms in the input." (p. 20)

Besides meaning and verb type, the early use of infinitives itself has been related to properties of the input. Wijnen, Kempen & Gillis (2001) tried to answer the question whether or not Dutch children's early use of infinitives can be explained as an input-effect. Wijnen et al.'s study is, in part, a replication of a study carried out by Klein (1974). According to Klein, the early use of sentence-final infinitives by Dutch children could be analysed as an effect of frequencies in the input. Unlike Klein (and unlike Pine et al. for English), Wijnen et al. did *not* find a significant relation between frequency of sentence-final infinitives in the input and children's early and frequent use of RIs. Because of this observation, the conclusion was that not frequency alone, but also other characteristics of infinitival forms in the input of Dutch-speaking children contribute to the explanation of the relative ease with which children learn infinitives. Wijnen et al. mention position, information load and semantic transparency. The infinitive's peripheral position in Dutch makes it a salient form that is presumably easier to detect

for the child than forms that appear in the middle of a sentence, such as finite verb forms in Dutch.[28] The sentence-final infinitives in the input data Wijnen et al. examined, displayed a type-token ratio that was much higher than that of the finite forms. This high type-token ratio indicates that infinitives often carry new and unpredictable information, and are, therefore, important for the child to pay attention to.[29] A final determinant may be that the lexical items that appear as infinitives are often semantically transparent, in the sense that they denote perceivable actions. In this respect, they differ from the finite verbs that denote states, the denotation of which is relatively abstract.[30]

In the input-oriented studies about RIs various similarities have been observed between what the children hear and what they produce. Children use modal RIs and hear sentence-final infinitives in the input most frequently in sentences with a modal meaning (Ingram & Thompson, 1996; Blom et al., 1998). Children show no overlap during the OI stage and they hear do not overlap in the input either (Pine et al. 1998). The interpretation of the observation about the input is very much dependent on the point of view one takes as different claims are corroborated by the input. If one takes a behavioristic point of view, the conclusion may be that these observations indicate that children imitate the input, and that they learn language by copying what they hear. The interpretation from a mentalist point of view, though, would be that children have adult language knowledge. Wijnen et al. (2001) showed that frequency alone does not explain why Dutch children use sentence-final infinitives so early. They argued that there is a complex of other (innate) factors that guide Dutch children to pick up and use the infinitival form earlier than other verb forms.

---

[28] This calls to mind Slobin's (1973) 'pay attention to the end of word' principle for child language.

[29] In this respect, these infinitives are almost embedded in repetitive carrier phrases. Peters (1982) argued that such phrases, as for instance "Look a *ball*", "Say *ball*" or "Gimme the *ball*" help the child to discover prosodic boundaries.

[30] Like in Pinker's (1984) semantic bootstrapping hypothesis, the underlying assumption seems to be that perceptible or observable cues are easier to pick up for young children than more abstract cues.

1.8   Summary

This summary focuses on three properties of RIs that have been explained in the literature in different ways. RIs show a preference for dropped subjects, whereas the finite sentences used alongside them do not (cf. Krämer, 1993; Haegeman, 1995; Hamann & Plunkett, 1998). This general observation could be explained by various variants of the FCH: Truncation accounts (Rizzi, 1992, 1994; Weissenborn, 1994; Haegeman, 1995) as well as Underspecification accounts (Sano & Hyams, 1994; Wexler, 1992, 1994; Schütze & Wexler, 1996; Schütze, 1997). The observation that children use systematically different verbs in their RIs and in their finite sentences, has been accounted for by the incremental version of the RCH (De Haan, 1987), by versions of the FCH (Ferdinand, 1996; Wijnen, 1997; Gavruseva, 2001, 2002) and by input-driven approaches (Pine, Lieven & Rowland, 1998). The temporal, modal and aspectual meaning of RIs is explained as an effect of structural properties of RIs (Ferdinand, 1996; Hoekstra & Hyams, 1998; Hyams, 2001a, b) and this, too, has been considered from the perspective of the input (Ingram & Thompson, 1996; Blom, Wijnen & Gillis, 1998).

1.9   Outline of the thesis

Chapter 1 shows in how many ways the absence of finiteness in early child language can be approached. As the number of theories outranks the amount of robust empirical findings, I have chosen to emphasise the empirical side of RIs in this thesis.

After a chapter devoted to the statistical methods applied in this thesis (Chapter 2), I will start the actual investigation in Chapter 3 by testing the hypotheses regarding the temporal, modal and aspectual denotation of RIs. The claims that will be tested are: (i) RIs can receive any temporal interpretation, (ii) RIs are modal, and (iii) RIs are not completed. Naturalistic data from six Dutch-speaking children and results from Dutch and English child language elicited in a controlled experimental setting are analysed. This cross-linguistic comparison is motivated by Hoekstra & Hyams' (1998) theory that infinitival morphology has a specific modal denotation (see:

§ 1.6.1). By implication, RIs that contain bare stems (e.g. RIs in early child English) have a meaning different from that of RIs containing true infinitival forms (e.g. RIs in early child Dutch). Testing this claim against the experimental results will help us to determine whether it is feasible and hence, whether the conclusion holds that children using RIs already encode modal distinctions.

The findings and conclusions in Chapter 3 provide a basis for the Chapters 4 and 5. In Chapter 4, it will be shown that the temporal/modal denotation of RIs, more specifically, the different modal usages, restrict verb selection in RIs. With reference to earlier proposals, I will argue that the modal properties of RIs lead to the marginal appearance of stative RIs (Ferdinand, 1996; Hoekstra & Hyams, 1998). To support this claim, I freely borrow from the observations reported in Chapter 3. My claim is that various factors lead, independent of each other, to the absence of stative RIs. First of all, like adults, children between two and three years old know that certain modalities are incompatible with states. Secondly, Dutch children hear only few stative infinitives in the input (which is probably an effect of the selection restrictions on stative predicates). As long as children are unable to derive infinitival forms themselves, the absence of stative infinitives is directly linked to distributions in the input. Thirdly, the kind of modality that is compatible with states is not accessible to children in the RI-age due to cognitive immaturity..

In Chapter 5, I will expand on a number of observations from the preceding chapters. Earlier observations will be placed in a developmental perspective, a strategy that turns out to be very fruitful. It will be shown that the predominance of modal RIs is not an effect of children's early ability to encode modal distinctions. On the contrary, it is the effect of children's ability to encode *present tense*. There may be support, however, for the claim that RIs mark aspectual contrasts. Changes in lexical and morphological properties of the verb forms that Dutch children know, suggest that inflection is acquired fairly late, as was hypothesised in Chapter 4. In order to learn inflection, Dutch children have to look beyond misleading patterns in the input. The only way to achieve this is by collecting sufficient evidence. Since verbal inflection is identified as the trigger for the acquisition of verb movement, it does not come as a surprise that I will defend the claim that verb movement, which marks the endpoint of the developmental path from

root infinitives to finite sentences, is not available to Dutch children from the beginning on.

The thesis closes with a summary of the most important observations and conclusions, a brief evaluation of previous hypotheses and a section on the implications for future research.

Statistical Methods

Studies on early child language face the problem that analysable data are sparse. The division of data into developmental stages makes the problem even more serious, because few data per stage remain after division, particularly in the earliest stages. This brings us to a second problem, which is that the numbers of utterances differ considerably from stage to stage. The implication of the lack of data and the uneven distribution of the data over stages is that standard statistical tests cannot be used. Standard tests are suitable for data that are normally distributed. The data analysed in this thesis contain too little information to decide whether or not they are normally distributed, however. Therefore, a statistical method is applied that does not require any additional assumptions: the permutation test (Good, 1999). This chapter describes this method and how it is used on various occasions in this thesis.

2.1 Differences between stages

Chapter 5 deals with longitudinal analyses, but the present chapter provides information on the stages that are distinguished. It is explained how statistical methods will be applied in this thesis in order to test for changes over time. Additionally, it is explained which method is used to determine whether or not generalisations hold for the total sample of six children.

2.1.1 Four stages

I analysed transcriptions of spontaneous speech data of six monolingual Dutch-speaking children. All analysed data are available through the Child Language Data Exchange System or CHILDES (MacWhinney, 1995). Abel, Daan, Josse, Laura, Matthijs and Peter are part of the Groningen Corpus. Laura is part of the Van Kampen Corpus. The transcriptions in these corpora are based on audiotape recordings made at home, in an unstructured home setting. The children's age ranges are given in Table 2.1. This table shows furthermore the total number of utterances produced by each child in the selected files to give an idea about the size of the files that were examined.

Table 2.1: Children's age ranges and the total number of utterances in the selected files that are used for analysis, data from all six children

Child	Age range	Total number of utterances
Abel	1;10.03 - 2;07.29	2890
Daan	1;08.21 - 2;09.10	4859
Josse	2;00.07 - 2;08.18	3340
Laura	1;09.04 - 3;04.06	4241
Matthijs	1;09.30 - 2;11.19	4624
Peter	1;07.18 - 2;03.21	2349

Laura's sample contains files from an older age (above three) than the samples from the other children (under three). As Laura suffered from recurrent ear-infections, her development is delayed. Otherwise, Laura developed normally.[1]

From the corpus of each child, a number of files were selected. The data that were examined represent four developmental stages. I selected clusters

---

[1] See Van Kampen (1997) for detailed information about Laura's linguistic development.

of files that correspond roughly with the following frequencies of finite verbs. The right-most column gives the average MLU (Mean Length of Utterance) of the six children in a certain stage:

Table 2.2: 'Stages' in the development of finiteness: frequency of finite sentences per stage (as a percentage of the total number of sentences that contains a verb), average MLU per stage.

Stage	Frequency of finite verbs	Average MLU
I	0	1.125
II	< 30 %	1.387
III	50 %	1.972
IV	> 80 %	2.607

The selected files are listed in Appendix 2.1 (p. 237).[2] Appendix 2.2 (p. 238) contains detailed information about the numbers of RIs and finite sentences in the four stages and the average MLU per child per stage. Note that children do not jump from 0 % to 30 % to 50 % to 80 % finite sentences. Rather, the increase of finite sentences is gradual, and the stages are snapshots taken from this gradual development.

2.1.2   The Root Infinitive period

Stages I-IV represent the Root Infinitive period (RI-period), i.e. the phase in which children (i) use RIs more frequently than adult speakers do (quantitative criterion), and (ii) use RIs differently from adult speakers (qualitative criterion). Stages II-IV constitute the Optional Infinitive stage (OI-stage) as defined by Wexler (1992, 1994). According to Wexler's

---

[2] The age of the child is given per file as well as the Mean Length of Utterance (MLU). The MLU is the number of morphemes that the child produced in a file divided by the number of utterances the child produced. This gives an estimation of the length of the utterances that a child produces. When the children get older, the MLU goes up.

definition, finite sentences and RIs co-occur in the OI-stage. The OI-stage ends when the frequency with which children use RIs equals the frequency with which adult speakers use RIs. Bol & Wijnen (1993) determined that approximately 5 % of the utterances in a spontaneous speech sample of adult Dutch are RIs. It is expected that when the quantitative criterion for the end of the OI-stage is met, the qualitative criterion is met as well. Since finiteness is only truly optional if children use the same (types of) verbs in RIs and finite sentences, Wijnen (2000) states that children enter the OI-stage when they are able to use the same verbs in finite sentences and RIs. In Chapter 5, it will be shown that this is not the case in stage II, even though the children use finite sentences next to their RIs. This observation shows that the choice between the application of Wexler's or Wijnen's definition is not arbitrary: it has consequences for the interpretation of child data, as Wijnen's definition leads to the discovery of a substage that remains obscured when Wexler's definition of the OI-stage is applied. My concern is changes over time. Therefore, Wijnen's definition is more appropriate than the definition given by Wexler. The label 'OI-stage' as defined by Wijnen, applies to stages III/IV in my data.

Developmental stages stress the dynamicity of language learning. They are the main topic of Chapter 5. In this chapter, data from more children and data that from a wider range of phenomena (than discussed in other longitudinal studies on early child Dutch) indicate that structural knowledge of finiteness (e.g. inflection and verb movement) is acquired by Dutch children in a piecemeal fashion. The four stages become less arbitrary than they may look now. It will be shown that at each stage represents a step in the direction of the target grammar of Dutch, and that each step gives rise to various simultaneous developments. In the third and fourth chapter, I will abstract away from the stages I-IV. In these chapters, properties of RIs will be compared to properties of finite sentences that are used alongside RIs. Changes over time *within* the set of RIs (and finite sentences) remain for the fifth chapter.

2.1.3   Two developments

Over time, the meaning of RIs changes. This change can be modelled by an S-shaped curve: when the children get older, RIs are more frequently modal

as compared to the preceding stages, that is, RIs are more frequently used to describe desires, intentions and to give commands and relatively less often to describe events that are ongoing at speech time. More specifically, between stages II and III, a *modal shift* takes place (Blom & Wijnen, 2000). The permutation test will be applied to determine whether or not the modal use of RIs in stage II differs from stage III/IV in the data of Abel, Daan, Josse, Laura, Matthijs and Peter. A second development concerns the dropping of subjects in RIs. Between stages I and III, the relative number of dropped subjects in RIs shows a decrease while between stages III and IV, the relative number of dropped subjects in RIs starts to go up again. This development can be modelled by a U-shaped curve. The permutation test will be applied here to test whether or not stages II/IV differ significantly from stage III. Stage I is left out of the statistical analysis, because the number of observations in this stage is so marginal that they may have a disproportionate effect on the outcome.

To test if the developmental differences as they visually present themselves are statistically significant, it will be estimated to what extent the observed differences between the stages are due to chance. In order to do this, I will start by spelling out the null hypothesis: "there is no difference between the stages". More specifically, in the case of the modal shift, the null hypothesis is that there is no difference between stage II and stages III and IV; the results from stage II, on the one hand, and stages III and IV, on the other hand, are drawn from one single distribution. In the case of subject drop, the null hypothesis states that there is no difference between stage III, on the one hand, and stages II and IV, on the other. Then, the null hypothesis will be simulated by randomly shuffling the observations per child over the different stages; recall that there is no difference between the stages, hence this random distribution is the distribution given the null hypothesis. We can now calculate the expected difference if all stages were similar. For the modal shift, the relevant difference is between the proportion of modal RIs in stage II and the proportion of modal RIs in stages III/IV. For subject drop, the relevant difference is that between the proportion of subject drop in stage II/IV and the proportion of subject drop in stage III. As I am interested in differences between *stages* and not in differences between children, I will calculate the average difference of the six children in both cases.

The above procedure gives us the average simulated difference between the stages given the null hypothesis. In addition, the average observed difference is calculated. This means that for each child the difference between the proportion of modal use in stage II and stages III/IV is calculated. The same procedure is applied to the proportions of subject drop in stages II/IV and stage III. As before, the average difference over the sample of six children will be estimated. We now have two values: a simulated average difference that is expected given the null hypothesis and an observed average difference. Since the number of all possible permutations and combinations for the six children is very big, we approximate the distribution of the differences based on chance by randomly shuffling the observed proportions 1000 times (or 5000, or more), which provides 1000 values (or 5000, or more) for the test statistic based on chance (Monte Carlo simulation). By counting the number of chance differences that are equal to or bigger than the observed difference, a very close approximation of the probability is obtained that the observed values are based on chance. This probability is the p-value. The p-value can be interpreted as follows. If the value is very small (say lower than a criterion value of 0.05), we consider it very unlikely that the stages are in fact one undifferentiated stage as far as meaning or subject drop is concerned. The null hypothesis is not confirmed; the alternative hypothesis provides a model that fits the observations better. The conclusion is that the changes over time are statistically significant patterns that have to be explained. If the p-value is very big (e.g. p > 0.95), chance will lead to a difference that is bigger than the observed difference in most of the cases. Thus, the difference is much smaller than would be expected on the basis of chance alone. This case is the opposite of the first (i.e. the small p-values). At the end of section 2.2, I give an example of the strategy that I follow in the case of unexpected high p-values.

2.1.4   "Jack knife" technique

To test whether it is possible to generalise over the six children, I applied a so-called Jack knife technique. This is a method, which is used to determine if the obtained p-value for the total sample of six children is influenced by the deviating results of one particular child. A simple example will illustrate the Jack knife technique. Imagine a study in which cultural differences

between children's preferences for certain shapes are examined. Assume that Japanese children prefer round forms; we want to know whether or not this preference for round forms within Japanese children develops over time. Four children are tested at the age of 1, 2 and 3 years old to find out if a change occurs and if the preference for round shapes is stronger at an older age than at a younger age.

Table 2.3: Imaginary results of longitudinal study of preferences for round shapes of four Japanese children at the ages of 1, 2 and 3

Age	Child A	Child B	Child C	Child D	Average
1	0.3	0.6	0.1	0.4	0.35
2	0.2	0.1	0.3	0	0.15
3	0.7	0.9	0.8	0.7	0.78

The results in Table 2.3 suggest a change in preference between the age of two and three. In order to test if this difference is statistically significant, the results from the ages 1/2 are compared to the results from age 3. The null hypothesis states that there is no significant difference between the average preference for round shapes of the four children at the ages 1/2 and the average preference of these four children at the age of 3. With the help of a simulation of the null hypothesis, a p-value of 0.047 is obtained that the observed difference between average preferences for round shapes at the ages 1/2 is drawn from the same distribution as the average preference at the age of 3. Though the difference is not highly significant, it is just below the significance level of $p = 0.05$. Hence, it can be concluded that there is a change over time.

It may be that this value is due to one individual participant (generalisability), however. This may happen especially since so few children are followed and the results are sparse. To test this, the probability is recalculated four times, each time the results of one particular child are omitted: first Child A is left out and only Child B, C and D are included, then Child B is left out and only child A, C and D are included, etcetera. The results are given in Table 2.4:

Table 2.4: Probabilities that the observed average difference between preferences for round shapes at the ages of 1/2 and at the age of 3 is due to chance. The probabilities are calculated over the entire group of four children and over 4 groups of three children, with one specific child omitted at a time

	p-value
All children	**0.047**
Omitted	
Child A	0.041
Child B	0.045
Child C	0.061
Child D	0.037

In most cases, p is still fairly low, i.e. below the criterion value (of 0.05). In one case, i.e. when child C is omitted, the p value is slightly above 0.05. As this is only slightly above the critical value, it is not the case that the data from one particular child influence the obtained p value for the total sample of four children disproportionally. It seems that we can generalise over the children and conclude that the preference for round shapes of Japanese children develops over time. This development occurs when the children are between two and three years old.

2.2   Differences between groups

In Chapter 3, experimental data on the meaning of Dutch and English child RIs will be presented. Apart from testing whether or not there is a preference for a certain meaning within the set of RIs in the two languages, I will test for differences between the two languages, i.e. a difference between groups. The null hypothesis states that the Dutch and the English children are in fact one undivided group as far as their preference is

concerned. Like before, the permutation test is appropriate, given that I collected relatively few data.

How does the permutation test work when differences between groups are tested? Imagine that we want to know whether Japanese children like round shapes more than English children do. In order to test whether or not there is a difference between children from the two groups, a simple experimental setting is designed and 12 Japanese and 12 English children are tested. The experiment contained 10 test items: each item consisted of two forms, a round form and a square form, and each child had to choose between the two forms. In Table 2.5 the results are given as the proportion of round preferences (number of test items divided by number of round forms that were chosen):

Table 2.5: Results of an imaginary experiment on the preference for round or square shapes with Japanese and English children, average and standard deviation (SD)

Subject	Proportion preferences for round shapes in Japanese	Proportion preferences for round shapes in English
1	0.6	0.2
2	1	0.1
3	0.5	0.6
4	1	0.4
5	0.7	0.3
6	0.9	0.4
7	0.6	0.5
8	0.5	0.5
9	0.8	1
10	0.3	0.3

11	0.6	0.1
12	0.5	0
Average	*0.67*	*0.37*
SD	*0.22*	*0.27*

Recall that the null hypothesis states that there is no difference between Japanese and English children. Another way of putting this is to say that the relation between a particular preference and being Japanese or English is completely accidental. That is, any combination between a particular preference and a nationality (Japanese or English) is just as good as any other. Any such combination will result in a particular average preference for the Japanese and an average preference for the English children. These averages will of course differ, but each of these differences will be accidental. For differences to be meaningful and to reflect a real distinction between Japanese and English children, it should be bigger than the differences that result from the arbitrary combinations of preferences with nationality labels. In practice, the difference is said to be meaningful if it is bigger than the great majority of such accidental differences (usually set to at least 95 %). I will go briefly over the statistical procedure once more. First, the preferences are combined into one group. Secondly, we randomly shuffle the position of the preferences, like randomly shuffling a deck of cards. Third, we assign the first half of the randomly shuffled preferences to the Japanese, the second half to the English group. Fourth, the preferences of the Japanese and the English group are calculated and the difference between these preferences is determined (Japanese difference minus English preference). Fifth, the procedure is repeated many times (1000 or 5000). Sixth, we count the number of times that this difference is bigger than or as big as the difference that we observed in our experiment and divide this by the number of repetitions (1000 or 5000). This results in the p-value, the estimation of the probability that the observed difference is due to chance. This procedure results in $p = 0.004$, which is clearly below 0.05. We can conclude that the English and Japanese children come from different groups as far as the preference for round shapes is concerned. The Japanese children show a greater preference for round shapes than English children do.

Besides the difference between Japanese and English children, the data in Table 2.5 can also be used to test whether the Japanese children (or English children) show a preference for round shapes (or for square forms). The null hypothesis can be simulated by randomly distributing the proportions of round and square shapes in Table 2.5 for each child: we observed that child 1 had a proportion of 0.6 preference for round shapes. However, given the null hypothesis, the chance that the preference for square shapes is also 0.6, is large; there is no preference for either the one or the other.

Table 2.6: Results of imaginary experiment, proportions of preferences for round and square shapes of Japanese children

Subject	Proportions of preferences for round shapes	Proportions of preferences for square shapes
1	0.6	0.4
2	1	0
3	0.5	0.5
4	1	0
5	0.7	0.3
6	0.9	0.1
7	0.6	0.4
8	0.5	0.5
9	0.8	0.2
10	0.3	0.7
11	0.6	0.4
12	0.5	0.5
Sum	*8*	*4*

In order to simulate the random distribution that is expected under the null hypothesis, the results in Table 2.6 are shuffled for each child. The sum of the proportion of preferences for round shapes of all children is calculated;

the same procedure is applied to preferences for square shapes and then the difference between these two sums is calculated. This value is calculated 1000 times. When the observed difference is compared to the chance distribution, the probability that our observed difference is found by chance is 0.017. This is below the 0.05, hence the difference is statistically significant: the Japanese children choose more often round forms than square forms. If we follow the same procedure for English, an extreme p-value of 0.96 is obtained. Based on this, we can conclude that the English children do not show a preference for either round or square forms. However, the difference is much smaller than should be expected on the basis of chance alone. How to interpret this unexpected high p-value? I have now tested for a preference of round as compared to square. However, if I had tested for the reverse, that is, for a preference for square shapes as compared to round shapes, the outcome for English would have been 1 - 0.96 = 0.04 (and for Japanese 1 - 0.017 = 0.983). The conclusion of this test would be that the English children show a preference for square forms, whereas the Japanese children show no preference. The general rule that I will follow with regard to the extreme values of p is that if $p < 0.05$ the null hypothesis is confirmed. If $p > 0.95$ there is support for the hypothesis that states the opposite.

2.3 Summary

I dedicated a chapter to the statistical method used in this thesis, i.e. the permutation test, because this technique is not commonly used in studies on language acquisition. A problem that occurs in all studies that focus on children's earliest utterances and that make use of empirical data, is that there are few useful data available and it is unclear whether or not these data are normally distributed. The advantage of the permutation test is that it works without the assumption that the data are normally distributed. I illustrated how the permutation test can be applied to test for differences between groups and differences between stages. These two kinds of differences play a role in Chapters 3 and 5.

CHAPTER 3

Temporal, Modal and Aspectual Denotation

In 1917, Van Ginneken observed that Dutch children use RIs to express wishes and desires. Almost a century later, Krämer (1993) and Wijnen (1997) confirmed the observation that RIs in Dutch child language predominantly express unrealised events. Around the same time, Ingram & Thompson (1996) and Ferdinand (1996) reported similar observations for German and French speaking children. The impossibility to generalise the specific meaning of RIs to English child language lead Hoekstra & Hyams (1998) and Hyams (2001a, b) to the proposal that infinitival morphology, which is present in Dutch, German and French but not in English, carries the modal meaning that is assigned to RIs.[1] The present chapter addresses the issue of the temporal, modal and/or aspectual properties of RIs. It provides more insight into the way in which Dutch-speaking children use RIs. Throughout this chapter, I will analyse and re-analyse both naturalistic and experimental data in order to clarify the nature of differences between the meaning that is assigned to two different types of RIs, those represented by RIs in Dutch and English child language, respectively. The topic of this chapter gives rise to various methodological questions. Therefore, I will pay a considerable amount of attention to research strategies.

---

[1] Henceforth, I refer only to Hoekstra & Hyams (1998) with regard to this claim. In Hyams (2001a, b), the claim is worked out in more detail, but the basic claim regarding differences between Dutch and English remained similar.

3.1 Outline of this chapter

Let me start by warning the reader that this chapter is fairly long. For keeping the overview, it may be helpful to make a three-way division. In the first part, that is, section 3.2, definitions are given and the relevant hypotheses are repeated. The second part discusses corpus results from early child Dutch are discussed (§ 3.3). In the third part, differences between the meaning of RIs in early child Dutch and early child English are examined in-depth (sections 3.4 to 3.7). This includes the presentation of experimental results as well as re-analyses of the Dutch corpus results. Section 3.8 contains a summary of the chapter.

3.2 Three hypotheses

The claims that have been made about the meaning of RIs in child Dutch can be divided into hypotheses about the temporal, modal or aspectual properties of RIs, respectively: the *No Tense Hypothesis* (Behrens, 1993; Wijnen, 1997), the *Modal Hypothesis* (Ferdinand, 1996; Ingram & Thompson, 1996; Hoekstra & Hyams, 1998) and the hypothesis that states that RIs obey a *Non-Completedness Constraint* (Lasser, 1997).

The notions tense, modality and aspect are well known for their complexity. The terminology becomes particularly confusing in studies about young children's speech since all terms originated as labels for various fine-grained semantic distinctions that adult speakers make and that children between the two and three - commonly the age in which they use RIs - cannot make (at least, not in the language they use).

3.2.1 Tense

Tense is the grammatical expression of a relation between speech time and event time, with present tense (e.g. speech time) as the unmarked tense. The absolute tenses - i.e. past, present and future - denote a relation between speech time and event time (Reichenbach, 1947). They do not make a claim about the length or span of the event. Comrie (1985: 38) wrote:

"Present tense refers only to a situation holding at the present moment, even where that situation is part of a larger situation that occupies more than just the present moment."

while past tense

"only locates the situation in the past, without saying anything about whether that situation continues to the present or in the future, although there is often a conversational implicature that it does not continue to or beyond the present." (p. 41)

Following the common strategy in research on RIs, I do not make a distinction between tense and finiteness.[2] Note that this idea about the relation between finiteness and tense may be a simplification, however. A number of studies that concentrate on finiteness in adult language suggest a more complex relation. Non-finite constructions that have their own temporal domain are possible, as pointed out by Erb (2001). This suggests that finiteness and tense are dissociated. Other linguists identify finiteness as a prerequisite for tense (and also mood) marking: the functional finiteness head, which is claimed to be part of the C-domain, provides a sentence's anchoring point to speech time (Holmberg & Platzack, 1995; Rizzi, 1997; Rousseau, 1998). Klein (1994) defines finiteness as a complex notion that contains tense and assertion information.[3]

---

[2] The observation that finite verbs contain tense and agreement features, could lead to the hypothesis that finiteness also includes agreement. There are various reasons to exclude agreement from finiteness, however. Finiteness is a semantic concept, while agreement describes a relation between two words (subject and verb) and is syntactic in nature rather than semantic. The exclusion of agreement from the concept of finiteness is corroborated by the occurrence of inflected infinitives that are marked for agreement in Hungarian (Tóth, 2001) and Portuguese (Raposo, 1994).

[3] More specifically, finiteness is associated with topic time, i.e. 'the time span to which the speaker's claim on this occasion is confined' (Klein, 1994: 4) and with assertion. Both topic time and assertion can be emphasised by contrastive focus (marked with capitals in the examples below); see (i) and (ii), respectively. In both cases contrastive focus lies on the finite verb. In example (i) the topic time is emphasised, while in (ii) the positive claim that Peter was at the party is emphasised:

Amongst others, Wexler (1992, 1994), Rizzi (1992, 1994) and Hyams (1996) argued that the absence of a finite verb in RIs is one of the indications that RIs lack tense.[4] Wijnen (1997) tested this hypothesis and argued that *if* tense is absent, RIs must have a free temporal reference. He interpreted the existence of RIs with past, present and future tense interpretations in the data of four Dutch-speaking children as support for the claim that RIs are untensed, irrespective of the uneven proportions of the temporal interpretations and the strong preference for a future or intentional usage. Wijnen's Dutch findings resembled the findings for German child RIs that were reported by Behrens (1993) and Lasser (1997).

(1)   *No Tense Hypothesis*
        RIs lack tense. Hence, they can be used to denote past, present and
        future tense (Behrens, 1993; Wijnen, 1997; Lasser, 1997).

3.2.2   Modality

Modality has to do with possibility and necessity (Kratzer, 1981: 39). The prototypical examples of modal expressions are the modal verbs *can* and *must*, the former denoting possibility and the latter necessity. There are many different modalities (see Palmer, 1986, for an overview). One of the most well known distinctions is that between *epistemic* and *deontic* (also: *root* or *circumstantial*) modality. In (2) and (3), this distinction is illustrated with the verbs *can* and *must*; in (2), epistemic examples are given while the examples in (3) exemplify deontic modality:

(i)   A: The book is on the table
           B: No, the book WAS on the table

(ii)   A: Peter was not at the party yesterday
           B: Sure, he WAS at the party (at some point). He came late and left early.

---

[4] Behrens' (1993) observation that German children that omit finite verbs are able to use temporal adverbials corroborates the conclusion that RIs need a grammatical explanation. Absence of tense is not the effect of children's inability to understand tense. A similar observation holds for Dutch children in the RI-age. The children I have examined use adverbials and particles to introduce topic time or refer to a future action like *nu* 'now', *even* 'just'. *straks* 'later on'.

(2)   a   John must be the murderer
            'Based on everything I know, I conclude that it is necessary that
            John is the murderer'
       b.   John can be the murderer
            'Based on everything I know, I conclude that it is possible that
            John is the murderer'

(3)   a.   Mary must leave the room
            'It is required that Mary leaves the room'
       b.   Mary can leave the room
            'It is permitted that Mary leaves the room'

Obligation, requirement and permission are deontic modal notions. According to Lyons (1977), these modal notions always involve a morally responsible agent that acts necessarily or possibly. Epistemic modality is speaker-oriented and refers to a speaker's beliefs or judgements. Often, reasoning is involved. Palmer (1986) makes a distinction between deontic and *dynamic* modality (both can be considered as root or circumstantial modalities). Dynamic modality is subject-oriented, as opposed to deontic modality. Palmer uses the label 'dynamic modality' for wishes or abilities, which refer to a necessity that is somehow internal to the subject (also: *volition* or *boulemaic* modality) or possibility, respectively. The distinction between dynamic and deontic modality can be illustrated with the sentence in (4):

(4)   I can play the trombone, but I cannot play the trombone now
       'I know how to play the trombone, but I am not allowed to play the
        trombone now'

The first clause of (4) expresses dynamic modality. The subject tells us that he has learned to play the trombone, hence (s)he is able to play the trombone. The second clause expresses deontic modality. The subject announces that he is not allowed to play the trombone at the time he utters (4), because he is not permitted to do so, maybe because he is in a place where silence is required. Note that the paraphrase in (4) is not the only accessible interpretation. Other possible interpretations are: 'I am allowed to play the trombone, but I am not in the physical condition to play the trombone' or 'I know how to play the trombone, but I am not in the

physical condition to play the trombone', etc. Kratzer (1981: 42) emphasised the influence of the context for the assignment a modal interpretation:

> "Modals are context-dependent expressions since their interpretation depends on a conversational background which usually has to be provided by the utterance situation."

Confusingly, the terms 'modality' and 'mood' are sometimes used as synonyms. I will make a clear distinction and take mood as a *grammatical* category (unlike modality, which is a semantic category) that surfaces in verbal inflection and/or verb placement. Though the formal properties of mood seem quite clear, it is less easy to describe what mood actually is, especially since there are different mood classifications.[5] Most well known are the two-way mood distinction between indicative and subjunctive mood (verbal mood) and the three-way mood distinction between declarative, interrogative and imperative sentences (sentence mood). In Indo-European languages, indicative and subjunctive mood are marked with inflection on the verb. The indicative is used to express perceived reality whereas the subjunctive is used to express doubt, probability, certainty, etc. The Spanish examples in (5) illustrate the distinction. In (5a), the speaker knows that his car is not working, whereas in (5b), the speaker does not know whether his car is defect or not.

(5)   a.   Llegaré aunque mi carro no **funciona**
             I will arrive even though my car is not running
       b.   Llegaré aunque mi carro no **funcione**
             I will arrive even if my car is not running

The indicative/subjunctive distinction can be seen as the grammaticalisation of epistemic modality, as a speaker's certainty or belief is marked on the verb through inflection (Lyons, 1977; Palmer, 1986; Bybee & Fleischman, 1995). The three-way distinction between the sentence-types declarative, interrogative and imperative can also be described as a mood distinction. Like verbal mood, sentence mood is closely related to modality, although the relation has an entirely different character. Imperative mood and deontic modality are closely related as imperatives are prototypically interpreted as

---

[5] In a clarifying overview article, Portner (1998) discussed various different classifications.

deontic modal expressions. In (6a), an imperative sentence is given, while (6b) contains a deontic declarative sentence:

(6) a. Ga nu!
         leave now
         'Leave now!'
     b. Je moet nu gaan
         you must now go
         'You have to leave now'

In the debate on the semantic properties of RIs, the claim that RIs are modal is most clearly put forward in the work of Ferdinand (1996), Ingram & Thompson (1996) and Hoekstra & Hyams (1998). In these studies, the term 'modality' does not only collapse intentions and various, more specific, modal interpretations, such as volition and deontic modality, but also includes temporal and mood properties of RIs. I will turn to the relation between tense and modality in the following paragraph.

In future tense, modality/mood and tense meet. It is commonly assumed that future tense is modal, because it is impossible to make a claim about our knowledge of the future (Lyons, 1977; Palmer, 1986). As we can make a claim about our knowledge of the past and the present, there is in this respect a distinction between past and present tense, on the one hand, and future tense, on the other. In various languages, future tense is expressed as a mood distinction with verbal inflection. In other languages, auxiliaries express both modality and future tense (cf. Chung & Timberlake, 1985; Palmer, 1986). For the present discussion, the relevant implication of the overlap between the notions of tense and modality is that RIs that are assigned a future tense interpretation count as modal. In fact, the only formulation of the Modal Hypothesis that captures all more specific formulations that have been given, defines the modality in RIs in terms of tense:

(7) *Modal Hypothesis*
     RIs denote modal events and do not denote past and present events
     (Ferdinand, 1996; Ingram & Thompson, 1996; Hoekstra & Hyams, 1998)

According to Ingram & Thompson (1996: 102), modal RIs express that "some activity will, can or should occur". According to Ferdinand (1996), modal RIs denote deontic modality and inchoative aspect. Hoekstra & Hyams (1998) include deontic modal RIs and volitional RIs in the set of modal RIs. According to them, modal RIs denote events that are not yet realised, because the verb form in RIs, i.e. the infinitive, contains the feature [-realised]. Note that the Modal Hypothesis proposed by Hoekstra & Hyams has a cross-linguistic dimension and, according to this theory, the modal meaning of RIs cannot be generalised to English child language. I will take up this issue in sections 3.4 and 3.5. The (working) definitions of Ingram & Thompson (1996), Ferdinand (1996) and Hoekstra & Hyams (1998) all exclude RIs that denote events that took place prior to speech time (past) or events that are taking place at speech time (present) from the set of modal RIs. Hence, (7) adequately summarises these proposals. Apart from studies that explicitly deal with the modality of RIs in terms of one or the other variant of the Modal Hypothesis above, there are some studies in which the modality of RIs is assumed to be an effect of the No Tense Hypothesis (Schönenberger, Pierce, Wexler & Wijnen, 1995; Schütze, 1997; Wijnen, 1997). In these studies, the idea that the modal interpretation is the reading that follows if tense cannot be grammatically bound, that is, the unanchored reading.

I will conclude this section with a brief note on sentence mood that serves as an additional illustration of how modality is interpreted in this thesis. With regard to their function, imperatives and deontic modal declarative sentences overlap: both (6a) and (6b) are used to order or request something.[6] Grammatically, they differ, however, as different verb forms are used and the main verb is placed in different positions in imperative and deontic modal declarative sentences.[7] Lasser (1997) divided RIs in German child language into declaratives, interrogatives and imperatives. In her study,

---

[6] A similar claim is made by Fries (1999).

[7] Following Levinson (1983), I make a distinction between sentences and utterances. Levinson (p. 243) writes:

> " […] we must be careful to distinguish the set of terms *imperative*, *interrogative* and *declarative* from the set of terms *order* (or *request*), *question* and *assertion* (or *statement*). The first set are linguistic categories that pertain to sentences, the second set are categories that pertain only to the use of sentences (i.e. to utterances and utterance-types)."

these labels do not refer to types of sentences or sentence mood, but to three different illocutionary forces (making a claim, soliciting information from an addressee and soliciting actions from an addressee, respectively). Hence, Lasser uses the notions declarative, interrogative and imperative to describe properties of the content of RIs and not to describe formal grammatical properties. From this functional perspective, imperative RIs are necessarily modal. I will not address the question whether RIs are declarative, interrogative or imperative in this thesis. I take the formal perspective on these sentence types, which leads to the conclusion that RIs can be used, theoretically, in all three moods. Verb form (i.e. infinitive) and verb placement (i.e. sentence-final) in RIs is compatible with all three sentence-types. This is illustrated in (8): declarative, interrogative and imperative sentences can all contain a sentence-final infinitive in adult Dutch:

(8) a.   Ga je je nu **aankleden**?
            go-*fin* you yourself now dress-*inf*
            'Do you get dressed now?'
     b.   Je gaat je nu **aankleden**
            you go-*fin* yourself now dress-*inf*
            'You are going to get dressed now'
     c.   Ga je nu **aankleden**!
            go-*fin* you now dressed-*inf*
            'Get dressed now!'

The examples in (8) show that sentence mood can be morpho-syntactically marked in adult Dutch by placing the finite verb in first or second position.[8] However, the infinitive does not give any information about sentence-mood, since it is placed in final position, irrespective of a particular mood. As RIs only contain a sentence-final infinitival main verb, I conclude that

---

[8] In both imperatives and interrogatives, the finite verb is inverted (in the first sentence-position), whereas in declaratives, it is in second position. The examples in (8) do not show a morphological difference between imperatives and interrogatives. However, the finite verb in a Dutch imperative is always a bare verb form. The finite verb form in interrogatives is normally inflected, with the only exception being second person singular. In declarative sentences, the finite verb inflected for second person singular has a suffix [-*t*], while in interrogative sentences, there is no suffix.

morpho-syntactic properties do not suffice as a criterion for determining the sentence mood of RIs.[9]

3.2.3   Aspect

Comrie (1985) called aspect the *internal temporal structure of an event*. Notions like 'ongoing', 'progressive', 'inchoative', 'prospective', 'perfective', 'telic' or 'punctual' describe aspectual properties. Although there is resemblance between aspect and tense, the two notions are distinct.[10] A syntactic difference is that that tense (like modality and mood) is located in at a high position in the syntactic structure, that is, the C-domain. Tense takes scope over the entire proposition expressed in a sentence, that is, subject as well as predicate. Aspect is related to the VP level, however.

In the discussion about semantic properties of RIs, aspect plays a twofold role. In this chapter, I will investigate whether or not there is evidence for claim that RIs contain structural features, either phonetically realised or not, that determine their aspectual meaning. This investigation is motivated by Lasser's (1997: 64) Non-Completedness Constraint, which states that the denotation of RIs is aspectually restricted:

(9)   *Non-Completedness Constraint*
         The predicate of an RI cannot refer to a completed event

Although Lasser formulates the constraint in aspectual terms, she makes no distinction between past tense RIs and completed aspect RIs in the interpretation of child RIs, presumably because this distinction cannot be made on the basis of the available information. In fact, Lasser (1997: 197-198) takes completion at speech time as a criterion for past interpretation. One could infer that Hoekstra & Hyams' (1998) feature [-realised] (cf. § 3.2.2) makes a prediction similar to the constraint in (9). The feature applies to events that are not finished or that have not yet been completed. On closer inspection, this conclusion is incorrect. If [-realised] had the

---

[9] Prosodic information would be very helpful in this respect: questions, for instance, are prosodically marked by a rising intonation.

[10] See Klein (1994) for the relation between tense and aspect.

same effect as (9), the prediction would be that [-realised] RIs allow for a present tense reading, because an event that takes place at speech time is not completed at speech time. Hence, provided that this feature is understood as [-completed], present tense and [-realised] are compatible. From their analyses - a definition of the feature is lacking - , however, it becomes clear that Hoekstra & Hyams exclude present tense from [-realised]. Therefore, I conclude that [-realised] does not mean 'not yet realised', but rather 'not yet started'.[11] Excluding past as well as present tense, it makes as different prediction as Lasser's Non-Completedness Constraint: it falls under the Modal Hypothesis in (7) and not under the constraint in (9).

In the following chapter, I will turn to a second kind of aspect, that is, aspectual properties of RIs that are due to the lexical meaning of the selected verb. Earlier findings indicate that children do not use all types of verbs in their RIs and the constraints seem to be linked to lexical aspect. The issues discussed in the chapter on verb type are closely related to the issues discussed in the present chapter, as several scholars have argued that constraints on verb selection in RIs follow from the absence or tense or from the predominance of modal meanings in RIs.

3.2.4   Tense, modality and aspect in adult Dutch

On the surface, children's RIs lack words for tense, modality and aspect. Nevertheless, or maybe because of this, the meaning assigned to RIs is described in terms of temporal reference, modal and aspectual interpretation, as we just have seen. One of the issues in the debate on the semantic properties of RIs is how and at which level the meaning (that is obviously assigned to RIs) is represented. In order to understand the claims that have been made by others and that will be made in this thesis, some insight in the Dutch system of tense, modality and aspect marking is a prerequisite.

---

[11] See Blom (2002) for more discussion of this issue.

In adult Dutch, tense distinctions surface as inflection on the verb.[12] Present tense, notated in (10) as an empty suffix, is the unmarked tense (speech time), whereas past tense is marked by a suffix (regular) or changes in the stem of the verb (irregular).

(10)   sneeuw-**ø**   sneeuw-**de**
        snow-*pres-sg*   snow-*past-sg*

(11)   zoek-**ø**   **zocht**
        search-*pres-sg*   search-*past-sg*

Auxiliaries that take an infinitival complement mark future tense and modality. In Dutch, there are two types of future tense, exemplified in (12a) and (12b).[13] Semantically, there is a distinction between the two future forms: (12a) tends to be used for the expression of epistemic modality whereas (12b) expresses near future. In (13), some examples of Dutch modal markers are given:

(12)   a.   **zal** sneeuwen
              fut-snow-*inf*
              'will snow'
       b.   **gaat** sneeuwen
              fut-snow-*inf*
              'is going to snow'

(13)   a.   **moet** zoeken
              mod-search-*inf*
              'have/has to search'
       b.   **wil** weten
              mod-know-*inf*
              'want/s to search'

---

[12] Note that inflection does not only encode tense, but also carries agreement features. Agreement will not play a role in the discussion until Chapter 5, therefore I will not discuss the properties of agreement here. More information about agreement inflection can be found in section 5.4.5.

[13] In addition, simple present tense forms can also be used with a future denotation in Dutch.

c.   **kan** dansen
mod-know-*inf*
'am/is able to dance'

The (deontic) modal markers bear past, present or future tense features. As present tense is not overtly marked (as opposed to past tense), it is the unmarked tense in Dutch that receives the default value of speech time. Present tense *is* lexically marked by an auxiliary if the verbal predicate expresses (im)perfective aspect, however. In the examples below, the event expressed by the verbal predicate is completed at speech time. The link to speech time is made by the auxiliaries *hebben* 'have' and *zijn* 'be'. Both tense auxiliaries take a past participle as their complement:

(14)   a.   **heeft** gesneeuwd
pres snow-*part-past*
'has snowed'

b.   **is** gebeurd
*pre*s happen-*part-past*
'has happened'

In sentences that contain a so-called *prepositional infinitival complement* (*prepositional infinitival constructions* or PICs), the auxiliary *zijn* 'be' seems to have a similar function to the example given in (14b). In this case, the verbal predicate is marked for ongoing aspect and *zijn* links the event expressed by the verbal predicate to speech time:[14]

(15)   **is** aan het sneeuwen
pres on the snow
'is snowing'

In adult Dutch, finite (tensed) verbs move to first or second position, whereas infinitives remain in sentence-final base position. See (16):

---

[14] Felser (2000) analyses the preposition *aan* 'on' as the aspectual marker.

(16)   a.   De mus **valt** van het dak
               the sparrow fall-*fin* off the roof
               'The sparrow is falling off the roof'
       b.   Gisteren **viel** de mus van het dak
               yesterday fall-*fin* the sparrow off the roof
               'Yesterday, the sparrow fell off the roof'
       c.   De mus wil van het dak **springen**
               the sparrow want-*fin* off the roof jump-*inf*
               'The sparrow wants to jump of off the roof'

More details on verb movement in Dutch can be found in Chapter 1 (§ 1.2). I conclude this section with a form that plays a very important role throughout this thesis: the infinitive. In adult Dutch, the infinitive is a form that is morphologically marked with the suffix [*-en*]. Syntactically, the infinitive is distinguishable from finite verbs as the infinitive is placed in final position, whereas finite verbs are moved; see example (16c) with the sentence-final infinitival form *springen* 'jump'.

3.3   RIs in early child Dutch: a corpus study

Having introduced the relevant hypotheses and definitions, I will turn to the data of Abel, Daan, Josse, Laura, Matthijs and Peter. What do their RIs have to say about the claims that have been made? Recall that there were three hypotheses, that is: (i) the No Tense Hypothesis according to which RIs lack tense and hence, they can be used to denote past, present and future tense (Behrens, 1993; Wijnen, 1997; Lasser, 1997), (ii) the Modal Hypothesis stating that RIs denote modal events and do not denote past and present events (Ferdinand, 1996; Ingram & Thompson, 1996; Hoekstra & Hyams, 1998), and, finally, (iii) the Non-Completedness Constraint that bans RIs that denote completed events (Lasser, 1997). In this section, I put these hypotheses to the test. The assignment of a specific interpretation to RIs could indicate that the construction contains structural features that determine the denotation and hence, could be used as an argument for early linguistic knowledge. In order to carry out this investigation, it must first be decided what counts as an RI (§ 3.3.1), and what counts as a modal, present, past or completed RI (§ 3.3.2).

3.3.1   The selection of RIs

In Dutch, there are two cues to distinguish RIs from finite sentences: the infinitive is morphologically marked and is placed in a distinct position. Nevertheless, when I analysed the child data it was not always unambiguously clear whether or not a child utterance was really finite. The Dutch infinitive is similar to the present plural form and therefore ambiguities arise, as in the case of *wij eten* 'we eat'. I excluded such utterances. A problem was posed by one-word-utterances, since these are ambiguous: a bare *eten* 'eat' can either be an RI with a dropped subject or a present plural with dropped subject. I decided to count these examples as RIs, for the following reasons: in finite sentences with overt subjects, the children used few plurals. The subject nearly always referred to the speaker or the addressee (respectively first and second person singular). Given this observation, I concluded that the likelihood that one-word-utterances with a verb form ending on [*-en*] were present plurals, was low and that in this case a conservative strategy would lead to the unnecessary exclusion of valuable data. Parasitic RIs that are used to give an answer to a question were excluded, as they are not syntactically independent.[15] This dependency is illustrated in (17) and (18). The question in (17) would yield a non-modal use of the infinitive, while in (18) the infinitive would be modal.

(17)   Q: Wat ben je aan het doen?
              'What are you doing?'
         A: Fietsen
               'Cycling'

(18)   Q: Wat wil je vandaag doen?
              'What do you want to do today?
         A: Eten
               'Eat'

---

[15] With this exclusion, I follow other studies on RIs (e.g. Lasser, 1997).

3.3.2   The criteria for assigning interpretations

In the transcripts, utterances of Abel, Daan, Josse, Laura, Matthijs and Peter
have been annotated with codes for verb forms that appear in the child
utterances.[16] Two people assigned modal/temporal and aspectual
interpretations independently. Recall that according to the definition I used,
the modal and temporal interpretation can be derived from each other. First
of all, an utterance is marked as being modal or non-modal. Utterances that
received a non-modal interpretation were provided with more precise
aspectual codes for events that were completed, ongoing or prospective at
speech time. The coding-system is represented in the diagram on the
opposite page. The system is based on Palmer's (1986) division into
epistemic, deontic and dynamic modality.

The schema is applied to utterances with verbs. The available information
from the transcripts was used in order to determine the interpretation of an
utterance; this was mainly linguistic information and to a lesser degree
circumstantial or contextual information. Utterances labelled 'modal' had to
meet one of the following criteria: preceding or subsequent parental
utterances suggested a modal interpretation, or the contexts suggested a
modal interpretation. In Appendix 3.1 (p. 240) some examples from the
transcripts are given that illustrate the codes and show how the
interpretations were assigned. The specific modal labels in schema 3.1 are
irrelevant in section 3.3 (because only a distinction between modal and non-
modal use of an utterance will be made), but they will play a role in section
3.4, particularly, the observation that modal RIs are used to express *I want*
(dynamic necessity or volition) or *You must* (deontic necessity) becomes
important later on in this chapter. These two modal expressions correspond
with two usages: the 'dynamic necessity' RIs are used by the children to
express a wish, while the 'deontic necessity' RIs are used for commands.
Modal distinctions are relevant in Chapter 4, as different modalities appear
with different verb types.

---

[16] The codes follow the CHAT conventions (MacWhinney, 1995). The CHILDES tools
could be used for various searches and counts. This facilitated the analysis of the data.

```
                          ┌──────────────┐
                          │   Utterance   │
                          └──────────────┘
                                 │
                 ┌───────────────┴───────────────┐
        ┌────────────────────┐          ┌────────────────────┐
        │       Modal         │          │     Non-modal       │
        │  ┌──────────────┐   │          │  ┌──────────────┐   │
        │  │   dynamic     │   │          │  │  completed    │   │
        │  │               │   │          │  └──────────────┘   │
        │  │ a subject's wish (necessity)  │  ┌──────────────┐  │
        │  │ a subject's ability (possibility)│ │  ongoing    │  │
        │  └──────────────┘   │          │  └──────────────┘   │
        │  ┌──────────────┐   │          │  ┌──────────────┐   │
        │  │   deontic     │   │          │  │  prospective  │   │
        │  │ obligation, requirement (necessity)│ └──────────┘  │
        │  │ permission (possibility)│  │                      │
        │  └──────────────┘   │          │                     │
        │  ┌──────────────┐   │          │                     │
        │  │  (epistemic)  │   │          │                     │
        │  │ a proposition is necessarily or possibly │        │
        │  │    true/probability │  │                          │
        │  └──────────────┘   │          │                     │
        └────────────────────┘          └────────────────────┘
```
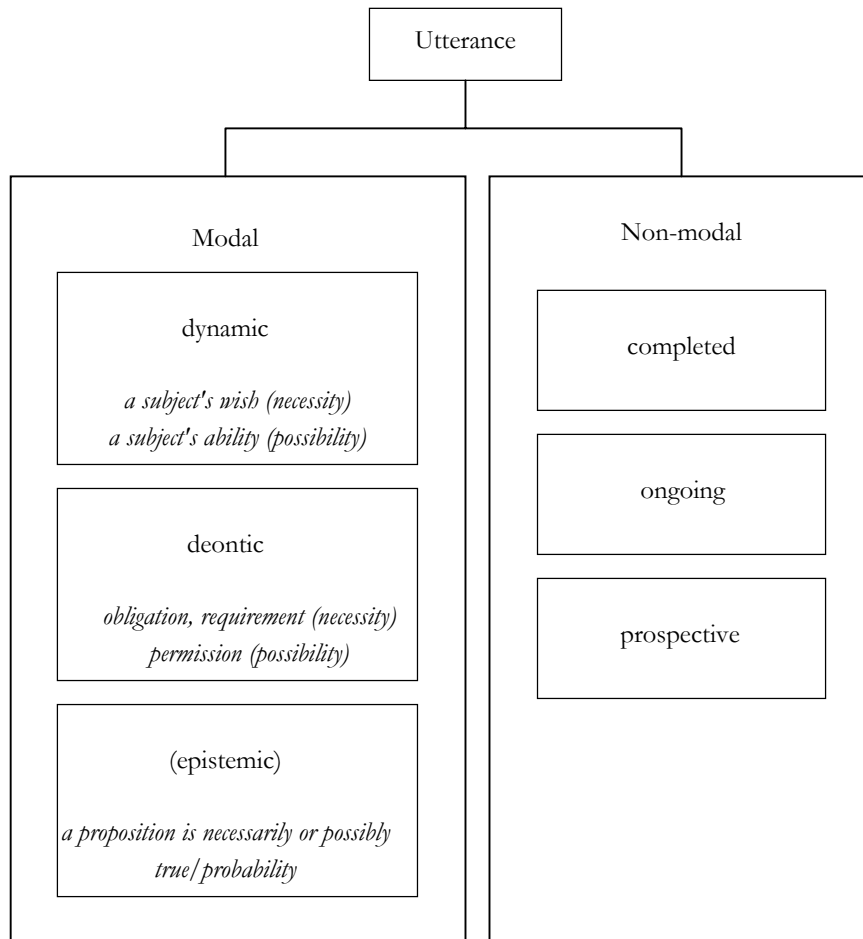
Figure 3.1: Modal coding system (utterances with verbs)

As for a *definition* of modality, I will use the definition in (7): RIs denote modal events and do not denote past and present events. This is a working definition that does not state what modality is but rather what modality does or does not do: modal utterances do not describe an event that occurs simultaneous to speech time or happened prior to speech time. In modal utterances, the predicate denotes an event that (possibly) takes place after speech time. Core modal notions such as possibility and necessity do not play an important role in theories on the modal denotation of RIs; modality is basically defined in temporal terms, which is directly reflected in the

working definition. The definition is definition is based on previous work on the interpretation of RIs. It captures future tense, and more specific usages, such as commands, desires and intentions. Epistemic modality is excluded, which might seem inappropriate from a theoretical point of view as it implies an adjustment of the notion modality. However, it suffices for a concise description of the meaning of RIs, as children in the RI age - roughly between the two and three years old - do not yet use epistemic modality. In Chapter 4, I will expand on the issue of cognitive immaturity of children in the RI age, the role it plays in their language production, and particularly, on how cognitive factors influence the verb types that appear in RIs (see § 4.4.5).

The implication of my definition of modality is that RIs that were coded as 'non-modal prospective' are included in the set of modal RIs.[17] Under the strict interpretation of modality on which the coding system was based, these RIs are not modal as modality is about possibility and necessity. According to the working definition, prospective RIs denote an event that (probably) takes place after speech time. Hence, these RIs are counted as modal. Unclear cases, i.e. the RIs that could not be assigned an unambiguous interpretation, were not included in the analysis. The application of a conservative strategy lead to the exclusion of 317 RIs out of a total number of 1565.

3.3.3   The denotation of RIs

Figure 3.2 gives the percentage of RIs that received a modal interpretation; this is the percentage out of the total number of interpretable RIs. Table 3.1 gives the corresponding numbers. For all children, modal RIs are more frequently used than non-modal RIs.

---

[17] These RIs can be said to express an intention. With respect to the interpretation of RIs, it was often impossible to decide whether an RI expressed a wish or an intention. As both are included in the set of modal RIs, this distinction (or the impossibility to make this distinction) is irrelevant for the counting of modal RIs. The distinction does become relevant in the next chapter, though. as sentences that express a wish or an intention behave differently with regard to the selection of verb types.
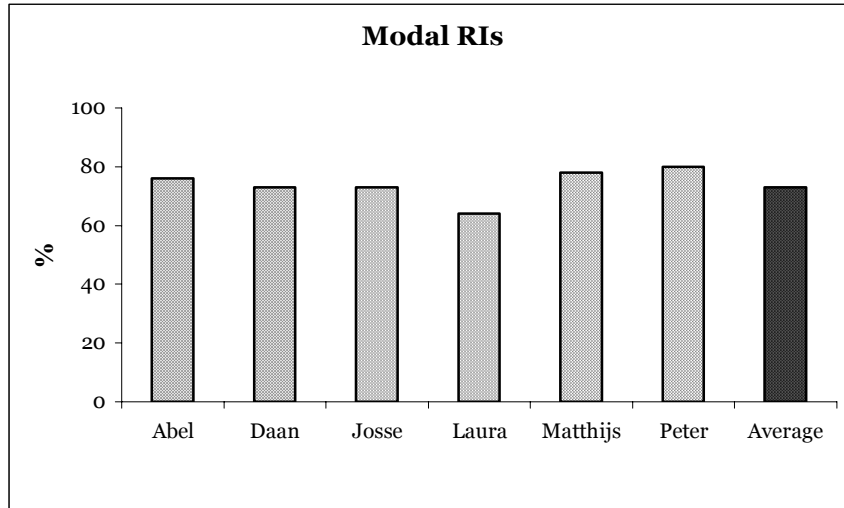
Figure 3.2: Percentage of modal RIs, data from all six children and average percentage

Table 3.1: Number of interpretable RIs, number of modal RIs and percentage of modal RIs, data from all six children

	N_{RI}	N_{MODAL}	$\%_{MODAL}$
Abel	120	91	76
Daan	157	115	73
Josse	206	150	73
Laura	314	200	64
Matthijs	254	199	78
Peter	197	157	80
Sum	*1248*	*912*	-
Average	-	-	*73*
SD	-	-	*6*

Examples of modal RIs and non-modal RIs are given in (19) and (20), respectively:

(19) a.   Peter bal pakken   Peter 2;01.27
            Peter ball get-*inf*
            context: Peter wants to get the ball
     b.   vrachtwagen emmer doen   Matthijs 2;04.24
            truck basket do-*inf*
            context: Matthijs wants the investigator to put the truck in the
            basket
     c.   op kist zitten   Josse 2;08.04
            on box sit-*inf*
            context: Josse wants his mother to sit on the box

(20) a.   ah, mij bril vallen [= mijn bril valt]   Abel 2;05.27
            ah, my glasses fall-*inf*
            context: his glasses are falling
     b.   poffie ginke [= koffie drinken]!   Daan 2;01.21
            coffee drink-*inf*
            context: people on television are drinking coffee
     c.   boot svaje [= varen]   Laura 2;04.01
            boat sail-*inf*
            context: refers to a picture with a sailing boat

3.3.4   A closer look at the non-modal RIs

The *non-modal* RIs represented in Figure 3.2 and Table 3.1 nearly always, i.e.
in 95% of the cases, refer to events that are simultaneous to speech time:
they are ongoing and denote present tense. In the files of all six children, I
found only 17 RIs that denoted an event that took place *prior* to speech time
and that was completed at speech time (which is the only criterion that can
be used for past/completed interpretations). These 17 RIs are included in
the non-modal RIs of the preceding section, and cover 5% of the non-
modal RIs and 1 % of all interpretable RIs. Some examples are given in (21)
and (22):

(21) *MAT: Ieke spugen !   Matthijs 1;11.24
             Ieke [= baby sister] throw-*inf* up
      *MOT: ging Ieke spugen ?
             went Ieke throw up

(22) *MOT: vond je dat leuk ?   Laura 3;04.06
               find you that nice
      *LAU: ja
               yes
      *MOT: in grote zwembad
               in big swimmingpool
      *LAU: met mij emmertjes spelen
               with me baskets-*dim* play-*inf*

Thus, the nonmodal RIs of the Dutch children are predominantly present/ongoing events and hardly ever denote past/completed events.

3.3.5   Comparison with other studies

The findings confirm various previously reported results from Dutch and German child language. My findings strongly resemble those of Lasser (1997) for German in the sense that there are only very few past/completed RIs. With regard to the temporal and modal interpretation of RIs, the results indicate a free temporal/modal reference, as RIs are not restricted modal usage. Although there is a preponderance of modal RIs - a Modal Reference Effect or MRE as Hoekstra & Hyams (1998) called it - RIs are also used for non-modal events taking place at speech time or took place before speech time. Haegeman (1995) and Wijnen (1997) reported similar observations for Dutch child language, Behrens (1993) and Lasser (1997) found similar results in German child language.

Haegeman's (1995) case study of the Dutch child, Hein (2;4-3;1), shows that 56 % (50/89) of his RIs have a modal interpretation, 7 % (6/89) are non-modal and the rest is ambiguous.[18] Wijnen (1997) reports a higher percentage of modal RIs in the data of four Dutch children. His results are given below:

---

[18] In addition, 7% (6/89) are elliptical answers of the type *Wat ben je aan doen? lopen (* 'What are you doing? walking'). In Dutch, the elliptical answer contains an infinitive and does not contain a finite form. Such an utterance could count as an RI. However, these elliptical answers should be distinguished from RIs because they do not occur as independent utterances (unlike RIs).

Table 3.2: The modal/temporal interpretation of RIs in the spontaneous speech data of four Dutch-speaking children, data Wijnen (1997)[19]

	N_{RI}	% MODAL/FUTURE	% PRESENT	% PAST
Josse	272	73	18	9
Matthijs	677	86	12	2
Niek	348	94.3	4.3	1.4
Peter	493	87	9	4
Average		*85*	*11*	*4*

The data Wijnen analysed partially overlap with my data; we both examined utterances from Josse, Matthijs and Peter. For Josse, we found the same percentage of modal RIs, but for Matthijs and Peter, my percentages of modal RIs are lower: 79 % (Wijnen: 86 %) and 76 % (Wijnen: 87 %), respectively. In § 3.4.1, I will argue that such differences may easily occur due to the methodology that is used, i.e. the interpretation of transcriptions of spontaneous speech data.

The findings for German child RIs are similar to the Dutch findings. In her study of the temporal reference in early German, Behrens (1993) studied six German children. She found that the temporal reference of RIs was free. Behrens provided quantitative data from only one child, Simone, since only this corpus was extensive enough for a quantitative analysis. Lasser analysed data from the German children, Simone and Andreas, and found that respectively 73 % (N = 130) and 69 % (N = 108) of all interpretable RIs in

---

[19] Below, the age ranges of the children Wijnen examined, are given:

Child	*Age range*
Josse	2;0.7 - 2;06.22
Matthijs	1;11.10 - 2;8.5
Niek	2;7 - 3;2.13
Peter	1;9.6 - 2;1.26

these data were modal. The rest received a present (predominant) or past (rare) interpretation.

3.3.6   The denotation of finite sentences

Is the temporal, modal or aspectual interpretation of RIs a specific property that relates to the absence of a finite verb? If this is the case, then it is expected that sentences that are entirely comparable with RIs but that are *finite* instead of non-finite, receive a different interpretation. In the files of the six children, an abundant number of finite sentences is available that can be compared. There are two types of finite sentences that differ from RIs only with respect to finiteness: simple finite sentences (SFs) and sentences with periphrastic verbs (PVs). The two constructions are illustrated in (23a) and (23b):

(23)   *Simple finite verb* (SF)
    a.   Lola rent de hele dag
          Lola run-*fin* the whole day
          'Lola is running the whole day'
    Periphrastic verb (PV)
    b.   Agent Cooper moet de wet volgen
          Agent Cooper must-*fin* the law follow-*inf*
          'Agent Cooper has to follow the law'

Copula and auxiliaries do not appear in RIs (De Haan, 1987; Jordens, 1990; Wijnen, 1997). Therefore, I excluded SFs with copula and auxiliaries and narrowed the set of SFs down to sentences with a *thematic* verb. Henceforth, this subset of the set of SFs will be referred to as SF_θ, because thematic verbs assign theta-roles (θ). As the main verb in PVs is thematic, this selection does not apply to PVs. I furthermore restricted the set of PVs to PVs that stand in close relation to RIs, hence the PVs with an infinitival complement; clauses containing auxiliary + past participle are excluded. Note that with the label 'PVs with infinitival complements' I generalise over genuine PVs and PICs, i.e. prepositional infinitival complements. Although these are clearly a different kind of construction in adult Dutch, Dutch children initially do not seem to make a distinction between the two. I will give some examples to illustrate this point. In adult Dutch, PICs contain the

preposition *aan* 'on'. *Aan* selects a nominal infinitive; the infinitive in this construction is preceded by the definite determine *het*:

(24)   De poppen zijn aan het dansen
          the puppets are on the dance-*inf*
          'The puppets are dancing'

In my sample, all children omitted the determiner and nearly always the preposition:

(25)   a.   nee, die s een boek leese papa   Laura 2;5
                   no, that is-*fin* a book read-*inf* daddy
                   'No, daddy is reading a book'
          b.   is takelen   Matthijs 2;5.1
                   is-*fin* hoist-*inf*
                   'The machine is hoisting'
          c.   zijn put openmaken   Josse 2;7.20
                   are-*fin* ditch open make-*inf*
                   'They are opening the ditch'

The surface structures of the sentences in (25) do not show whether or not the children distinguish between PVs and the precursors of PICs in (25) in their underlying representation: the children say *wil lezen* 'wants read-*inf*' next to *is lezen* 'is read-*inf*', with the auxiliaries denoting modality and present tense, respectively. The preposition and nominal infinitive that characterise PICs in adult Dutch appear only later in child Dutch.[20,21] Differences in

---

[20] Only Peter, in the final files I examined from him, used the preposition *aan* 'on', exemplified in (i):

(i)   varken is trui aan breien   Peter 2;3.21
           pig is-*fin* pullover on knit-*inf*
           'The pig is knitting a pullover'

[21] Note that in adult Dutch, it is possible to use constructions like the ones in (ii). These constructions have the specific connotation that the agent is absent from the current setting:

(ii)   Jan is lezen
           Jan is read-*inf*
           'John is away to read'

meaning could indicate a formal distinction. However, as will be shown shortly, not all PVs in early Dutch child language are modal: *is* + *inf* 'is + inf' (which I interpret as a precursor of PICs, a pre-PIC) and PVs, like *gaat* + *inf* 'goes + inf' or *doet* + *inf* 'does + inf', are all used by children for events that are ongoing at speech time. It is unclear if the children use the pre-PIC to describe the ongoingness (and mark aspect) and the PVs to denote present tense.[22] Before I turn to the quantitative results on the children's interpretation of PVs, I will first describe their use of the other finite sentence type: SF_θ.
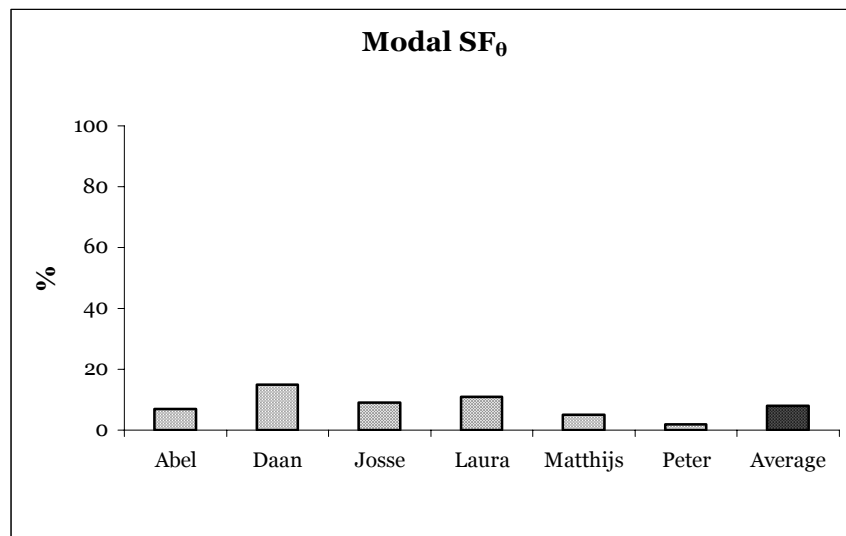
Figure 3.3: Percentage of modal SF_θ, data from all six children and average percentage

Young children do not use the construction with this specific meaning, though, e.g. the action described in (ii) is performed in the presence of the speaker.

[22] According to Felser (2000), the preposition *aan* 'on' is the aspectual marker in this construction. The implication of the early omission of *aan* may be the aspectual meaning of PICs is specified later on and that the precursors of PICs denote only present tense. This would imply that Dutch children make a 'tense-before-aspect' distinction, though considering that there are various observations in support of the 'aspect-before-tense' hypothesis (cf. Bronckart & Sinclair, 1973; Antinucci & Miller, 1976; Wagner, 1998), this hypothesis may be unlikely. In general, there is not much known about the development of PICs in child Dutch, as far as I know. It would be an interesting subject for future research.

A comparison between Figures 3.2 and 3.3 reveals a major contrast between RIs and SF_θ with respect to modality: RIs occur most frequently in modal contexts, while SF_θ are most often non-modal. The average percentage of modal use of RIs is 73 % (SD = 6) while on average only 8 % of the SF_θ is used modally (SD = 5). Thus, the difference is highly significant: the modal use of RIs is approximately 9 standard deviations larger than the modal use of SF_θ.

Table 3.3 gives the exact numbers and percentages that correspond to Figure 3.3.

Table 3.3: Number of interpretable SF_θ, number of modal SF_θ and percentage of modal SF_θ, data from all six children

	$N_{SF\theta}$	N MODAL	% MODAL
Abel	119	9	7
Daan	289	43	15
Josse	137	12	9
Laura	157	17	11
Matthijs	95	5	5
Peter	424	7	2
Sum	*1221*	*93*	-
Average	-	-	*8*
SD	-	-	*5*

(26) and (27) give examples of modal and non-modal SF_θ, respectively:

(26) a.   oh, valt bijna om Abel   Abel 2; 07.15
            oh, fall-*fin* almost down
            'Oh, it almost falls down'
      b.   ik ga ook naar de dok, of niet?   Laura 3;04.06
            I go-*fin* also to the doctor, or not
            'I will go to the doctor too, don't I?'

(27) a.   ik hoor paatje niet   Laura 2;04.15
              I hear-*fin* horse-*dim* not
              'I do not hear the little horse'
      b.   Daan ligt in de wieg   Daan 2;04.14
              Daan lie-*fin* in the crib
              'Daan is lying in the crib'

If we now turn to the PVs, and do not consider the deviating results from Peter, a picture emerges that is the opposite from Figure 3.3. In (28), some examples of PVs used by the children, are given.

Figure 3.4: Percentage of modal PVs, data from all six children and average percentage

Table 3.4: Number of interpretable PVs, number of modal PVs and percentage of modal PVs, data from all six children

	N_{PV}	N_{MODAL}	$\%_{MODAL}$
Abel	79	68	86
Daan	108	92	85
Josse	108	90	83
Laura	156	102	65
Matthijs	149	128	86
Peter	91	32	35
Sum	*691*	*512*	*-*
Average	*-*	*-*	*73*
SD	*-*	*-*	*20*

(28) a. Ik moet daar zitten he   Matthijs 2;10.22
         I must-*fin* there sit-*inf* huh
         'I have to sit there, haven't I?
     b. mag ik grote blok bouwen?   Abel 2;07.15
         may-*fin* I big block build-*inf*
         'May I build the big block?'

The difference between SF_0's and RIs with regard to modal use is highly significant. For PVs and RIs, this is less clear, partially because the average of the six children is affected by Peter's strikingly low proportion of modal PVs: when Peter is excluded the average percentage of modal PVs goes up from 73 % to 81 %. When we use a permutation test to see whether or not there is a difference between RIs and PVs with regard to modality, it turns out that inclusion of Peter leads to no difference between RIs and PVs ($p = 0.518$). When Peter is excluded and we compare the PVs and RIs of the other five children, there is a significant difference and PVs are more often modally used ($p = 0.02$). As to the question why Peter deviates, it turns out that there are two factors that contribute to his high number of non-modal PVs. Firstly, Peter uses *gaat* + inf 'goes + inf' more than any of the other children to denote ongoing events. Secondly, he uses many

precursors of PICs that have the format *is (aan)* + inf 'is (on) + inf'. This is especially true for the last files that I examined from him (i.e. 2;03.21) where he is describing pictures in a booklet. In (29), two examples are given from the ongoing use of the auxiliary *gaan* 'go' and the precursors of PICs, respectively:

(29)   a.   hij gaat zo hoepla doen   Peter 2;03.21
                he go-*fin* so hoopla do-*inf*
                'He is now doing hoopla'

        b.   kikker is aan lopen   Peter 2;03.21
                frog is-*fin* on walk-*inf*
                'Frog is walking'

Summarising the observations so far, we can say that Dutch-speaking children use RIs for modal as well as non-modal events. They use RIs more frequently to denote modal events than non-modal events. In this respect, there is a Modal Reference Effect (MRE) in early child Dutch. Non-modal RIs are nearly always used to denote activities that are ongoing at speech time and hence, have a present tense reference. The denotation of RIs contrasts with the denotation of children's early finite sentences: SF_θ 's are restricted to present tense while PVs are predominantly modal. In general, PVs are slightly more frequently modal than RIs, although there is variation per child. In my sample of six children, Peter uses many non-modal PVs.

3.3.7   Evaluating hypotheses

Even though the RIs of the six Dutch-speaking children appear more often in modal than in non-modal contexts, the Modal Hypothesis is not confirmed because there is no restriction to modal usage. The theories underlying the Modal Hypothesis assume that the modal meaning assigned to RIs is structurally represented. The modal meaning is encoded by either a covert modal auxiliary (Ferdinand, 1996; Ingram & Thompson, 1996) or an infinitival suffix (Hoekstra & Hyams, 1998). In particular, the strong claim made by the second proposal, which I will dub the *Infinitival Morphology Hypothesis* or simply IMH, is not confirmed. According to the IMH, infinitival morphology yields a modal meaning. Hence, as each RI contains

infinitival morphology, the prediction is that each RI must be modal.[23] Even a low percentage of non-modal RIs in a language in which the infinitive has infinitival morphology provide evidence against this hypothesis. Behrens (1993) addressed the issue of what counts as 'restricted to a certain meaning' and proposed a 90 % criterion.[24] I will apply a somewhat lower threshold and assume that Hoekstra & Hyams (1998) own "criterion" of 85 % modal use suffices.[25] The rate of modal RIs in my data differs per child; it ranges from 64 % to 78 %. A calculation of the average rate shows that approximately one quarter of the RIs in Dutch child language is non-modal, which is 10 % below the percentage that is required to conclude that the IMH hypothesis is confirmed and even 15 % below the 90 % criterion. The No Tense Hypothesis, on the other hand, is confirmed since children use RIs with different temporal denotations. The percentage of past/completed RIs is so low (less than 5 %) that the Non-Completedness Constraint (NCC) is confirmed.

---

[23] Ferdinand (1996) argued that RIs contain a null modal. However, in the original AUX drop proposal of Boser et al. (1992), the dropped auxiliary could either be a null modal or a dummy *do* or *go*. These dummy verbs would give rise to a non-modal interpretation, thus, non-modal use of RIs is not per se excluded in the AUX drop account. Recall that Ferdinand needed the modal drop hypothesis to explain the eventivity constraint. An interesting question would be whether she could maintain her claim if the matrix verb is a dummy or light verb instead of a modal. Thus, unlike the IMH, the AUX drop hypothesis does not necessarily predict a modal meaning. In this respect, the AUX drop hypothesis explains the range of interpretations assigned to RIs more adequately than the IMH. However, as has already been pointed out in Chapter 1, I think that the AUX drop hypothesis faces methodological problems: in Chapter 5, it will be shown that when children start to use RIs, they do not use any combinations of auxiliary and infinitive. Hence, there is no indication whatsoever that the auxiliary is dropped in RIs.

[24] An independent criterion is needed. Modal use in RIs could be compared to modal use in the set of all finite sentences or FINs (collapsed SFs and PVs). The modal use of FINs could give an indication of whether or not RIs are disproportionally often modal. In Chapter 5, however, I will argue that modality in RIs and FINs is not independent: the more specified modal FINs there are, the less modal RIs there will be. To decide whether or not RIs are relatively often modal, the modal RIs have to be calculated as a proportion from modal RIs and FINs (all utterances with verbs that children use to talk about events). In § 3.4.9, I come back to this issue.

[25] Although Hoekstra & Hyams (1998) do not use any explicit criteria for what counts as modal, they take Wijnen's (1997) average of 85 % modal use in the data of Dutch children as evidence for the MRE. Thus, according to them, the predicate 'modal' can be applied to RIs even though 15 % of the RIs are not modal.

3.3.8 Towards a new hypothesis

By putting interpretations of RIs into broad categories (like modal vs. non-modal), one abstracts away from the more specific interpretations that are actually assigned to RIs. Based on close examination of the RIs produced by the six children, I conclude that Dutch child RIs receive at least six different interpretations. Consider Daan's RIs in (30):

(30) boekje lezen   Daan 2;01.21
      book-*dim* read-*inf*

Daan utters (30) with a neutral intonation; there is no prosodic cue that marks a particular illocutionary force, such as a rise, which would suggest that Daan is asking a question.[26] As explained before, the morpho-syntactic properties of RIs exclude that RIs are imperatives. Hence, I (as an interpreter) take (30) to be a declarative sentence. A first possible interpretation of (30) is paraphrased in (31):

(31)   You have to read a book   DEONTIC NECESSITY (modal)

The deontic necessity interpretation arises from cues given by (i) the context (Daan walks to somebody and hands her the book, for instance), (ii) from what I know of Daan (he cannot read himself) and (iii) from the shared history of Daan and the other person (she always reads books to him). Based on the same cues, (30) can also receive a slightly different, volitional, interpretation:

(32)   I want to read a book   VOLITION (modal)

Note that it can be hazardous or even impossible to decide between (31) and (32). Commands and desires are closely related and Daan can use (30) not to make a statement but also as an indirect command. According to Lyons (1977: 826):

---

[26] In the CHILDES corpora I studied, utterances that have a question intonation are marked with a question mark.

"It seems clear that these two functions are ontogenetically basic, in the sense that they are associated with language from the very earliest stage of its development in the child. It is equally clear that they are very closely connected. It is a small step from a desiderative utterance meaning "I want the book" to an instrumental utterance meaning "Give me the book"; parents will commonly interpret the child's early desiderative utterances and mands, thereby reinforcing, if not actually creating, the child's developing awareness that he can use language in order to get others to satisfy his wants and desires."

Daan's behaviour leads me to a third interpretation. If he is taking steps to start reading a book (and he is not reading yet), I may infer that (30) denotes an intention:

(33)   I am going to read a book   INTENTION (modal)

According to the definition of modality that I apply, (31), (32) and (33) are modal. However, when Daan utters (30) while he himself or someone else in the room is reading a book, one of the two non-modal interpretations in (34) or (35) can be assigned:

(34)   X is reading a book   DESCRIPTION (non-modal)

(35)   This is 'reading a book'   PRESENTATIONAL (non-modal)

In the interpretation of (34), the predicate *boekje lezen* is used to denote the event 'read book'. In the interpretation of (35), Daan performs a metalinguistic act and modifies the predicate *boekje lezen* 'read book'. Note that 'modification' does not only mean that information is added and the content of the predicate changes; 'modification' can also mean that the content of a predicate is confirmed. An interpretation as in (35) is assigned when Daan utters (30) in the middle of a labelling game, for instance. Although the difference between (34) and (35) is interesting and may even be crucial to our understanding of RIs, I do not see how they can be distinguished systematically. Therefore, I did not distinguish between the

two readings.[27] Finally, I encountered RIs in the corpus data that were assigned a past interpretation (when Daan is telling a story about a past event):

(36)   X was reading a book   NARRATION (non-modal)

The examples above indicate that Dutch child RIs allow for a wide range of interpretations and that the interpretations assigned to children's RIs are based on a complex of extra-linguistic factors such as contextual information, shared history, knowledge of the world, etc. In this respect, the example in (37) is illustrative. In (37), Laura says *pakke* 'glue'. With this utterance, she refers to the event of gluing. The sequence of questions that follows shows how her mother tries to understand what Laura exactly wants to express. Her attempt shows that RIs allow for various different interpretations.

---

[27] In Blom (2002), I argue that one possible strategy to circumvent the problem of the existence of non-modal RIs is to assume that the non-modal RIs differ fundamentally from the modal RIs. One way in which they could differ is with respect to their syntactic category. In adult Dutch, infinitives can be verbal as well as nominal. It could be hypothesised that this difference is reflected in the RIs of Dutch children: modal RIs are verbal and non-modal RIs, depicting an ongoing activity, are used as labels and hence, are nominal. I hypothesised that the nominal infinitives appear less often with subjects/objects than the verbal RIs. However, although I found that non-modal RIs contained less subjects than modal RIs, there were various non-modal RIs with subjects. It was argued that the difference in subject use could be a developmental effect as most non-modal RIs were used in the earliest stages (see Chapter 5) and studies on subject drop (Haegeman, 1995) show that subject drop is also a property of Dutch child language that is related to the earliest stages. However, the developmental observations that will be presented in Chapter 5 shed a different light on this problem, because the development of subject drop in RIs differs from the development of subject drop in general. Between stages III and IV, subject drop in RIs shows an increase, which implies that the appearance of fewer overt subjects in non-modal RIs than in modal RIs cannot be explained as a developmental effect.

(37) *LAU: **pakke, hoor**
                    glue-inf, *interjection*
        *MOT: ga je plakken?
                    go-*fin* you glue-*inf*
                    'Are you going to glue?'
                    […]
        *MOT: is dat plakken?
                    is-*fin* that glue-*inf*
                    'Is that gluing'
        *MOT: ben je daarmee aan het plakken?
                    are-*fin* you therewith on the glue-*inf*
                    'Are you glueing with that?'
        *LAU: eh.   Laura 2;06.10

Example (37) exemplifies the semantic unspecificity of RIs and the result of this, namely how we (as interpreters) are guessing what children intend to say when they use an RI.

However, suppose that the meaning of RIs is intrinsically free. Considering the results presented in this chapter, two questions arise: how is it that (i) completed RIs hardly occur and obey the NCC, and (ii) that modal RIs are predominant and that RIs display a MRE? I will take up these two questions in Chapter 5 (§ 5.3).

3.4   Differences across languages: experimental data

As a follow-up on the corpus study, I decided to conduct an experiment.[28] An important motivation for an experiment was the difficulties that arose when interpretations were assigned to the RIs in the corpus data. In § 3.4.1, these difficulties will be explained. Secondly, I was struck by a difference in the interpretation of Dutch and English RIs discussed in the work of

---

[28] The experiment was conducted in collaboration with Evelien Krikhaar (Groningen University). We carried out the pilot studies together. After this, Evelien Krikhaar took care of the execution of the final experimental design with Dutch-speaking children while I tested the English-speaking children.

Hoekstra & Hyams (1998). An experiment would make it possible to compare the two languages in a more principled and controlled way than the previously reported corpus data allowed for (§ 3.4.2).

3.4.1   More reliable interpretations

Recall that Wijnen (1997) reported more modal use in RIs in the data of two Dutch-speaking children than I did, even though we examined data from the same children (§ 3.3.5). In this section, I will illustrate that such differences can easily occur with the methodology we used, i.e. assigning interpretations to corpus data. The difference between our studies illustrates the severe methodological difficulties of corpus analysis and the limitations of the analysis of corpus data with regard to semantic questions. Results that are based on interpretations of corpus data may show quite some variation due to (i) too little information, (ii) subjective interpretations, and (iii) the situation and the activities that were carried out when the children were taped (yielding variation over sessions and children).

Corpus data that are transcriptions of audiotapes often provide little extra-linguistic information.[29] Especially circumstantial information is necessary to interpret elliptic utterances such as RIs (§ 3.3.8). A simple example will illustrate this problem. Assume that a child says *bal gooien* ('ball throw'). As long as the interpreter does not know anything about the state or position of the ball, this utterance can be modal as well as non-modal: the ball could be going to be thrown, could have to be thrown, would be thrown or could have been thrown (when the child tells a story).[30] To minimise the risk of assigning a wrong interpretation, more than one researcher must interpret the utterance (and the interpreting must be done independently), and the application of a conservative strategy ('exclude the utterance from the analysis in case of uncertainty or ambiguity') is advisable. Nevertheless, even if the risk of interpretation mistakes is minimised,

---

[29] Video recordings are, for instance, much more informative than audio-tapes as they show the surrounding space, direction of the gaze and gestures, which are all useful sources of information.

[30] In paraphrase, the modal version could be 'you have to throw the ball' or 'I am going to throw the ball'. The non-modal version would either be 'You are throwing the ball' or 'I am throwing the ball'.

problems still remain. The researcher's interpretation is often deduced from an adult response to the child's utterance. This response is based on the adult's interpretation and may be incorrect. Other disadvantages are that studies are difficult to compare as researchers use their own criteria and definitions (the *subjectivity factor*). Second, the data are not only influenced by the interpreter but also by the circumstances in which the data are collected. Not every situation yields a similar amount of modality (the *situation factor*). For example, a play situation elicits much modality as the child expresses wishes and gives commands to the adult while playing. When adult and child read a booklet or look at pictures, there is less modality. Presumably, the distribution of situations over transcripts differs when different files are studied. As a consequence, the number of modal utterances may be different per file.[31] Coming back to the discrepancy between Wijnen's results and mine (§ 3.3.5), the subjectivity factor may play a role here. But also the situation factor, since we did not study exactly the same files of these two children.

Many of the problems signalled above can be circumvented in an experiment. With an elicitation task in which the context, i.e. the interpretation that is to be encoded, is given, it is possible to test which verb forms children use to express modal or non-modal meanings (under the assumption that the child interprets the depicted situation in an adult-like way). Doing a corpus study, the researcher assigns the interpretation afterwards to a given verb form. In an experiment, this situation can be turned around: the interpretation is given and a specific verb form is triggered.

3.4.2   Valid cross-linguistic comparisons

In their study on the meaning of RIs, Hoekstra & Hyams (1998) compared Dutch/German RIs with English RIs and concluded that the data showed a significant cross-linguistic difference. Their conclusion was that RIs in Dutch and German child language are modal while English RIs are not. They related the difference to a morphological difference between Dutch

---

[31] This implies that a meta-analysis is required that compares datasets with regard to the amount of modality that is used.

and German RIs, on the one hand, and English RIs on the other. The former contain infinitival morphology, whereas the latter contain bare stems. As their claim is that infinitival morphology denotes a modal meaning, the cross-linguistic predictions of the *Infinitival Morphology Hypothesis* (IMH) are:

(38)   The IMH cross-linguistically
     (i)   Dutch RIs are restricted to modal use
     (ii)   English RIs are not restricted to modal use and have a free
           reference

The empirical fundament for (38) comes from a comparison of data reported in various studies.[32] This comparison focuses on percentages reported by others and only little is known about the data behind the percentages. As discussed before, this is a risky method with a number of uncontrolled factors that possibly influence the results.[33] The use of the same experimental design for different languages, however, enables a methodologically valid cross-linguistic comparison. In this way, it can be tested in a relatively controlled and principled fashion whether or not there is a difference between the meaning of Dutch and English RIs. Previously, Schönenberger, Pierce, Wexler & Wijnen (1995) attempted to compare the interpretation of RIs in Dutch and English child language through an experiment. In § 3.4.13, I will discuss this study and compare the obtained results with my results.

Before I turn to the experimental design, the protocol, the subjects and the results, I want to comment on the relation between the IMH, its predictions in (38) and my conclusions earlier in this chapter with regard to Dutch child language. Earlier in this chapter, it was concluded that (38i) is not borne out: Dutch corpus data show that there is no restriction to modal use. This means that there is no explanation for the size of the MRE, because the theories we have available predict an MRE that is larger than the observed

---

[32] Wijnen (1997) and Behrens (1993) for Dutch and German, respectively, Deen (1997) and Madsen & Gilkerson, (1999) for English.

[33] Especially about the English data only marginal information is available. It is not clear which utterances are included, what criteria have been applied for interpretations and whether or not more than one interpreter assigned an interpretation, for instance.

MRE. In the fifth chapter, I will propose an alternative account that captures the corpus results reported in this chapter as well as other data that will be presented in Chapter 5. The implication of rejecting the IMH account is that if the experiment reveals differences in meaning between Dutch and English RIs, an alternative explanation must be given for this cross-linguistic difference. If the experiments do not suggest any difference, it must be explained why the corpus and experimental results are deviant.

3.4.3   Pilots

To test the children, we used a design to elicit verb forms: a *controlled elicitation task*. For an experiment, this design is fairly unconstrained, as will be explained later on. The major difference between the experiment and the corpus method is that in the experiment the meaning is controlled. The design we ultimately settled on followed a series of pilot studies. This section provides a summary of the most important conclusions from the pilot studies.

Assigning an interpretation to naturalistic corpus data is problematic. However, testing two- and three-year-olds to find out more about the semantics of the forms they use, is difficult for other reasons. During the pilots we observed (i) that a completion task may lead to imitation, (ii) that a comprehension task was often not correctly understood, (iii) that dynamic movies are preferred to static pictures for the present topic, and (iv) that a condition intended to depict modality requires more explanation than a condition expressing an ongoing activity. Most of these problems relate to the young age group we tried to test. In this section, I will discuss these points one by one.

In the pilots, we tested whether a completion task qualified as a pre-test. The aim of the pre-test was to pick out the children in the RI-period. We showed the children two contrasting pictures in two conditions, i.e. ongoing and modal, described the first picture and stimulated the children to complete the description of the second picture. For instance, in the ongoing condition we presented the children a picture with an eating man and a drinking man. The experimenter said: *Deze man eet en deze man …?* ('This man is eating and this man …?'). The sentence was uttered with a rising intonation so that the child was stimulated to fill in the event in the second

clause. The children we aimed to select, had to insert an infinitive in at least one of the items that were offered. As various children that used RIs when they talked spontaneously produced no RIs in this test (which may be an effect of the finite triggering sentence), we decided that this test was not an adequate pre-test to select subjects for the experiment.[34]

According to the first set-up, a production as well as a comprehension task was carried out to investigate the meaning properties of RIs. Even though I am still convinced that this combination of tests gives optimal results, we did not include comprehension in the final experiments. Subjects who were able to do the production task were often incapable of doing the comprehension task and there were few subjects that understood both tasks. Most misunderstandings were of the following type: the children that were asked to pick out the picture that matched a sentence (either ongoing or modal) selected the picture that simply matched the event. This means that the ongoing picture (showing the event) was the preferred picture, for both the ongoing as well as the modal sentence. The picture showing the modal event (hence, that did not show the event) did not match any sentence. Therefore, it was less often chosen. For instance, the triggering sentence in the modal condition was *Peter wil **eten*** 'Peter wants to **eat**'. The ongoing variant was *Peter **eet*** 'Peter **eats**'. In both cases, the children tended to select the picture that showed Peter eating and not the picture that showed Peter's intention or desire to eat. This tendency indicated that the children concentrated on the event and the picture that represented the event; the children neglected the ongoingness or modality that distinguished the two conditions from each other. More discussion on this issue can be found in § 3.4.13.

It turned out to be difficult to make young children understand the modal condition. Two things were helpful in this respect: the use of movies and the implementation of causality in the modal condition. Movies were preferred to pictures, as the children we examined tended to interpret describing pictures as a labelling task: they did not pay attention to the modal-ongoing contrast. Instead, they interpreted the modal action as a 'not-action' or a different action instead of a 'going to be action'. Thus,

---

[34] We decided not to use RIs to elicit child completions, because this lead to unnatural triggering sentences.

many responses were of the type: *walks - not walks/stands* and not from the type *walks - wants to walk*. Movies elicited less labelling utterances and relatively more multi-word sentences. To emphasise modality, we built in causality: the necessity to act was stressed because the agent had a reason to act. For instance, a dirty hungry dog had to wash himself because he was only allowed to enter the house (where the food was) if he was clean. In this case, the washing was motivated by the dog's filthiness and hunger. Another illustration of a compelling reason to act (more specifically to run away) was the approaching of a car while a boy played on the street. Although this strategy helped the children to understand the modality involved, it required insight in causal relations between events. This requirement excluded children of the youngest age group.

Our initial idea was to create a controlled procedure in which we would collect an equal number of responses for each subject as well as an equal number of responses in each condition. In the elicitation task that we finally carried out, we gave this kind of control up as it led to the exclusion of many informative responses. The movies represented activities in an ongoing and a modal condition. In the former condition, the action was shown. In the latter condition, the desire or necessity to perform the action was stressed through causality; the action itself was not shown. However, the children did not always talk about the specific actions that we primed on. When we were focussing on washing, the children talked about swimming, taking a bath, becoming clean or splashing. Therefore, the decision was made to allow a fairly broad interpretation of the actions depicted in the movies, and hence to include utterances about swimming, taking a bath, becoming clean and splashing as descriptions for the 'washing' movies. Furthermore, a movie contained often more than one action, especially the movies in the modal condition. For instance, in the modal 'running' movie, a moving car was to be seen. In the modal 'drinking' movie, a girl walks to her mother. When the children talked about what happened in the movies, they mentioned such side-activities as well. We decided that it was important to collect as many verb forms as possible that denoted an activity of which we were certain that, at speech time, it was either ongoing or modal, irrespective of whether it was the action we initially focussed on.

3.4.4   Pre-test

Before the experimental task started, each child was shown a big picture on a laptop computer with lots of things to talk about. In this picture, three subsequent actions or movements could be animated: a bird flew away, a ball rolled on the grass and someone was driving a car. Presenting this picture served several purposes. For us it served to elicit spontaneous speech from a child. On the basis of this sample, we decided whether or not the child was in the RI-phase. Children that only used finite sentences in the pre-test were not selected for the actual experiment. Apart from being representative of the children's linguistic behaviour, other advantages of this pre-test were that it made the children familiar with the laptop computer, the setting and the task.

3.4.5   Test items

We designed eight animated movies that concentrate on four actions. Each action was shown in a modal and a non-modal condition. In the modal condition, only the intention or wish to act was shown, whereas in the non-modal condition the ongoing action was shown. I will describe the test items below. The choice of verbs denoting particular actions was based on (i) frequent use of the verbs in Dutch corpus data, and (ii) appropriateness of the denoted actions for short movies. This lead to the selection of *wassen* (washing), *rennen* (running), *drinken* (drinking), and *bellen* (calling).

Table 3.5: Description of the test items, used in the controlled elicitation task to test the denotation of verb forms in Dutch and English child language

Action	Ongoing Condition	Modal Condition
Washing	A dirty pig, covered with mud, stands next to a bathtub that is filled with water and foam. The pig jumps in the tub. He is washing himself: he turns around several times, is splashing water and the dirt	A dirty dog, covered with mud, stands on a road that leads to a house, which can be seen on the background. A bathtub is in front of the house. The dog walks in

	gradually disappears.	the direction of the house.
Running	A girl is running from the left to the right side of the screen. It is raining.	A boy is playing on the street. He throws a ball up in the air several times. Suddenly a car is approaching him.
Drinking	A boy is sitting at a table. He picks up a glass filled with lemonade that stands in front of him. He brings the glass to his mouth and drinks until the glass is empty. This is shown by a gradually reducing amount of lemonade.	In a room there are a girl and a mother. The girl has an empty glass in her hands. She walks to her mother and lifts the empty glass.
Calling	A boy is sitting at a table. In front of him is a telephone. The boy picks up the phone and talks for a while, which is shown by his moving lips.	A little girl is standing next to a cupboard with a phone on it. She tries to reach the phone but is too small.

We made the movies simple in order to reduce the chance that the children misunderstood the purpose of the movie. Occasionally, however, the children drew their own conclusions from what they saw. An anecdote illustrates the unforeseen interpretations of young children: one boy insisted that getting clean did not help the dog to enter the house. When we asked him why, he answered that the dog first had to become smaller (because of the perspective the house in the background was much smaller than the dog in the foreground).

3.4.6 Protocol

The experiments were run with two experimenters. While one experimenter talked to the child and told the story of the movie, the other observed and took notes. In addition, the sessions were audio-taped, to be transcribed in CHAT-format (MacWhinney, 1995) afterwards and expanded with semantic information. We used the following protocol. First, the experimenter intro-

duced the characters. The human beings were given common names like Peter or Lisa. The movie was shown while the experimenter told the child what happened. Then, the movie was shown a second time and the child was asked some questions to test whether the movie was understood (Can you tell me what you see? What does Peter want? What is the doggy going to do? Etc.).

If so, the experimenter showed the movie again and asked the child to say what happened. The experimenter interfered as little as possible. Note that the story telling was especially important in the modal condition, as the intended meaning (i.e. wish or requirement) could not be derived from the movie alone. Below are two examples, *washing* in the modal and the non-modal condition respectively. The rest of the stories are in Appendix 3.2 (p. 243).

DIRTY DOG (modal)

This is the story of the dirty hungry dog that wants to go into the house to eat. He is so hungry and his food is in the house. But he is much too dirty to go into the house. Look how dirty he is! The doggy really has to wash himself. Look! Next to the house is a bathtub, where he can go and wash himself.

WASHING PIG (non-modal)

This is the story about the dirty pig. The pig is waiting to get in the bath. See, now he jumps into the bathtub. He is washing and washing. He washes himself until he is completely clean. You see?

3.4.7   Subjects

Table 3.6 gives information about the number of subjects that participated, their ages and MLU scores. All subjects produced RIs in the pre-test, and were selected on the basis of this pre-test. As our analyses concentrated on RIs, we excluded subjects that did produce RIs in the pre-test but used no RIs at all in the test itself. The subjects in Table 3.6 used at least one RI. Subjects that did not seem to respond on modal contexts were excluded.

Table 3.6: Numbers, ages, MLU of the Dutch and English-speaking subjects in the experiment

	Dutch subjects	English subjects[35,36]
Number	26	29
Age range	1;11 - 3;10	1;11 - 3;6
Mean age	2;10	2;7
MLU range	1.57 - 4.9	2 - 5.62
Mean MLU	2.63	3.44

Given that (Dutch) children start to produce RIs around the age of two (see Chapter 5 of this thesis), the subjects in the experiment are fairly old: the mean age of the Dutch subjects is 2 years and 10 months.

3.4.8 Analysed data

To be included in the analysis, RIs had to meet the following criteria: (i) they had to be part of a multi-word utterance in order to exclude nominative infinitives that are used to label actions, and (ii) they were not parasitic or elliptic. Thus, infinitives that were answers to a question asked by the experimenter, such as in the examples (17) and (18) in § 3.3.1, were excluded. In English, the elicited utterance must have a third person singular subject, as this requires the only finite verb form that is morphologically distinct from the bare verb form with either the suffix [-s] in the present tense or [-ed] for past tense. In the experiment, this criterion was easily met, as the characters in the movie were all third person singular subjects. A total number of 198 and 158 RIs have been excluded in the English and Dutch results. Additionally, 91 root participles (i.e. non-finite clauses like RIs,

---

[35] The subjects were English-speaking American toddlers, predominantly monolinguals. The bi-lingual subjects that were included did not show any deficiencies in their English. Before the children were tested, we asked the children's caregivers to fill out a questionnaire about the language situation at home (languages of the caregivers, languages spoken at home, etc.).

[36] In Blom & Krikhaar (2002), there are 30 English subjects. However, I later excluded one subject when it became clear on closer inspection that it used no modal utterances at all.

containing a past participle instead of an RI) have been left out from the English data. Very few finite sentences, i.e. less than 10, were excluded.

3.4.9 Two different questions

The experimental data can be approached in two ways, dependent on the question that is asked. Both meaning and form can be taken as constants or variables. When meaning is taken as a constant and form is the variable, the question follows: Given a meaning, what is the probability that a child uses an RI? I will call this the *semantic approach*, for this approach isolates a meaning, say modal meanings, and the probability is calculated that a particular form (e.g. RIs, SFs or PVs) is used to denote this meaning. This is turned around in the second approach, i.e. the *syntactic approach*. In the syntactic approach, forms, say RIs, are isolated and the probability is calculated that this form is used with a certain denotation (e.g. modal or ongoing).

The syntactic approach is applied to the corpus results. Thus, in order to compare the experimental results with the previously reported corpus results, the syntactic approach is preferred, even though it seems counter-intuitive at first sight. The semantic approach has other advantages, however, and is preferred for reasons other than just being comparable to the corpus results. The first advantage of the semantic approach is that modality within RIs is calculated as a proportion of the total amount of modal utterances that a child produces (the same applies to the proportion of ongoing RIs). Thus, the percentages of modal RIs and ongoing RIs are independently calculated. This analysis controls for the effect that either a high or low percentage of modal RIs is the effect of much or little modal talk in general (this method controls for the situation factor discussed in (§ 3.4.1). Secondly, the first analysis provides insight in the full range of verb forms that a child has available. It relates the properties of RIs to other forms and hence, gives additional valuable information. Therefore, I will reconsider the corpus data and apply the semantic approach. The results based on this method will be presented in the fifth chapter (§ 5.3).

The question may arise why the semantic approach is not applied to the corpus data.[37] As far as I may judge, the main reason for application of the syntactic approach is a focus on RIs as a construction. Researchers that examined semantic properties of RIs were interested in discovering and defining properties of this specific form that is frequent in child language but rare in adult speech. Given the syntactic point of view, the question 'What is the meaning of an RI?' is more obvious than the question 'What is the proportion of RIs in the set of utterances that denote modal events?'.

3.4.10 The semantic approach to the results

The semantic approach takes the whole range of verb forms that children use into account. In the experiment, the Dutch subjects used four different forms: RIs, simple finite sentences (SFs), periphrastic verbs (PVs) and prepositional infinitival complements (PICs):

(40) a. Het meisje rennen   RI
          the girl run-*inf*
      b. Het meisje rent   SF
          the girl run-*fin*
          'The girl is running'
      c. Het meisje moet/wil/gaat rennen   PV
          the girl must/wants to/goes run-*inf*
          'The girl must/wants to/is going to run'
      d. Het meisje is aan het rennen   PIC
          the girl is on the run-*inf*
          'The girl is running'

The English subjects used five different forms: RIs, SFs, PVs and two forms that contain a present participle, namely 'normal' finite sentences (abbreviated as FPs 'finite participles') and root participles (RPs). Examples of the five forms are given in (41):

---

[37] This question mainly concerns Hoekstra & Hyams' study (and the studies that Hoekstra & Hyams refer to), because they make a strong claim about the meaning of RIs but only use the contrast between RIs and SFs to show that RIs are modal. RIs are not compared to PVs. There is no information about the proportion of RIs in the sets of modal and ongoing utterances, and, hence no evidence showing that the modality in RIs is disproportional.

(41)   a.   The girl run   RI
       b.   The girl runs   SF
       c.   The girl has to/wants to/is going to run   PV
       d.   The girl is running   FP
       e.   The girl running   RP

The Tables 3.7 and 3.8 show the distribution of these forms over the modal and ongoing condition in the Dutch and English data:

Table 3.7: Distribution of forms in the modal and ongoing condition, Dutch subjects (N = 26)

Form	N	Modal condition (N = 298)		Ongoing condition (N = 562)	
		N	%	N	%
RI	149	101	34	48	9
SF	487	42	14	445	79
PV	216	155	52	61	11
PIC	8	0	0	8	1

Table 3.8: Distribution of forms in the modal and ongoing condition, English subjects (N = 29)

Form	N	Modal condition (N = 243)		Ongoing condition (N = 440)	
		N	%	N	%
RI	162	71	29	91	21
SF	79	14	6	65	15
PV	153	134	55	19	4
FP	201	13	5	188	43
RP	87	11	5	76	17

What do these distributions show? Let us start with the forms that show a rather clear usage. In the Dutch results, these are SFs and PVs. The first are very likely to appear in the ongoing condition, whereas the latter prefer the modal condition. In the English results, FPs and PVs show a clear usage. The first are likely to appear in the ongoing condition, whereas PVs prefer the modal condition (like in Dutch). PICs in early Dutch child language are marginal, SFs are less frequently used by English children than by Dutch children and the English children use a variant of RIs, namely RPs.[38] A closer look at the class of PVs in Dutch indicates that PVs with a modal auxiliary like *moet* 'must' or *wil* 'want' have to be distinguished from PVs that contain the auxiliary *gaat* 'goes'. In adult Dutch, *gaat* is used to denote (near) future but the children use this auxiliary in a different way. Recall that this was mentioned before, in § 3.3.6, with regard to the interpretation of the PVs of one of the six Dutch-speaking children that I examined, Peter. In the experiment, all PVs used in the ongoing condition were instances of 'go'. In the literature on Dutch child language, this non-adultlike use of *gaat* is often reported (Schaerlakens & Gillis, 1987; Verhulst-Schlichting, 1985; Jordens, 1990; Evers & Van Kampen, 1995; Hollebrandse & Roeper, 1996; Van Kampen, 1997; Zuckerman, 2001). In Chapter 5, I will discuss this ongoing, or present tense, use of *gaat* as an effect of development.

Compared to these clear forms, the behaviour of RIs looks diffuse, especially in the English results. Recall that the predictions generated by the hypothesis with which I started this section, i.e. the IMH, were that Dutch RIs are restricted to a modal meaning whereas English RIs are unrestricted with regard to modality. Translating this into predictions that are consistent with the semantic approach, I arrive at the following three predictions: it is expected that (i) in Dutch, the likelihood that RIs appear in the modal condition is larger than in the ongoing condition, (ii) in English, there is no difference between the two conditions, and (iii) it is more likely in Dutch that RIs appear in the modal condition than in English. At first blush, it seems more likely that Dutch RIs appear in the modal condition than in the ongoing condition, because 34 % of the modal forms are RIs, whereas only

[38] Note that the comparison between the different verb forms serves a methodological purpose. The clear differences between the forms indicate that the experimental results are reliable. If the subjects did not understand the experimental conditions (for instance the modal condition), less clear-cut differences between the forms would have been expected.

9 % of the ongoing forms are RIs. In English, there is less of a difference between the two conditions: 29 % of the modal forms are RIs, whereas 21 % of the ongoing forms are RIs. With regard to the cross-linguistic difference, the results suggest a difference between the probability that RIs appear in the modal condition, although the difference is small: 34 % of the modal forms in the Dutch data are RIs, whereas 29 % of the modal forms in the English data are RIs.

Are the differences and similarities as they present themselves in the percentages statistically reliable? Table 3.9 gives an overview over the probabilities that the various H0's (which follow from the three predictions that I formulated) are confirmed. The permutation test is used to calculate the p-values:[39]

Table 3.9: Probabilities that there is no difference between the Dutch and English results in the modal and ongoing condition and that there is no difference between the modal and ongoing condition in Dutch and English

H0	p-value
There is no difference between the conditions in Dutch	0.001
There is no difference between the conditions in English	0.03
There is no difference between Dutch and English in the modal condition	0.001

Given the probabilities in Table 3.9, it can be concluded that statistical analyses confirm the first impressions. In Dutch, there is a difference between the two conditions. As the p-value is small, i.e. $p = 0.001 < p = 0.05$, it can be concluded that it is more likely that an RI appears in the modal condition than in the ongoing condition. The same applies to the English results: there is a difference between the two conditions ($p = 0.03 < p = 0.05$). The larger probability that RIs show up in the modal condition in both languages is the result of the children using relatively more forms (other than RIs) in the ongoing condition than in the modal condition. In

---

[39] Proportions RIs in the modal and ongoing condition are calculated for each subject. I have taken the average of these proportions and difference between languages/conditions.

the Dutch data, these productive ongoing forms are predominantly SFs. In the English data, progressives, i.e. FPs and RPs, are responsible. When the two languages are compared, it turns out that there is a difference: in Dutch, it is more likely that an RI occurs in the modal condition than in English ($p = 0.001 < p = 0.05$).

3.4.11 The syntactic approach to the results

In this section, RIs are isolated (as is done with the corpus results) and the number of modal RIs is calculated as a proportion of the total number of interpretable RIs. This is done for both languages, Dutch as well as English. The results are plotted in a box-plot. The percentiles provide information about the distribution of the results. When the box is small, the results are very similar and the standard deviation will be low. When the box is large, the children's responses diverge clearly and the standard deviation will be high. The box-plot in Figure 3.5 depicts the percentiles 0.05, 0.25, 0.5, 0.75 and 0.95. The percentiles 0.05 and 0.95 are the two borders of the 90 % interval; in the figure, this interval is marked with two horizontal stripes. The percentiles 0.25 and 0.75 are the two borders of the 50 % interval; this interval is marked with the two small open circles. The percentile 0.5 is the exact middle of the distribution, or the median; this is marked with the black square. The big open circle is the average proportion of modally used RIs.
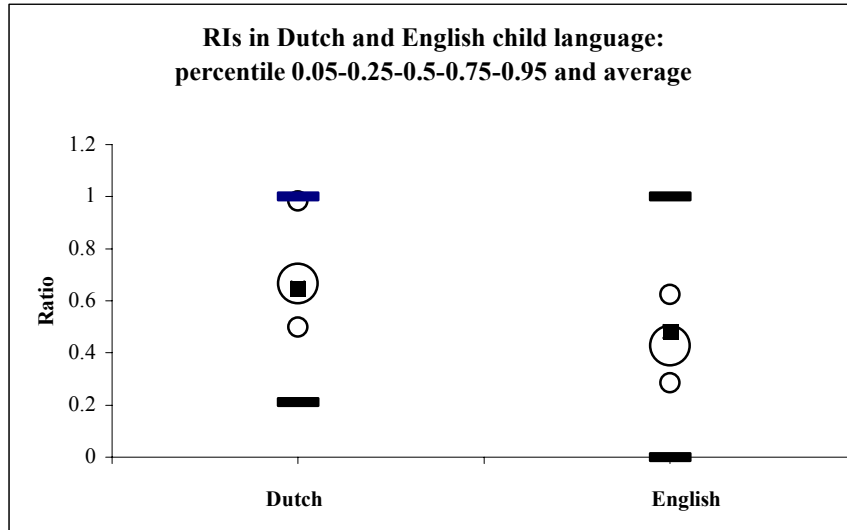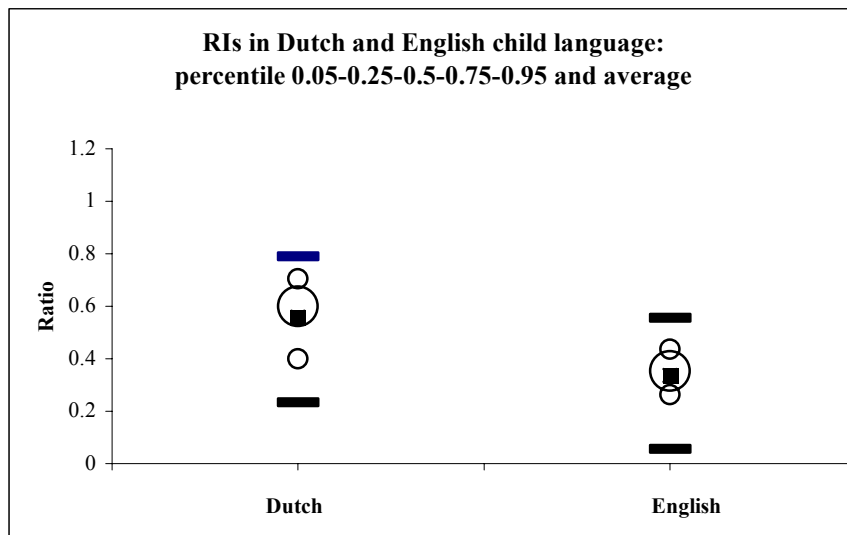
Figure 3.5: Use of RIs in the modal condition, results of Dutch and English-speaking children that used at least 1 RI, percentile 0.05-0.25-0.5-0.75-0.95 and average

—	0.05 - 0.95 percentile
○	0.25 - 0.75 percentile
■	median
○	average
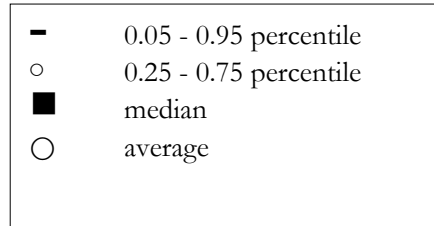
Table 3.10 gives the total number of interpretable RIs, the proportion of RIs used in the modal condition (with the number of modally used RIs in parenthesis), the average modal proportion over the total sample of children and the standard deviation.

Table 3.10: Use of RIs in modal condition, results of Dutch and English-speaking children that produced at least 1 RI, number of interpretable RIs, percentage of modally used RIs (number), and standard deviation

	N_{RI}	$N_{SUBJECTS}$	% $_{MODAL (N)}$	SD
Dutch	149	26	68 % (101)	29
English	162	29	44 % (71)	31

With the aid of the permutation test, I calculated (i) whether the Dutch children show a preference with regard to the use of RIs in the modal

condition, (ii) whether the English children prefer to use RIs in one of the two conditions, and (iii) whether there is a difference between the Dutch and the English results. The Dutch children show a preference for using RIs to denote modal events. When a chance distribution is simulated and the observations are shuffled per subject, it turns out to be unlikely that the observed distribution and the chance distribution are drawn from a single distribution ($p = 0.013$). The English subjects do not show a preference ($p = 0.62$). The English p-value is not extremely high (smaller than 0.95). Hence, I conclude that there is no support for the claim that English RIs show a preference to denote ongoing events with RIs. If we compare the Dutch and English results and estimate the probability that the distributions in the two languages are drawn from one single distribution, a p-value of 0.024 is obtained, which is below the criterion value of 0.05. This indicates that there is difference between the usages of RIs in two languages: RIs in child Dutch are more often modal than RIs in child English.

If, in the experiment, a child used only one RI and this RI was used in the modal condition, 100 % of the RIs of this particular child were classified as modal. Many of such subjects may affect the results. Therefore, I carried out an additional post hoc selection that includes only children that used 5 or more RIs and 5 or more finite sentences, i.e. FINs, (to be sure that the children are also able to use forms other than infinitives or bare stems).[40] Although the sample of subjects after this selection is much smaller, the data may be more reliable than the data just described: the effect of children that use only one RI (and this one RIs is either 100 % modal or 100 % non-modal/ongoing) no longer occurs.

Considering Figure 3.6 and Table 3.12, it seems that the distributions in both languages are indeed affected by the exclusion of children who used very few RIs/FINs. The 90 % and the 50 % intervals are smaller than before and the standard deviation is lower. In general, the results are less spread out than before.

---

[40] Although this number remains rather arbitrary, is gives some idea of the effect.

Table 3.11: Numbers, ages, MLU of Dutch and English-speaking subjects in the experiment that produced 5 or more RIs and 5 or more FINs

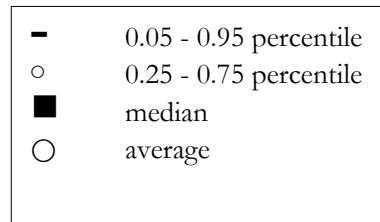|	Dutch subjects	English subjects
Number	9	12
Age range	1;11 - 3;10	2;1 - 3;5
Mean age	2;09	2;06
MLU range	1.57 - 4.25	2.74 - 5.62
Mean MLU	2.68	4.08

Figure 3.6: Use of RIs in modal condition, results of Dutch and English-speaking children that used at least 5 RIs and 5 FINs, percentile 0.05-0.25-0.5-0.75-0.95 and average

Table 3.12: Use of RIs in the modal condition, results of Dutch and English-speaking children that produced at least 5 RIs and 5 FINs, number of interpretable RIs, percentage of RIs used in modal condition (number), and standard deviation

	N_{RI}	$N_{SUBJECTS}$	% MODAL (N)	SD
Dutch	87	9	61 % (53)	22
English	90	12	36 % (32)	17

Statistical tests confirm the first informal observations in so far as that the results for each language have changed in comparison to the first analysis that included all children. The difference between the two languages remains statistically significant (p = 0.006). What is the difference between the two analyses? For the Dutch subjects, the probability that the observed distribution and a chance distribution are drawn from the same distribution is large (p = 0.22). They do not show a preference for the modal condition anymore. The p-value is smaller than 0.95, which suggests that there is no support for the claim the Dutch subjects show a preference for the ongoing condition. For the English subjects, the probability that the children show a modal preference is small: p = 0.005. This p-value is smaller than 0.05. Hence, I conclude that these data show a preference to use RIs to denote ongoing events.

In sum, according to both tests, there is a difference between the two languages:

Table 3.13: Probabilities that there is no difference between the use of Dutch and English RIs in the modal condition (i.e. null hypothesis), for all children (that used at least one RI) and for the productive children (i.e. the children that used at least 5 RIs and 5 FINs).

H0	P-value All children	P-value Productive children
No difference between the use of Dutch and English RIs in the modal condition	p = 0.024	p = 0.006

The outcome of the two analyses is similar: Dutch RIs are more often modal than English RIs. With regard to the question whether Dutch RIs show a preference for modal events or English RIs show a preference for ongoing events, the results are ambiguous. When the larger sample of children is taken and all subjects are included, there is a modal preference in the Dutch RIs and no preference for either modal or ongoing events in the English RIs. When results of a small sample of productive children are analysed, there is an ongoing preference in the English RIs and no modal or ongoing preference in the Dutch RIs.

3.4.12 Discussion

I approached the experimental data in two ways. The question that underlies the semantic approach is which form(s) the children use when they talk about either a modal or an ongoing event. More specifically, how probable is it that Dutch and English-speaking children in the RI-age use an RI in the modal condition given a set of various different verb forms? It turned out that in both languages, it was more likely that RIs appeared in the modal condition than in the ongoing condition. A comparison between the two languages showed that the probability that Dutch children used RIs in the modal condition was significantly larger than that the English children used RIs in this condition. In the syntactic approach, I collapsed the RIs produced in the two conditions in order to determine the proportion of modal versus ongoing RIs *within the set of RIs* in the two languages. Like the

results from the semantic approach, the results from the syntactic approach confirm cross-linguistic differences. On the basis of the experimental results, it is unclear if the difference should be formulated as 'Dutch RIs show a modal preference and English RIs do not' or 'English RIs show an ongoing preference and Dutch RIs do not'. The direction of the difference is similar in both formulations, however.

According to Hoekstra & Hyams's (1998) theory, the IMH, the proportion of modal RIs in Dutch child language is insensitive to the frequencies of other forms that children use besides RIs. In the semantic approach, the proportions of RIs are relative to the other forms that the children used in the experiment, however. Because of this discrepancy, I am reluctant to draw any conclusions on the basis of the outcome of the semantic analysis with regard to the study that motivated me to do the experiment in the first place. Nevertheless, an observation that is in line with the IMH is that there is a difference in meaning between Dutch and English RIs, more specifically that Dutch RIs are more frequently modal than English RIs. Taking a look at the data in the Tables 3.7 and 3.8, I conclude that the experimental results do not seem to support the IMH unequivocally, in spite of this cross-linguistic difference. The tests that I applied tested differences, which I translated as preferences. The IMH, however, makes a prediction with a more absolute character for Dutch: it is expected that the Dutch children use RIs in the modal condition, but do not use RIs in the ongoing condition (§ 3.3.7). This prediction is not borne out as the Dutch subjects use RIs in both conditions: in the modal condition 34 % of the responses are RIs, in the ongoing condition 9 % are RIs.

3.4.13 Comparison with other studies

How do the cross-linguistic experimental results relate to other and earlier findings? As far as I know, there is only one other experimental study in which the use of verb forms in child Dutch and English has been compared. Schönenberger, Pierce, Wexler & Wijnen (1995) report data from a sentence-picture-matching task carried out with Dutch and English-speaking children. Despite the fact that the design of the experiment causes a non-modal bias - I will discuss this bias later on in this section - and that the number of English children that is tested is too low (N = 5) to draw any far-reaching conclusions, it is interesting to take a look at their results.

The aim of the Schönenberger et al. study was to investigate the interpretation that children assign to RIs (modal vs. ongoing). In the task, the child functions as an intermediary between puppet Kermit the Frog and the experimenter. The experimenter tells a story involving two pictures: one depicting an ongoing event, the other depicting a modal event. A puppet (Kermit the Frog) is listening to the story as well. After the story has finished, the experimenter asks Kermit which picture he likes best. However, Kermit cannot point. Therefore, he has to explain with a descriptive sentence. The experimenter does not understand Kermit and so the child, who can understand him, has to make Kermit's choice clear to the experimenter by pointing at the picture that fits Kermit's description. In the experiment, Kermit uses three kinds of sentences: either with a finite main verb, a modal auxiliary or an RI. In the English version, the present progressive is used instead of a simple finite main verb as this sounds much more natural. In this way, this experiment elicited the interpretation (i.e. modal or ongoing) of three kinds of constructions. Table 3.14 below contains the percentages of selection of the picture that depicted the ongoing event for three different sentences:[41]

Table 3.14: Results from Dutch-English experiment on the interpretation of RIs, Schönenberger et al. (1995), the percentages of selections of pictures that depicted ongoing events for respectively finite verbs, non-finite verbs (RIs) and modal verbs

	% FIN.VERB	% NON-FIN.VERB (RI)	% MODAL VERB
Dutch	91	62	43
English	95	95	40

[41] Eight English-speaking children joined the experiment. However, only results of five children could be retained. The age-range holds for all eight children though.

	Number of subjects	Age range
Dutch	18	1; 11 - 3; 4
English	5	2; 8 - 3; 5

These results show the ongoing bias that we also noticed in our pilots when we asked the children to point to the picture that matched a sentence (§ 3.4.3). Various children tended to ignore the modality and focused on the action. This action is expressed in all stimuli, but it is only shown in the picture that depicts the ongoing action. The picture that depicts the modal action does not show the action itself. Hence, for all stimuli the ongoing picture is the best match. The results in Table 3.14 illustrate this effect: even for the stimulus sentence with a modal verb (e.g. *Koekiemonster wil de koekjes hebben/Cookie Monster wants to have the cookies*), both Dutch and English-speaking children point to the picture that shows the ongoing event (thus, to the picture in which Cookie Monster has cookies) in approximately 40 % of the cases. Despite this non-modal bias, the results in Table 3.14 do show an effect. Finite verbs (simple present tense in Dutch and present progressive in English) and modal verbs are interpreted similarly in Dutch and English: finite verbs are nearly always ongoing, while for modal verbs, the modal ongoing picture is chosen approximately 60 % of the time. Interestingly, non-finite verbs (i.e. the infinitive in Dutch and bare stem in English) yield a different interpretation in the two languages. English children tend to point more often to the ongoing picture if the stimulus sentence is an RI than the Dutch children do (95 % vs. 62 %).

In Table 3.15, an overview of the results from various studies is given. I do not aim to compare the different studies in detail, as there are too many differences between these studies. In each study a different methodology is applied, and, as pointed out before, each methodology has its own problems that may affect the results.[42] The size of these effects is unknown, and, therefore, a detailed comparison is meaningless. In spite of these methodological differences and problems, across studies a similar difference is measured. In all studies, it is found that Dutch RIs are less often ongoing/more often modal than English RIs. In the overview below, I informally corrected Schönenberger et al.'s results and subtracted 40 % (which may approach the size of the non-modal bias, given the children's responses on stimuli that contained a modal auxiliary):

---

[42] At least three factors may influence the interpretation of corpus RIs: the information factor, subjectivity factor and situation factor (§ 3.4.1). The comprehension experiment suffers from a bias for the ongoing condition.

Table 3.15: Modal RIs in Dutch and English child language, corpus data (present study for Dutch data; the English percentage is the average of the results from Deen, 1997 and Madsen & Gilkerson, 1999), experimental results from a production task (present study) and experimental results from a comprehension task (Schönenberger et al., 1995)

	% ONGOING corpus study	% ONGOING experimental study (production)	% ONGOING experimental study (comprehension)
Dutch	26	32 - 39	22
English	89	45 - 36	45

Besides the cross-linguistic difference, there is another observation that catches the eye: the asymmetry between the two experimental studies, on the one hand, and the corpus study, on the other hand, with regard to the preponderance of ongoing RIs in the English data. In the corpus data, the percentage of RIs with an ongoing reading is strikingly high. A similar high proportion of ongoing RIs, also based on analysis of corpus data, is reported by Harris & Wexler (1996). Harris & Wexler interpreted the bare stems of English children and compared their temporal reference to forms ending with an [-s] suffix. They found a clearly dominant present tense ('ongoing') use of the bare stems of 82 %.

In § 3.5.1, I will give an answer to the question why specifically English *corpus* RIs are more frequently ongoing than the experimental results. In § 3.5.2, an alternative explanation for differences between the interpretation of RIs in Dutch and English child language will be given that does not restrict RIs in Dutch child language to a modal meaning (and that, therefore, accounts more successfully for the results than the IMH in this respect).

3.4.14 Summary

In this section, I discussed experimental results of a cross-linguistic experiment carried out with Dutch and English subjects in the RI-age. The motivation for this experiment were the predictions that follow from a theory developed by Hoekstra & Hyams' (1998), the IMH, and naturalistic

child data Hoekstra & Hyams discuss in order to support the IMH. Basically, the idea is that RIs in Dutch child language are modal because they contain a morphologically marked infinitival verb form, whereas RIs in English child language (lacking infinitival morphology) have a free modal usage. I argued that the empirical foundation of this claim, which is provided by a comparison between data from various different studies, is unstable and that it is therefore desirable to conduct an experiment. In an attempt to design a suitable experiment, I encountered a number of difficulties, which led to a fairly uncontrolled final experimental design. The most important advantages of the experimental data are that the denotation of a child utterance is more controlled and that a valid comparison can be made between Dutch and English. The results confirm a difference in meaning between RIs in Dutch and English child language: Dutch RIs are more often modal than English RIs. This observation is compatible with the IMH. Given the IMH, it is not expected that Dutch children use RIs for ongoing events. The Dutch subjects, however, used ongoing RIs and hence, the prediction from the IMH is not borne out (as concluded before with regard to the Dutch corpus data).

3.5   The ongoingness of English RIs

In the remainder of this chapter, Dutch and English RIs will be discussed in greater detail in order to explain the large variation between the observations reported in the studies that concentrate on this topic and to gain more insight into cross-linguistic differences and similarities. I first address differences between corpus studies and experimental studies. In the following section, I will argue that differences between the meaning of RIs in Dutch and English child language are expected, given certain morphological and syntactic properties of the two languages.

3.6   A methodological artefact

Why are English corpus RIs so often non-modal? The answer to this question follows from the fact that the English corpus RIs are restricted to utterances with third person singular subjects while Dutch corpus RIs are

utterances with all kind of subjects. Exclusion of utterances with first and second person singular subjects leads to a considerable decrease of, particularly, modal utterances. What is the reason for the asymmetry between the Dutch and English corpus data? Before I turn to a reanalysis of the Dutch data, I will explain why it is perfectly well understandable that the set of English RIs contrasts to the set of Dutch RIs.

English has a very poor inflectional paradigm. In the present tense there is only one suffix, which is the third person singular [*-s*] as in *walks* or *drinks*. All other forms in the paradigm are bare. Thus, except for utterances with third person singular subjects, RIs cannot be distinguished from finite utterances:

(42) a.   Peter **get** ball   RI
      b.   Peter **gets** ball   FIN
      c.   I/you/we/they **get** ball   ??

In Dutch, the inflectional paradigm is not only richer, but the position of finite verb and infinitive is distinct as well. Recall that finite forms are placed in sentence-initial position while infinitives are placed sentence-finally. Even if morphology is not sufficient (as in the plural), syntax provides a cue to decide whether an utterance is finite or non-finite. Present tense plural is marked by a morpheme that is equivalent to the infinitival morpheme: [*-en*]. However, if present tense plural is used, the verb is moved to sentence-initial position and precedes the object, while in RIs the verb follows the object:

(43) a.   Peter bal **pakken**   **RI**
            Peter ball get-*inf*
      b.   Peter/jij **pakt** bal   FIN
            Peter/you get-*fin-2/3sg* ball
      c.   Ik **pak** de bal   FIN
            I get-*fin-1sg* the ball
      d.   Wij/zij **pakken** de bal   FIN
            we/they get-*fin-pl* the ball

There seems to be no methodological reason to exclude RIs with specific subjects from the Dutch dataset, as long as it is unambiguously clear that the utterance is a RI and no finite utterance. And, indeed RIs with different

subjects are included, in all the Dutch data reported so far while in the English samples, only RIs with third person singular subjects are analysed. As soon as the interpretation of RIs is involved, and especially a modal interpretation, subjects in RIs become important however. The modality children use in RIs is strongly connected to discourse subjects, that is, to first and second person subjects or 'I' and 'you '. The specific interpretation that have been assigned to the RIs in my database, showed that the children used modal RIs nearly always to express desires/intentions or commands: they expressed either *I want + inf* or *You must + inf.* Thus, the modal RIs were either of the type 'dynamic necessity' and expressed a wish of the child, i.e. the speaker, ('volition') or were of the type 'deontic necessity' and expressed a command to an addressee. In addition, the children used prospective RIs for intentions (see § 3.3.1 for more information about the interpretations). I grouped the volitional and intentional RIs under the '*I want to + inf* -RIs'. These were often difficult to distinguish from each other, and for the present purpose, the distinction is irrelevant since the children connect them both to first person singular subjects. To conclude, the exclusion of first and second person subjects may lead to a dramatic decrease in the proportion of modal RIs.

In an attempt to give an idea of the size of this effect, I will re-analyse the Dutch corpus data presented earlier this chapter (section 3.3). My expectation is that when the Dutch RIs are made more comparable to the English RIs, the two languages do no longer exhibit such differences as suggested by the previously reported corpus data.[43] Moreover, the corpus data will be more like the experimental data, as the experimental data do not suffer from the methodological artefact: all characters are third person singular subjects. I want to emphasise that this reanalysis serves as an illustration. A re-analysis of the Dutch corpus data cannot provide a valid estimation of the actual size of the effect in the English corpus data.

---

[43] In addition to my attempt to make the Dutch and English RIs more comparable by re-analysing the Dutch data, further research has to point out what happens if the English data are re-analysed and modal use within all utterances with bare stems is calculated.

3.6.1   The corpus results re-analysed

To make the Dutch data comparable to the English data, the set of Dutch RIs must be narrowed down to RIs with third person singular subjects only. To facilitate this selection, subject use in RIs has been coded. In Appendix 3.1 (p. 240), some examples of these codes can be found. Null subjects whose interpretation was unclear, are excluded from the analysis. Table 3.16 gives the 'old' and 'new', i.e. re-analysed, data:

Table 3.16: Modal use of RIs, reanalysed Dutch corpus data on the basis of exclusion of first and second person subjects, data from all six children

	'Old data' Incl. first and second person subjects			'New' data Excl. first and second person subjects		
	N_{RI}	N_{MOD}	$\%_{MOD}$	N_{RI}	N_{MOD}	$\%_{MOD}$
Abel	120	91	76	16	7	44
Daan	157	115	73	27	13	48
Josse	206	150	73	26	18	69
Laura	314	200	64	82	28	34
Matthijs	254	199	78	76	34	45
Peter	197	157	80	93	67	72
Sum	*1248*	*912*	*-*	*320*	*167*	*-*
Average	*-*	*-*	*74*	*-*	*-*	*52*
SD	*-*	*-*	*6*	*-*	*-*	*15*

Re-analysis shows that the number of RIs dramatically decreases when first and second person subjects are excluded: the total number of RIs goes down from 1248 to 320. And, more importantly for our purposes: for four out of six children (Abel, Daan, Laura and Matthijs), there is no predominant modal use anymore. Their modal ratio is below the 50% level. Two children still show a preference for modal use. However, for Josse there are relatively few data available. The average modal ratio for the six

children goes down from 74 % to 52 %. Peter's case is peculiar in another way. He very often tends to use a proper name where adults would use discourse pronouns. In (44) two examples are given where an adult would use *I* and *you*:

(44) a   **Peter** pakken   Peter 1;11.13
           Peter get-*inf*
      b.   **mamma** doen   Peter 1;11.13
           mama do-*inf*

Table 3.17 illustrates how often Peter uses a proper name, his own name, instead of the pronoun *I*. The other children hardly ever do this, but Peter does so in 44 out of 93 RIs. The first column contains the number of interpretable RIs.

Table 3.17: Use of proper name instead of first person singular subjects in RIs, data from all six children

	N_{RI}	RIs with proper name instead of first pers. sg. pronoun
Abel	16	0
Daan	27	0
Josse	26	4
Laura	82	0
Matthijs	76	4
Peter	93	44

If Peter's data are modified and the 44 cases where a proper name is used instead of a first person singular pronoun are excluded, Peter's percentage of modal RIs goes down to 62 % (32/52). The average of all six children goes down from 52 % to 50 % modal use of RIs (SD = 13). Presumably, this percentage decreases even more if Peter's frequent use of *mama* 'mama' instead of second person singular *you* is excluded (there are 29 examples). However, due to the fact that corpus data do not provide any insight into the speech situation, one can never be sure if the proper name is really used

instead of *you*. Since most often a third adult is present as well, Peter may also be speaking to this adult and not directly addressing his mother. Therefore, I will not make this second modification. I assume that the first modification illustrated the described effect already very clearly.

3.6.2   Conclusion

In this section, I argued that the methodology used in corpus studies on RIs in English child language distorts the comparison between the interpretations assigned to Dutch and English (corpus) RIs. RIs in the studies on English RIs do not include RIs with first and second person singular subjects. This selection is not made in the studies on RIs in Dutch child language. The modality that children use in RIs, however, is closely related to first and second person singular subjects. Therefore, it is expected that the English RIs are less modal than the Dutch RIs. A modification of the Dutch results confirmed this hypothesis. The exclusion of first and second person singular subjects in the Dutch corpus data had a significant effect: the average modal use of RIs went down from 74 % to exactly 50 %. Recall that this problem occurred in corpus studies, but not in the experiments: all RIs in the Dutch as well as English experimental results have third person singular subjects. This explains why a comparison between the meaning of Dutch and English corpus RIs (as made by Hoekstra & Hyams, 1998) shows a considerable larger difference than the experimental results (Schönenberger et al., 1995; results presented in this chapter).

3.7   The Heterogeneous Set Effect

In Blom & Krikhaar (2002), it was argued that the *Fuzzy Set Effect* is responsible for differences in meaning between Dutch and English RIs. Our proposal was that the set of RIs in English is heterogeneous and contains different forms, as opposed to the set of Dutch RIs. More specifically, English RIs are a collection of 'real' untensed RIs in the sense of Wexler (1994) and Harris & Wexler (1996) and of forms with randomly dropped

inflection.[44] By implication, English RIs are relatively frequently ongoing: apart from the expected proportion of ongoing RIs in a prototypical RI-language such as Dutch or German, English RIs contain finite forms that have an ongoing or present tense denotation. In the present study, I use the more appropriate name *Heterogeneous Set Effect* (HSE) for this phenomenon. 'Fuzzy' relates to the border of a set whereas 'heterogeneous' is about the content of a set.[45] As the effect I am aiming at is dependent on the content and not the borders of the set of RIs, the term 'heterogeneous' is more adequate.

Why are English RIs a heterogeneous set of utterances? In brief, if English-speaking children use a finite sentence but, for some reason, drop inflection, an RI remains:

(45)   Peter catch-es the ball

In Dutch, this does not happen. A finite sentence with dropped inflection would look like (46):

(46)   Peter vang-t de bal

There are two reasons to exclude (46) from the set of RIs. First, the verb is placed in front of the object. Second, the verb does not have infinitival morphology. The RI-variant of (46) would look like (47):

(47)   Peter de bal vangen

Unlike Dutch, English is not a Verb Second language. Thus, English finite and non-finite verbs are placed in the same position. Morphologically, the verb form that is used in RIs, a bare stem, is hardly distinguishable from finite verbs (only in third person singular contexts, see previous section). These two properties of the verb in English RIs, make it impossible to

---

[44] This is what Harris & Wexler (1996) call 'Hypothesis I'; a hypothesis stating that children randomly use the [-s] morpheme: "…inflectional marking is optionally added or deleted, subject to interference from processing-load demands." (Harris & Wexler 1996:9). Such a proposal would be in the spirit of Bloom (1990).

[45] Frank Wijnen pointed this out to me.

distinguish between an RI and an utterance that contains a verb form with dropped inflection.

3.7.1   Support from Dutch for the HSE

Do children drop inflection randomly? This question cannot be answered for English, but it can be answered for Dutch. It turns out that Dutch children do use forms that are moved and that are syntactically marked for tense, but that lack inflection. These forms can count as 'stripped' forms when agreement inflection is dropped. Note that they can also be first person singular forms, since in the Dutch agreement paradigm first person singular is marked with a bare stem. For the present purpose, this distinction is not relevant: it is relevant, however, that both the dropped inflection form as well as the overused first person singular forms are finite. I will refer to the bare forms that appear in inappropriate contexts as "overused bare stems". An exhaustive list of these items is given in Appendix 3.3, p. 245. I have excluded one-word utterances from this list and bare stems that are the result of dropped infinitival morphology. On the basis of relative placement of object/particle and verb, I decided whether the child dropped (non-finite) infinitival/participial or finite morphology. When object or particle precede the bare stem, infinitival morphology is dropped. When object or particle follow the bare stem, finite morphology is dropped. In (48) some examples are given:

(48)　a.　jij bouw trein   Abel 2;05.27
　　　　　you build-*ø* train
　　　b.　die heef snor   Daan 2;04.28
　　　　　that has-*ø* moustache
　　　c.　Audrey slaap nog een tijdje   Josse 2;07.20
　　　　　Audrey sleep-*ø* still for a while
　　　d.　dese hoor niet daar   Laura 3;03
　　　　　that belong-*ø* not there
　　　e.　zit ook pitten in   Matthijs 2;10.21
　　　　　sit-*ø* also pits in
　　　f.　hij zeg toettoet   Peter 2;03.07
　　　　　he say-*ø* toettoet

If the overused bare forms are finite, it is expected that utterances that contain forms with and without inflection have a similar use. More precisely, it is expected that all utterances in (48a-f) - and all other examples in Appendix 3.3 - denote present tense, just like their equivalents in which the verb carries a suffix [-*t*].[46] This prediction is borne out: the bare stems in Dutch do not display the free use of RIs, but they pattern like the inflected present tense forms.[47,48] Additional motivation for the claim that the overused bare stems in Dutch child language are finite comes from the observation that the percentage of states in the bare stem forms clearly exceeds the percentage of states in RIs. There are 37 stative bare stems (i.e. approximately 40 %), whereas states hardly appear in RIs (Wijnen, 1997; see Chapter 4). Based on these observations, it can be concluded that it is very plausible to assume that the set of English RIs contains finite utterances (i.e. overused bare stems that are either the result from inflection drop or from wrong inflection), and, that these utterances affect the overall interpretation assigned to the English RIs. To get some idea of the effect that the HSE may take, I calculated the average percentage of modal use in a set of utterances that contains RIs and utterances with overused-bare stems (BS). Again I want to stress that the aim of these counts is merely to illustrate the likelihood that the HSE has a noticeable effect that may even cause a statistically significant difference between the meaning assigned to RIs in Dutch and English child language. The result of the counts is not necessarily representative for the size that the HSE has on the interpretation of English RIs.

---

[46] Theoretically, past tense interpretations are also expected. However, children at this age use in these cases hardly any past tense forms. If they use past tense form, they use irregular forms. Past tense is not marked by inflection but it is expressed by vowel change. I will deal with this issue in Chapter 5.

[47] De Jong (1999) reports that overuse of bare stems is found in the speech of Dutch SLI children. He attributes this drop of the inflectional ending to a lack of processing resources. Extending this to the data from the normally developing, but clearly younger children, it can be hypothesised that the children in my sample drop inflection because of a lack of processing resources. An explanation along these lines is compatible with the random character of inflection drop. If there was an underlying grammatical deficit, it were expected that inflection is dropped across the board. However, children drop inflection sometimes and at other times, they don't.

[48] The use of bare stems that lack the overt signs of tense and agreement but have a specific temporal reference is also found in Swahili child language (Deen & Hyams, 2001).

Table 3.18: The Heterogeneous Set Effect illustrated with Dutch corpus data, number of overused bare stems (BS), sum of interpretable RIs and BS, number and percentage of modal use within (the combined set of) RIs and BS, average percentage and standard deviation, data from all six children

	N_{BS}[49]	$N_{BS + RI}$[50]	$N_{MOD. BS + RI}$	$\%_{MOD. BS + RI}$
Abel	10	26	7	27
Daan	16	43	13	30
Josse	19	45	18	40
Laura	22	104	28	27
Matthijs	8	84	34	40
Peter	17	69	32	46
Sum	*92*	*371*	*132*	*-*
Average	*-*	*-*	*-*	*35*
SD	*-*	*-*	*-*	*10*

In the Dutch data, the average percentage of modal utterances shows a clear decrease: it goes down from 50 % to 35 % if the HSE is included. Thus, if the Dutch data are made comparable to the English data, in the sense that only RIs with third person singular subjects are counted *and* if sentences that contain a bare stem are added, the apparent modal preference disappears and a preference for non-modal use is observed in the Dutch results. Extending this observation to the cross-linguistic comparison, I conclude that the HSE can theoretically account for the observation that RIs in English child language are less often modal than RIs in Dutch child language. Like the *Infinitival Morphology Hypothesis* (IMH) proposed by Hoekstra & Hyams (1998), the HSE predicts a difference between RIs in Dutch and English child language. The IMH relates this difference directly

---

[49] Utterances with first person singular subjects are not included as in Dutch these forms cannot be distinguished from bare forms with dropped inflection.

[50] These are the RIs that remain after all corrections. Thus, only third person singular subjects and Peter's data are corrected for use of proper name instead of *I*.

to presence vs. absence of infinitival morphology. According to the HSE, presence vs. absence of infinitival morphology combined with OV/VO-order explains the difference between the two languages. The empirical advantage of the HSE is that this account does not predict a restricted use of RIs in Dutch (or English) child language. Because of this, the HSE, unlike the IMH, does not ignore the occurrence of ongoing RIs in Dutch child language.

3.7.2   Other cross-linguistic differences explained

Wexler (1994) and Harris & Wexler (1996) argued that English children use RIs, just like Dutch, Danish, German, Swedish, Norwegian and French children do. Hoekstra & Hyams (1998) pointed out that there is a difference in the interpretation of English RI and Dutch and German RIs. In the previous section, I discussed a factor that contributes to the cross-linguistic difference in the meaning of RIs: the HSE.[51] In this section, I will point to a number of other differences between English RIs, on the one hand, and RIs in other languages (that is, more prototypical RI-languages that possess a recognisable infinitival form) that are explained by the HSE.

First, there is a difference in predicate type between English, on the one hand, and Dutch/French RIs, on the other. Deen (1997) and Madsen & Gilkerson (1999) find respectively 25 % and 40 % stative predicates in English RIs. The percentages Wijnen (1997) and Ferdinand (1996) report for Dutch and French are lower: less than 10 % of the RIs in these languages are stative. However, stative predicates are frequent in finite

---

[51] As mentioned before, I assume that the interpretation of RIs in intrinsically free. This applies to RIs in Dutch but also to the subset of RIs in English that contains 'real' RIs (and not finite items with dropped inflection). As other, newly acquired, verb forms take over the function of RIs, the order of acquisition and relative productivity of these new forms has an effect on the overall interpretation given to RIs. Thus, the interpretation (or use) of English RIs is not only influenced by the HSE, but also by another factor: the acquisition other forms. As far as I know, the development of different verb forms over time in English child language has not been studied yet. Therefore, I cannot determine whether or not this factor is relevant for the study of cross-linguistic differences. However, before detailed comparisons can be made between the different languages more insight into the acquisition of different verb forms in different languages is required.

sentences in both languages. As English RIs contain finite sentences with dropped inflection, it is expected that English RIs contain more stative predicates than Dutch and French RIs. A second difference concerns WH-words. Phillips (1995) summarises results from different studies: he reports that German, Dutch, Swedish and probably French show an asymmetry between finite sentences and RIs with respect to WH-questions. Finite sentences contain WH-questions while RIs do not (Kursawe, 1994 for German; Haegeman, 1994 for Dutch; Santelmann, 1994 for Swedish; Crisma, 1992 for French). In all these languages, RIs differ clearly from finite sentences in verb form and/or verb placement. Phillips did not find the same effect for English though: in child English, both finite sentences as well as RIs contain WH-questions. Given that in the English data the total number of RIs includes finite sentences, this difference is to be expected.[52]

3.7.3 Some final notes on bare stems in child English

Hyams (2001a, b) reports a contrast between the interpretation of bare verb forms in child English and forms that end on [-*s*], that is, third person singular SFs.[53] The former are predominantly ongoing whereas the latter are most often habitual (like in adult English). This observation might have the implication that RIs in English child language are more often habitual than RIs in Dutch child language, because English RIs contain forms that are, in their underlying semantic representation, SFs. I made an attempt to find out

---

[52] The reverse prediction is that WH-questions in Dutch are expected to appear with dropped inflection (like in English). The problem is that this is difficult to determine as in Dutch WH-questions with third person subjects, inflection and subject cliticise, as illustrated in (ii):

    (i)   wat doet hij/die?
                 what does-*fin* he
                 'What does he do?'

    (ii)   wat doetij/doetie

By implication, sentences with pronominal subjects cannot be used to determine whether or not inflection is really dropped. Sentences with full DP subjects, however, could provide a test case. I will leave this as a suggestion of future research.

[53] Hyams cites results from Madsen & Gilkerson (1999).

more about the habitual use of RIs (and other verb forms) in the Dutch corpus data, and found hardly any habitual RIs. I have to admit, though, that I found it very hard to define any criteria to decide whether or not an utterance was habitual or not. The only more or less clear cases occurred when children talked about animal sounds (the pig grunts, the dog barks, etc.). However, even when the children were looking at the same time at images of pigs or dogs, it was still questionable as to whether the utterance was habitual. Therefore, I desist from drawing any conclusions on this issue.[54]

One could also infer from the contrast that Hyams describes that the HSE may not adequately describe the difference between Dutch and English: Hyams' study suggests that relatively few SFs in child English denote ongoing events. Hence, the number of SFs with dropped inflection denoting ongoing events must also be small. By implication, the HSE may cover relatively few ongoing RIs in early child English. Until now, it is still unclear what the size and the nature of the difference between the interpretations of RIs in the two languages is, however. Thus, although an explanation of differences between the meaning of Dutch and English RIs like the HSE has to reckon with the different semantics of SFs in Dutch and English, future study has to point out whether or not an additional explanation for differences in meaning between Dutch and English RIs is really required.

Suppose that future research shows that the HSE does not suffice: what factors, other than the HSE, could explain the ongoingness of English RIs? The analysis that I will give for the Modal Reference Effect (MRE) of RIs in early child Dutch in Chapter 5 gives an idea for a possible solution. It turns out that the order of acquisition of various finite forms causes changes in the meaning of RIs. I will not go into the details of this analysis, but, given that finite forms in English differ from finite forms in Dutch, we could

---

[54] Given the kind of criteria that are used to decide whether or not RIs in child Dutch are modal - that is, no reference to an event that is ongoing at speech time or took place before speech time - it is furthermore possible that some of the modal RIs were, in fact, habitual. Both in the modal as well as in the habitual reading a reference is established to speech time of a time prior to speech time.

hypothesise that the developmental effect that takes place in Dutch and that leads to the MRE, does not take place in English.

I finish this section with a final remark on SFs in early child English that relates to the previous paragraph in which I mentioned changes over time in the children's system of verb forms. Considering the experimental data, the use of SFs by English-speaking children is ambiguous: the experimental results show that children have some idea of the habitual use of SFs specific to English, on the one hand, but they also suggest that a full grasp of the habitual restriction is lacking. The early awareness of the restrictions on SFs shows up in the children's preference to use progressives in the ongoing condition instead of SFs.[55] However, if the SFs were restricted to habitual readings and bare stems to ongoing readings (which is, basically, the claim made by Hyams), the expectation would be that the English subjects in the experiment used exclusively bare stems and no forms with [-s] in the ongoing condition. This expectation is not borne out, as 19 % of the verbs in the ongoing condition were bare, whereas 15 % carried an [-s] ending. Thus, the children used approximately as many SFs as RIs in ongoing condition. This may suggest that they do not have the full grasp of the target system and (over) use SFs relatively frequently to denote ongoing events.[56] I leave this as a suggestion for future research; preferably longitudinal research that shows whether forms with [-s] become more specific (i.e. habitual) over time.

3.8 Summary

In this chapter, I focused on the temporal, modal and aspectual denotation of RIs. Corpus data from six children acquiring Dutch showed that the children preferred to use RIs for modal events. Non-modal usage was

---

[55] The results in Table 3.8 (p. 93) indicate that it is more likely that the English children use a progressive form on [-ing] in the ongoing condition than a form that ends on [-s]: the percentages are 60 % and 15 %, respectively.

[56] As in adult English, verbs denoting states appear more often as finite forms on [-s] than verbs denoting events - and states even require [-s] and cannot appear as progressives - one could hypothesise that young children do not yet map aspect and tense properly and overuse the [-s] on the basis of a generalisation over state verbs.

certainly not excluded, however: this confirms that there is a Modal Reference Effect (MRE) in early Dutch child RIs. In this respect, I found more evidence for the No Tense Hypothesis (Behrens, 1993; Wijnen, 1997), than for the Modal Hypothesis (Ingram & Thompson, 1996; Ferdinand, 1996; Hoekstra & Hyams, 1998). The strikingly marginal number of past/completed RIs suggests that Dutch RIs obey a Non-Completedness Constraint or NCC (Lasser, 1997). Experimental results confirmed a difference between the meaning of RIs in Dutch and English child language: Dutch RIs are more often modal than English RIs. The observed difference, however, was less sizeable than the difference that Hoekstra & Hyams (1998) reported, in a corpus-based study. I pointed to methodological choices that are responsible for the difference between experimental data and corpus results, and bring about a non-modal bias in the English corpus analysis. It was furthermore argued that RIs in child English are expected to be more frequently ongoing than RIs in Dutch child language because of the Heterogeneous Set Effect (HSE). As opposed to Dutch RIs, English RIs contain finite sentences in which the bare stem form is "overused". This asymmetry between Dutch and English RIs is the effect of morphological and syntactic differences between the two languages. The HSE is motivated by data from Dutch that support the claim that children indeed overuse bare stems because they either drop inflection or use incorrect inflections. The awareness that the HSE interferes in cross-linguistic comparisons does not only enable a better understanding of cross-linguistic differences in the meaning assigned to RIs but also explains various other asymmetries between RIs in early child English and RIs in languages that have distinct infinitival forms.

Types of Verbs

We have seen that Dutch and English RIs differ with regard to their temporal and modal denotation. This is not the only difference between RIs in the two languages: RIs in early child English contain numerous state-denoting predicates, whereas the number of stative predicates in Dutch RIs can be counted on the fingers of one hand. Towards the end of the previous chapter, I mentioned this asymmetry in verb types as support for the claim that English RIs contain finite verbs, while Dutch RIs are non-finite across the board. Naturally, this explanation is only applicable given that simple finite sentences, SFs, are actually used by young children to denote states. In the present chapter, I will look more closely at the types of verbs that Dutch children use in their SFs and RIs. The data from Abel, Daan, Josse, Laura, Matthijs and Peter support the claim that Dutch RIs obey the Eventivity Constraint (EC) and contrast with SFs in this respect. The EC, however, is a *weak* constraint, since stative predicates are not excluded in RIs. It will be argued that, although the number of stative RIs is small in early child Dutch, they should not be ignored, as, remarkable as it may sound, the appearance of stative RIs provides us with a cue to understand why RIs are so often eventive. Following previous claims, it will be shown that the EC is related to the predominant modal use of RIs: in their RIs, children use kinds of modality that are semantically incompatible with stative predicates. Yet, at the same time, children do not use the kind of modality in RIs that goes well with states, because they lack the cognitive maturity that this particular kind of modality, i.e. epistemic modality, requires. In an attempt to disentangle the complex of factors involved in the EC, distributions in the input are found to be relevant as well.

4.1   Outline of this chapter

In the first chapter of this thesis, theories about the types of verbs that children use in RIs and SFs were divided into Dynamicity Hypotheses, on the one hand, and a Telicity Hypothesis, on the other hand. These theories do not only make different predictions with regard to the characterisation of the verb types in RIs and SFs, but also with regard to the overlap between the types of verbs that children use in RIs and SFs. In section 4.2, the predictions derived from the hypotheses are tested on the data of Abel, Daan, Josse, Laura, Matthijs and Peter. In 4.3, I evaluate the outcome of this test. On the basis of empirical arguments, the most promising explanation will be singled out for more discussion in section 4.4. Section 4.5 gives a summary of the findings and conclusions of this chapter.

4.2   Evidence for a weak Eventivity Constraint

Let me start with two recent claims. Ferdinand (1996), Wijnen (1997) and Hoekstra & Hyams (1998) argued that the difference between finite verbs in SFs and infinitives in RIs can be described as a distinction between states and events. Their conclusion is based on data from early child French (Ferdinand) and early child Dutch (Wijnen and Hoekstra & Hyams); the distinguishing feature is [±dynamic]:

(1)   *Dynamicity Hypotheses*
        (i)   Eventive predicates are excluded in early finite sentences
                (Ferdinand, 1996)
        (ii)   Stative predicates are excluded in RIs
                (Ferdinand, 1996; Wijnen, 1997; Hoekstra & Hyams, 1998)

Gavruseva (2001, 2002) defends another viewpoint.[1] She claims that the relevant difference is that between verbs that are inherently telic or atelic and verbs that become telic or atelic in combination with other words.

---

[1] Henceforth, I will refer to this theory as Gavruseva (2001), because the two studies make the same claim but the 2001-version is more elaborate than the 2002-version.

Skipping details, the second (underspecified) type of verb is expected to appear in RIs, whereas the first (specified) type of verb is expected to appear in finite forms (see § 4.2.3 for some discussion on the exact predictions of the Telicity Hypothesis).

(2)   *Telicity Hypothesis*
       Verbs that do not have an inherent telicity value appear in RIs,
       whereas [+telic] and [-telic] verbs appear in SFs. (Gavruseva, 2001)

Theories about the types of verbs that children use in RIs and SFs can be classified on the basis of the aspectual features dynamicity and telicity. A second division can be made on the basis of the predicted overlap between verb types in SFs and RIs. Ferdinand (1996) predicts a no-overlap pattern in the child data: states appear in SFs and events are used in RIs. As a verb is either stative or eventive, it is predicted that a verb type that appears in an SF does not appear in an RI, and vice versa. The other theories predict overlap: the same verb type may show up in RIs and SFs.

I continue with an introductory section on aspectual verb classes. This section contains background information and gives the criteria that were applied in this study to classify verb types in children's SFs and RIs.

4.2.1   Aspectual classes, terminology and tests

Vendler (1967) divided verbs into four classes: activities, accomplishments, achievements and states. He gave the following definitions (Vendler, 1967: 106):

"For activities: A was running at time t means that time instant t is on a time stretch in which A was running.

For accomplishments: A was drawing a circle at t means that t is on the time stretch in which A drew that circle.

For achievements: *A won a race between t_1 and t_2* means *the* time instant at which *A* won the race is between *t_1 and t_2.*

For states: A loved somebody from t_1 to t_2 means that at any instant between t_1 and t_2 A loved person"

States and achievements denote a process in time, whereas activities and accomplishments do not.[2] With respect to the state-event contrast that plays a role in the theories about child language, Vendler's classes cluster as follows: activities, accomplishments and achievements are events, and contrast in *dynamicity* with states. Comrie (1976: 48) suggested that events, but not states, require an input of energy for the maintenance of the eventuality. Pustejovsky (1991: 56) described a state as "a single event, which is evaluated relative to no other event". Summarising various claims in the literature, Wanner (1999) gave the following list of stative verbal predicates: perception verbs (*taste, feel, smell*), position verbs (*sit, lie, stand*), verbs of existence (*be, seem, exist*), psych verbs (*fear, like*) and verbs of possession (*have, possess, own*). In his study on child Dutch, Wijnen (1997) classified modal verbs that appear as main verbs in Dutch, as states. These are verbs like *kunnen* ('be able to'), *moeten* ('have to') or *willen* ('want to'). I follow Wijnen's approach as the criteria for states, as given below, apply to the Dutch modal main verbs. In particular in studies on child language, the modal verbs turn out to be important: they appear very early in Dutch child language.

There are several tests to distinguish between states and events. According to Lakoff (1966), the following contexts only allow eventive verbs: imperatives, sentences with the adverbs *deliberately* and *carefully*, and complements of verbs such as *force* or *persuade*. Another test that is often applied to distinguish states from events is the present progressive test: states are incompatible with present progressive. Some confusion may be caused by states that can switch their dynamicity value and shift from state to event (Dowty, 1979; Quirk, Greenbaum, Leech & Svartvik, 1985). Quirk et al. (1985: 201 ff.) write:

> "The definition of stative verbs is not so much that they are
> incompatible with progressive, as that when they are combined with

---

[2] Dowty (1979) gives the following formalisation of Vendler's classes:

STATE:	[x BE [z]]
ACTIVITY:	[x DO [z]]
ACHIEVEMENT:	[y BECOME [(AT) z]]
ACCOMPLISHMENT:	[[x DO [z_1]] CAUSE [y BECOME [(AT) z_2]]]

the progressive, some change of interpretation other than the addition of 'temporary' meaning of the progressive aspect is required. This change of interpretation can usually be classified as a transfer, or reclassification of the verb as dynamic, eg. as having a meaning of process or agentivity."

According to Dowty (1979), the set of stative verbs includes perception verbs. Wanner (1999) adds verbs of position. While the feature dynamicity distinguishes between states and the various events, the factor telicity distinguishes between states and activities, on the one hand, and achievements and accomplishments, on the other. Telic predicates can be modified with the adverbial phrase *in an hour* while atelic predicates can appear with *for an hour*. Traditional examples of achievements are the telic instantaneous verbs or VPs *die, arrive, find a wallet* or *recognise*. Examples of accomplishments are telic predicates that involve a process in time, such as *run a mile, read the book* or *travel from X to Y*. Verkuyl (1972) noticed that 'telicity' is compositionally derived. To distinguish between the telicity that is part of the lexical meaning of the verb, and the 'telicity' on the level of the VP or sentence, he introduced the term 'terminativity'. Terminativity contrasts with durativity. The examples in (3) show how properties of the verb's internal and external argument affect the terminativity or durativity of a sentence:

(3) a. Mary walks three miles   TERMINATIVE
     b. Mary walks miles and miles   DURATIVE
     c. Children walk three miles   DURATIVE

Having introduced the notions dynamicity and telicity, I will now turn to the patterns in child Dutch. Recall that one prediction is that there is no overlap in the verb types that appear in SFs and RIs and that the former contain stative predicates, whereas the latter are eventive (Ferdinand, 1996). Another prediction is that there is overlap between SFs and RIs: the former contain stative and eventive predicates, while the latter are restricted to event-denoting predicates (Wijnen, 1997; Hoekstra & Hyams, 1998). The prediction of this second view is described as a constraint, i.e. the Eventivity Constraint or EC ('RIs are restricted to event-denoting predicates'), a term that has been coined by Hoekstra & Hyams. Finally, it is predicted that there is overlap between SFs and RIs, in the sense that SFs contain [+telic] and

[-telic] verbs, whereas RIs contain [αtelic], [+telic] and [-telic] verbs (Gavruseva, 2001). As it is not obvious how this set of predictions follows from the Telicity Hypothesis, I will give an explanation in § 4.2.3.

4.2.2   Testing the Dynamicity Hypotheses

Ferdinand's no-overlap claim is not confirmed by the data from Abel, Daan, Josse, Laura, Matthijs and Peter, as there is overlap between the verb types that appear in SFs and RIs during the RI-period of the six children. Table 4.1 shows that this overlap is not even marginal: approximately half of the verb types in SFs overlap with verb types in RIs. In Table 4.2, the verb types that overlap are given. Appendices 4.1 (p. 248) and 4.2. (p. 251) contain exhaustive lists of the verb types that are used in RIs and SFs.

Table 4.1: Numbers of different verb types in RIs and SFs and numbers of overlapping verb types in RIs and SFs, data from all six children

Child	N verb types RI	N verb types SF	N overlap verb types RI - SF
Abel	46	24	10
Daan	53	35	13
Josse	53	21	8
Laura	55	23	8
Matthijs	73	35	16
Peter	33	33	18

Table 4.2: Overlapping verb types in RIs and SFs, data from all six children

Child	Overlapping verb types in RIs and SFs
Abel	doen 'do', gaan 'go', hebben 'have', komen 'come', maken 'repair', nemen 'take', staan 'stand', vallen 'fall', zetten 'put', zitten 'sit'
Daan	doen 'do', hebben 'have', horen 'hear', komen 'come', leggen 'put', liggen 'lie', maken 'repair', pakken 'get', slapen 'sleep', spelen 'play', stoppen 'stop', zit 'sit', zoeken 'search'
Josse	doen 'do', gaan 'go', klimmen 'climb', slapen 'sleep', springen 'jump', staan 'stand', vallen 'fall', zitten 'sit'
Laura	doen 'do', sooien [= gooien, EB] 'throw', hebben 'have', knoeien 'spill', komen 'come', zingen 'sing', zitten 'sit', zwemmen 'swim'
Matthijs	doen 'do', gaan 'go', geven 'give', gooien 'throw', hebben 'have, klimmen 'climb', komen 'come', liggen 'lie', maken 'repair', rijden 'ride', slapen 'sleep', vinden 'find', zetten 'set', zien 'see', zitten 'sit', zoeken 'search'
Peter	branden 'burn', doen 'do', draaien 'turn', dragen 'carry', geven 'give', gooien 'throw', hebben 'have', eten 'eat', kijken 'look', liggen 'lie', maken 'repair', pakken 'take', rijden 'ride', staan 'stand', vallen 'fall', zetten 'put', zien 'see', zitten 'sit'

The list of overlapping verb types in Table 4.2 contains states as well as events. This indicates that states and events appear in SFs and RIs. Given this observation, the conclusion might be draw that the EC gives an inadequate description of the data, because it excludes the appearance of states in RIs. Put differently, the EC, as it stands now, would predict that the predicates in Table 4.2 are restricted to eventive predicates (which is not the case). The EC could be maintained, but only in a weaker variant, if there is a clear asymmetry between SFs and RIs in the sense that the number of eventive SFs is significantly greater than the number of stative RIs. In order to determine whether or not there is evidence for such an adjustment of the EC, the size of the overlap-effect in the set of verb types in RIs and in the set of verb types in SFs has to be determined.

A comparison of the verb types in the Appendices 4.1 and 4.2 suggests that verb selection in SFs is fairly unconstrained: there is no preference for either states or events. In RIs, however, events are clearly favoured. Table 4.3 gives the numbers and percentages of states in RIs, whereas Table 4.4 shows the results for SFs. The number in parenthesis in the third column represents the stative verbs that are unambiguously stative, because they do not allow for type shifting to eventiveness (non-type shifting predicates, i.e. NTS).

Table 4.3: Numbers and percentages of stative predicates in RIs (qualitative results), data from all six children

	$N_{\text{verb type RI}}$	$N_{\text{stative RI (NTS)}}$		$\%_{\text{stative RI (NTS)}}$	
Abel	46	5	(2)	11 %	(4 %)
Daan	53	6	(1)	11 %	(2 %)
Josse	53	4	(1)	8 %	(2 %)
Laura	55	2	(1)	4 %	(2 %)
Matthijs	73	4	(1)	5 %	(1 %)
Peter	33	6	(1)	18 %	(3 %)

Table 4.4: Numbers and percentages of stative predicates in SFs (qualitative results), data from all six children

	$N_{\text{verb type SF}}$	$N_{\text{stative SF (NTS)}}$		$\%_{\text{stative SF (NTS)}}$	
Abel	24	15	(10)	63 %	(42 %)
Daan	35	17	(11)	49 %	(31 %)
Josse	21	12	(8)	57 %	(38 %)
Laura	23	12	(8)	52 %	(35 %)
Matthijs	35	15	(10)	43 %	(29 %)
Peter	33	14	(7)	42 %	(21 %)

All six children show an amazingly similar pattern, not only with regard to the asymmetry between RIs and SFs but also with regard to the proportions of states within RIs and SFs. On the basis of a comparison between the results in Tables 4.3 and 4.4, I conclude that the EC describes the pattern in children's RIs and early SFs adequately, provided that the EC is interpreted as a weak constraint: it describes children's *tendency* not to use states in RIs rather than that it forces the exclusion of states in RIs.

The studies that defend the Dynamicity Hypothesis deal differently with the weakness of the EC, or rather, they have different ways of ignoring the weakness. For Ferdinand (1996), the issue does not exist. She argues that all states in RIs in early child French are type-shifting predicates that denote events when they are used in RIs. This type-shifting is forced by the modal meaning of RIs, according to Ferdinand (1996). Wijnen (1997) observes that states do appear in RIs in early Dutch child language. He reports furthermore that states prefer modal RIs. Wijnen does not explain these observations, even though they pose a problem for his account. Hoekstra & Hyams (1998) make use of Wijnen's data, but neglect the small percentage of stative RIs. In sum, there seems to be an explanatory gap with regard to the occurrence of states in RIs. Later on in this chapter, the weakness of the EC is discussed in detail. It will be shown that specifically its weak character is crucial for understanding one of the factors that contribute to the absence of states in RIs.

4.2.3 Testing the Telicity Hypothesis

In order to yield predictions from Gavruseva's (2001) proposal, I first go into some of the details of her proposal that were not mentioned in the overview in Chapter 1. My interpretations may go beyond the original proposal; it will be mentioned when predictions are derived on the basis of my interpretations on issues where the theory is unclear.

As discussed in Chapter 1, Gavruseva claims that verbs that are underspecified for telicity, notated here as [αtelic][3], cannot appear as finite verbs in SFs as long as children are unable to derive the telicity value of the VP syntactically. As an effect, [αtelic] predicates surface as infinitives (resulting in RIs). Taking children's use of RIs as an indicator for the inability to compute the telicity value of the VP, I come to the following contrast: finite verbs in SFs are [+telic] or [-telic], whereas RIs contain verbs that are [αtelic]. As long as the children use RIs, [αtelic] predicates will not appear in SFs.

RIs in child language can be argued to be due to grammatical deficits. However, when *adults* use RIs, absence of grammatical knowledge cannot serve as an explanation for the omission of finite verbs, since adults have a full-fledged grammar. Thus, the occurrence of adult RIs suggests that the grammar allows for omission of finiteness.[4] Why do I turn so suddenly to RIs in adult language? Gavruseva makes reference to RIs in adult language (p. 54-55) may imply that she assumes that finiteness (or tense) is not grammatically required. If finiteness is not required, verbs that are inherently specified for telicity can freely appear in RIs. By implication, the Telicity Hypothesis predicts overlap between the verbs that appear in SFs and RIs in the sense that [+telic] verbs and [-telic] verbs are allowed in both sentence types. Adult use of RIs shows that RIs are not only indicators of an underspecified grammar, for the simple reason that adults do not have an underspecified grammar anymore. RIs that are not the output of an underspecified grammatical system are expected to occur in child language as well as adult language. Taking this rationale one step further, it can be argued that there is no reason whatsoever for these RIs to be restricted to [αtelic] predicates. The assumption can be made that this particular set of RIs, i.e. the set of RIs that are used by both children and adults, allows for all types of predicates. In sum, the finding that child RIs contain, apart form [αtelic] verbs, [+telic] or [-telic] verbs, does not contradict the Telicity Hypothesis. These [+telic] or [-telic] RIs are expected, however, to be

---

[3] The underspecified value must be distinguished from a negative specification, [-telic], which is a specified value.

[4] See Avrutin (1999), who gives an account for RIs in adult language, according to which finiteness is required by rules that are part of the language module regulating discourse rather than by grammatical rules.

similar to the RIs that are allowed in adult Dutch. If this expectation is not borne out and if children use RIs that are specified for telicity, on the one hand, but that differ from adult RIs, on the other hand, we are left with an unexplained category of RIs in child language. For this reason, it will be tested whether or not such non-adult RIs occur.

What do the data show? First of all, SFs that contain [αtelic] verbs are not ruled out during the RI-period of Dutch children, as illustrated by the examples in (4). Verb like *doen* 'do', *zingen* 'sing' or *draaien* 'turn' are not inherently telic but become telic when a direct object or particle is added, as in *de afwas doen* 'do the dishes', *een liedje zingen* 'sing a song' or *een rondje draaien* 'turn around in a circle':

(4) a.   hij doet zo   Abel 2;05.27
          he do-*fin* so

       b.   hij zingt   Daan 2;00.29
          he sings-*fin*

       c.   die draait   Peter 2;00.28
          that turns-*fin*

Secondly, is there evidence for the children using RIs with verbs that are intrinsically [+telic] or [-telic], but that cannot be used by adult speakers of Dutch? The RIs in (5) contain [+telic] verbs, whereas the RIs in (6) contain [-telic] verbs:

(5) a.   ik ook opschrijven   Abel 2;01.16
          I also down write-*inf*

       b.   dichtdoen !   Daan 2;00.29
          close-*inf*

       c.   Ab maken   Josse 2;00.07
          Ab make-*inf*

(6) a.   ja jij niet goed (be)waren !   Daan 2;08.13
          yes you not good keep-*inf*

       b.   mama niet hebbe   Laura 2;06.10
          mama not have-*inf*

       c.   woef liggen   Peter 1;09.20
          woef [= dog, EB] lie-*inf*

From the examples in (5) and (6), only (5b) is allowed in adult Dutch. Hence, it can be concluded that Gavruseva's proposal does not cover all the RIs that Dutch children use. There are two possible conclusions: either RIs in child language are caused by an additional lack of knowledge (besides their inability to syntactically derive telicity) or the Telicity Hypothesis is not on the right track. I decide for the latter option, because there are various problems with the Telicity Hypothesis other than the one just mentioned.

The Telicity Hypothesis makes the incorrect prediction that SFs with an undetermined telicity value do not occur during the RI-period. Secondly, the Telicity Hypothesis could not account for the observation reported by Bar-Shalom & Snyder (1998) that there is a difference in frequency of RIs between Polish and Russian child language. According to the Telicity Hypothesis, no difference between these two Slavic languages is predicted: they are expected to show equally few RIs. The Russian data are problematic as they suggest that Russian children use a considerable amount of RIs and that they are, in this respect, much more similar to children acquiring Germanic languages than expected given the Telicity Hypothesis (see Chapter 1, § 1.6.2, for a more elaborate discussion of this problem).[5]

4.2.4   Summary

The Dutch findings did not resemble Ferdinand's (1996) observations for child French. Ferdinand found a strict no-overlap pattern that matches the state-event contrast. Ferdinand writes that "a verb like *avoir* 'have' is always finite, both as an auxiliary and as a lexical verb with the possessive reading" (p. 102). Thus, when French children used type-shifting verbs in RIs, these *always* denoted the inception of a state. Although the Telicity Hypothesis predicts overlap, the results reported in this section did not support this hypothesis either. The most adequate description for the patterns of verb type selection in RIs and SFs is given by the EC, provided the EC is interpreted as a weak constraint. Put simply, the weak EC states that similar verb types are allowed in SFs and RIs, but there is an asymmetry between

---

[5] The use of RIs by Russian children is discussed also in the work of Bar-Shalom & Snyder (1999), Brun, Avrutin & Babyonishev (1999) and De Bode (1999).

SFs and RIs in the sense that SFs contain significantly more state-denoting predicates than RIs do.

4.3 Syntactic and semantic approaches to the EC

Earlier in this chapter, I mentioned two kinds of Dynamicity Hypotheses: those that predict no-overlap and those that allow overlap, more specifically, predict that only verb type selection in RIs is constrained by the EC (and that SFs show free selection). The Dutch results support the hypotheses that allow for overlap and argue for the EC. If we now turn to the Dynamicity Hypotheses that account for the EC, two different types of explanations can be distinguished. Both explanations are related to the issues discussed in Chapter 3: the EC is explained as an effect of the absence of Tense (i.e. the No Tense Hypothesis) or from the modal meaning of RIs (i.e. the Modal Hypothesis) (see Chapter 3, section 3.2 for more details about these hypotheses). Figure 4.1 gives an overview. In the previous section, § 4.2.2, I zoomed in on the upper part of Figure 4.1, that is, on the hypotheses that are not bold-faced. The present section concentrates on the lower part of the figure, i.e. the bold-faced hypotheses:
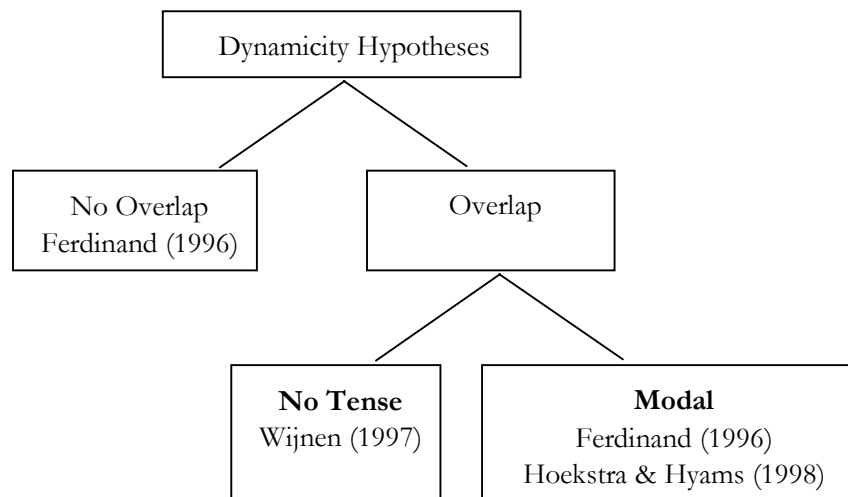
Figure 4.1: Dynamicity Hypotheses

The theory that relates the EC to the absence of Tense in RIs takes a syntactic point of view, in the sense that the syntactic representations that children can build (i.e. representations without TP) are incompatible with formal properties of states (i.e. no event variable). The theory that relates the EC to the modal meaning of RIs takes a semantic viewpoint, because it stresses the semantic interaction between modal verbs and the verb types that appear as the complements of modal verbs.

In the following sections, it will be shown that the semantic Modal Approach (MA) is more fruitful than the syntactic No Tense Approach (NTA). The explanatory potential of both models will be examined by means of an empirical test, followed by a discussion of the possibilities that the models offer for explaining two observations that relate to the EC. The outcome pleads for an account along the lines of the MA.

4.3.1   Test case PVs

Until now, only one type of finite utterances has been discussed, and that was SFs. However, children also use periphrastic verbs (PVs) in the RI-period, as we have seen in the previous chapter. PVs share with SFs the property that they contain a finite form and hence, that they are tensed. With regard to modality, the results in Chapter 3 show that PVs are more like RIs than like SFs, however. In Table 4.5, the temporal/modal interpretation of SFs, RIs and PVs is summarised to give an impression of the distributions of modality in the three sentences types. Recall, that PVs are PVs with an infinitival complement. I excluded PVs with participles (cf. Chapter 3, § 3.3.6). Unlike the SFs discussed in the previous chapter, the SFs examined here contain thematic verbs and non-thematic verbs like auxiliaries and copula.

Table 4.5: Modal use of interpretable SFs, PVs and RIs compared, data from all six children collapsed[6]

	N	N$_{MODAL}$	% $_{MODAL}$
SFs	3185	93	3 %
PVs	691	512	73 %
RIs	1248	912	73 %

The results in Table 4.5 show a very significant difference between the modal use of SFs and PVs: the former are non-modal (predominantly present tense) whereas the latter denote various modal meanings. The average percentage of modally used PVs equals the percentage of modal RIs. Peter's data lead to a decrease of modality in PVs because he uses a considerable number of ongoing PVs. If Peter is excluded from the sample, the average percentage of modal PVs increases to 81 % (whereas the average percentage of modally used RIs remains similar).

On the basis of the results in Table 4.5, it can be concluded that PVs provide a test case for the MA versus the NTA. If stativity is correlated with the absence of Tense (NTA) states are expected to be absent in RIs only and not in SFs and PVs. If stativity is negatively correlated with modality (MA), it is expected that states are absent in RIs and PVs, and not in SFs. Table 4.6 contains the results of this test:

---

[6] All children behave very similarly, and differences between the subjects are marginal. Therefore, the data are collapsed. The same holds for the results in Table 4.6.

Table 4.6: Distribution of stative predicates over SFs, PVs and RIs, data from all six children collapsed

	N	N$_{STATIVE}$	% $_{STATIVE}$
SFs	3242	2645	82 %
PVs	810	70	9 %[7]
RIs	1565	60	4%

The following schema summarises the observations:

Table 4.7: Tense, modality and eventivity in SFs, PVs and RIs, general patterning in early child Dutch

	Tense	Modal	Eventive
SFs	+	-	-
PVs	+	+	+
RIs	-	+	+

On the basis of the patterns in Table 4.7, it can be concluded that eventivity correlates with modality and not with absence of Tense. This provides an empirical argument in support of the MA.

4.3.2   States in RIs

A first observation that will be discussed in the light of the NTA and the MA is the weakness of the EC. Are the two approaches able to deal with the observation that states are allowed in RIs?

---

[7] Most states in PVs appear quite late, that is, when the children use 'can' and 'may' with an infinitival verb form that denotes a position (*kan staan* 'can stand'*, mag zitten* 'may sit', etc.). The children hardly use these modal meanings in their RIs.

All six children use states in their RIs, such as the verb of possession *hebben* 'have' and verbs that denote a location or position.[8] According to the NTA, RIs are sentences that lack Tense. RIs do not allow for states, because states need to be temporally anchored via Tense. Thus, the NTA cannot explain why states appear in RIs, even though the number of stative RIs is marginal. The MA can deal with weakness of the EC, however. Basically, the idea of the MA is that selection of the verbal predicate is constrained by the kind of modality that is expressed in an utterance. The results of the previous chapter show that there is an asymmetry with regard to the modal use of RIs and SFs in Dutch child language: RIs are most often modal, whereas SFs are not modal and nearly always denote ongoing present tense events. Based on the difference in the modal use of SFs and RIs, the MA predicts that SFs select freely for states, whereas RIs do not (being constrained by their modal character).[9] Turning now to the small set of states that appear in RIs, we can think of two possible explanations within the framework provided by the MA. The first option comes from the set of RIs that are *not* modal, which covers one quarter of all the interpretable RIs: the MA does not apply to the ongoing RIs and hence, it may be that the stative RIs are restricted to the set of ongoing RIs. A second possibility is that the set of modal RIs does not behave uniformly with respect to the selection of verbal predicates: one kind of modality may be compatible with states, whereas another kind of modality may not. Irrespective of which explanation is on the right track, the MA predicts an asymmetry between SFs and RIs, in the sense that SFs more frequently contain states than RIs, but the MA does not necessarily exclude states in RIs. Thus, in the light of the observed weakness of the EC, the MA seems more promising than the NTA.

---

[8] States that appear frequently in SFs such as modals, copula and various mental verbs do not appear in RIs. Therefore, the asymmetry in verb type between SFs and RIs concerns basically auxiliary-like verbs like modals and copula and some mental verbs.

[9] Note that, because of the topic of the present chapter, an additional analysis must be made that includes non-thematic verbs such as auxiliaries and copula in the set of SFs. The aim is here to find out whether or not there is a correlation between modality and verb type (and not, as in the previous chapter, to find out if RIs differ in meaning from comparable finite clauses).

4.3.3   The EC in adult Dutch

Before, it was concluded that the EC is not a particular property of RIs, but also that it applies to the PVs of Dutch children, and, hence applies to children's sentence-final infinitives in general (§ 4.3.1). We can generalise even further when adult Dutch is taken into account, because sentence-final infinitives in adult Dutch also obey the EC. An analysis of the caregivers' speech in the corpora of Abel, Daan, Josse, Laura, Matthijs and Peter illustrates this claim. If the data from the children's caregivers are examined, it turns out that eventive and stative predicates are unevenly distributed, more specifically, that the distributions of eventive and stative predicates resemble the distributions of these predicates in the children's data. In Table 4.8, the results of this investigation are summarised. FINs are all sentences with finite verbs in first or second position, whereas INFs comprise all sentences with a sentence-final infinitive. In the set of FINs, the finite verbs in SFs and PVs are collapsed. In the set of INFs, infinitives in PVs and adult RIs are collapsed.

Table 4.8: Distributions of stative predicates over finite and infinitive main verbs in Dutch child-directed speech

	N	$N_{STATIVE}$	$\%_{STATIVE}$
Finite verbs FINs	8915	7922	89 %
Infinitives INFs	3500	89	3 %

Table 4.8 displays an asymmetric distribution: there are hardly any infinitival states whereas finite forms are predominantly stative. Like in the child data, these finite states predominantly comprise auxiliaries, modals and copula (N = 5920), and the rest (N = 2002) are verbs of position (*liggen* 'lie', *zitten* 'sit',

zijn 'be')[10], possession (*hebben* 'have'), and mental states *(denken* 'think', *vinden* 'find/think', *weten* 'know', *bedoelen* 'mean', etc.).[11]

As before, the MA has the potential to cover a more extensive empirical domain than the NTA. According to the NTA, the eventivity of RIs is a particular property of child language, because children lack knowledge of Tense, whereas adults do not. By implication, the similarity between the EC in child and adult language must be apparent and, according to the NTA, the two ECs have different causes.[12] Theoretically speaking, the MA can account for both ECs, however. The eventivity of RIs follows from (i) the relation between deontic modality and eventivity, on the one hand, and epistemic modality and stativity, on the other hand, and (ii) the non-use of epistemic modality (as opposed to deontic modality and near future).[13] Provided that adult speakers of Dutch use relatively few epistemic modal

---

[10] I distinguished between the copula *be* ('he is mad') and the verb of location *be* ('he is at home').

[11] See for similar observations about the uneven distribution of verb types of finite forms and infinitives in adult Dutch and Dutch child-directed speech Schlichting (1996:188) and Wijnen, Kempen & Gillis (2001). Schlichting compared child data to data from adult Dutch and observed that there were significant correlations between the two with regard to the formal properties of the verbs and types of verbs (Schlichting compared the two on a lexical level). Pine, Lieven & Rowland (1998) report high correlations between the production data of English children and the production data of their caregivers with regard to verb form-verb type patterns (see also Chapter 1, § 1.7).

[12] Given that the NTA needs an additional explanation for the EC in adult sentence-final infinitives (which are most frequently PVs) anyway, it could be the case that this additional explanation also accounts for the EC in the PVs in child Dutch, mentioned earlier in this chapter.

[13] Note that there are, apart from deontic modality and near future (which are the contexts that Ferdinand and Hoekstra & Hyams mention), more selecting contexts that prefer or even select for eventive infinitival complements, and that, hence, contribute to the EC. Consider the progressive:

(i)   Hawk is sporen van de dader aan het zoeken
        Hawk is traces of the offender on the search-*inf*
        'Hawk is searching traces of the offender'

(ii)   * Andy is de vader van de baby aan het zijn
        Andy is the father of the baby on the be-*inf*
        'Andy is being the father of the baby'

utterances, the explanation for the EC in child Dutch can easily be extended to the EC in adult Dutch.

4.3.4   Summary

Two explanations for the EC have been evaluated: the syntax-oriented No Tense Approach (NTA) and the semantics-oriented Modal Approach (MA). On the basis of the pattern that PVs display, I concluded that the MA is preferred as an explanatory model. The two approaches were also evaluated in the light of two observations: (i) the (marginal) appearance of states in RIs, and (ii) the similarity between child and adult language with regard to the EC. In both cases, the NTA showed a lack of explanatory power, whereas the MA offered possible explanations. The overall conclusion is that the MA is more promising in terms of empirical coverage. Therefore, I will continue this chapter by zooming in on the MA.

4.4   The Modal Approach to the EC

Recall that the MA offers two possible explanations for the appearance of stative RIs: stative RIs could be limited to the RIs that are not modal (but ongoing at speech time or denoting the here-and-now at speech time) or it could be that modal RIs represent a heterogeneous set with regard to selection restrictions. These two explanations do not exclude each other. Hence, both of them may apply. In the following sections, the data of the six children will be examined more closely in order to answer the question as to why RIs are so often eventive, but are not eventive across the board.

4.4.1   States in the here-and-now

If all stative RIs are ongoing, the first explanation is supported. If all stative RIs appear in utterances qualified by a certain type of modality, the second explanation holds. If both explanations hold, it is expected that there are

only a number of non-modal RIs in the set of stative RIs and that the rest must be a specific type of modal RI.

In the literature, there are some relevant observations on this issue. Wijnen (1997:17) noticed that the states in Dutch child RIs show a preference for modal/future interpretations:

> "Non-eventive verbs are not entirely absent from root infinitivals, though their number is marginal (7 % at most). Importantly, the temporal interpretation of these non-eventive root infinitives is restricted to the future."

A similar preference has been observed in the experimental study of Schönenberger, Pierce, Wexler & Wijnen (1995).[14] In this study, the proportion of future interpretations was almost twice as high for stative verbs (50 %) than for eventive verbs (27 %). These observations indicate that what appeared to be the most obvious explanation for stative RIs at first sight does not account sufficiently for the data on closer inspection. Not all stative RIs denote states that are ongoing at speech time; there are also stative RIs that denote states that possibly or necessarily take place after speech time (future). The results mentioned in this paragraph even suggest that modal stative RIs are preferred.

The data from Abel, Daan, Josse, Laura, Matthijs and Peter confirm this preference, as shown in Table 4.9:

Table 4.9: Modal interpretations of stative RIs, data from all six children collapsed

	N_{MODAL} (%)		$N_{ONGOING}$ (%)	
Stative RIs	37	(62 %)	23	(38 %)

Stative RIs are predominantly modal. More specifically, they express volition. For instance, all 31 instances of RIs with *hebben* 'have' were desiderative and expressed a wish. Some examples are given in (7) below.

---

[14] See Chapter 3, § 3.4.13, for more details on the work carried out by Schönenberger et al.

RIs that contain verbs of position/location contain type-shifting predicates that can be stative as well as eventive. In deontic RIs used to give a command, these verbs receive eventive interpretations such as 'sit down', 'lie down' or 'go stand'. The verb refers to the inception of a state, hence to an event. A stative verb like *blijven* 'stay' is used as 'don't go' and implicitly denotes the event of going. In the RIs used to express a desire (8a-c) or used to describe the situation at speech time (8d-e) there is no shift towards eventivity.[15] The children talk about a desired state, and, sometimes, they describe a property of the situation at speech time. Thus, occasionally, stative RIs are ongoing RIs, but most often, stative RIs are modal RIs, particularly volitional RIs.

(7)   a.   Ik ook een hebben   Daan 2;09.10
             I also one have-*inf*
             'I also want to have one'

       b.   soel [= stoel, EB] hebben   Abel 2;07.15
             chair have-*inf*
             'I want to have the chair'

       c.   speen en poesje hebben   Matthijs 2;05.10
             pacifier and cat have-*inf*
             'I want to have my pacifier and cat'

(8)   a.   koekjes in zitten   Abel 2;07.15
             cookies in sit-*inf*
             'cookies are in there'

       b.   hier sjaan   Daan 2;01.21
             hier stand-*inf*
             'I want to stand here'

       c.   bugge zitten   Josse 2;03.28
             buggy sit-*inf*
             'I want to sit in the buggy'

       d.   Thijsje liggen   Matthijs 1;11.24
             Thijsje lie-*inf*
             'Thijsje is lying'

---

[15] This kind of type shifting is also mentioned by Ferdinand (1996) for child French.

e. Woef liggen   Peter 1;09.20
    dog lie-*inf*
    'The dog is lying'

States appear in RIs with a here-and-now denotation (e.g. verbs of
location/position) and in RIs that express a desire (e.g. verbs of possession).
States are unattested in deontic and intentional RIs. These observations
indicate that there is support for both explanations that the MA provides for
the allowance of stative RIs. The pattern within the set of modal RIs, that is,
the restriction of states to volitional RIs expressing a desire, suggests that we
are on the right track by assuming that modalities differ with regard to their
selection restrictions. In the following sections, I will take a closer look at
desired states (attested), required states (unattested) and intentional states
(unattested) in order to determine why children denote the first in their early
sentences, but not the second and third.

4.4.2   Desired states

Children's use of states as well as events in volitional sentences is in
accordance with adult Dutch, as shown in the examples in (9):

(9)   a.   James en Donna willen gelukkig zijn
           James and Donna want happy be-*inf*
           'James and Donna want to be happy'
      b.   Bobby en Shelly willen trouwen
           Bobby and Shelly want marry-*inf*
           'Bobby and Shelly want to marry'

Thus, when children use states in volitional RIs, they do not do anything
remarkable; they do just what they are supposed to do.

4.4.3   Required states

A similar conclusion can be drawn for deontic RIs that express required
states and are used by children to give commands. In adult Dutch,
commands and states are less easily combined than desires with states.

Hence, the absence of states in commanding RIs is not unexpected. When states are inserted in a sentence that is formally marked for imperative mood, the sentence expresses a warning rather than a command (the prototypical function of imperatives):

(10) a. Heb plezier in het leven!
          have-*imp* joy in the life
          'Enjoy life!'
      b. Weet het antwoord op mijn vraag!
          know-*imp* the answer on my question
          'Know the answer to my question!'

The difficulty arises as an effect of incongruity between the function of a command and the lexical meaning of many state verbs. Bickerton (1981: 157) put this as follows: "It would be bizarre if he [i.e. speaker, EB] sought [...] to influence the thought processes and emotions of others by commanding them to want, need, know, etc."

In an attempt to formalise this intuition, I will propose that this inappropriateness comes from a conflict between two different specifications of the feature [±control]: commands require [+control] predicates, whereas most states are inherently specified for [-control]. Examples of [-control] states are *zwanger zijn* 'be pregnant' or *bestaan* 'exist'; [+control] states are exemplified by predicates like *stil zijn* 'be silent' or *blijven zitten* 'remain sitting'. The following contrast in an imperative/commanding sentence between a [+control] state, like *stil zijn* 'be quiet', and a [-control] state, such as *zwanger zijn* 'be pregnant', confirms the relevancy of the feature [±control]:

(11) a. Wees stil!
          be-*imp* quiet
          'Be quiet!'
      b. * Wees zwanger!
          be-*imp* pregnant
          'Be pregnant!'

For the purpose of trying to understand why children do not use any states in their commanding RIs, while they do talk about desired states in RIs, it is important to realise how much overlap there is between the set of states and

the set of [-control] predicates. Levin & Rappaport (1995:170-171) already pointed out that it is difficult to find a test that establishes the state-event contrast. They observed that most tests distinguish between agentive and non-agentive predicates, which is pertaining to the distinction between [+control] and [-control].[16,17] Thus, in practice, the claim that RIs are relatively infrequently stative is highly similar to the conclusion that [-control] predicates are seldom used in RIs.

In spite of the amount of overlap between the set of states and the set of non-agentive predicates, the example in (11) shows that there are some [+control] statives. Thus, theoretically speaking, children should be able to use stative [+control] verbs in their commanding RIs. Given that the set of [+control] statives contains only few elements, the chance that such RIs indeed occur is very small. Turning to the data of the six children, it can be observed that the children use two kinds of state verbs: verbs of position/location and verbs of possession. In commanding RIs, the stative [-control] reading of verbs of position/location (sit, lie, stand, etc.) is blocked. As a consequence, the [+control] eventive reading is triggered.

---

[16] According to Levin & Rappaport, the test from Dowty (1979) that only non-statives have the ability to appear in *do*-constructions and Jackendoff's (1983) test that only non-statives can be inserted in the frame *What happened/occurred/took place was ?* are the only tests that distinguish between states and events.

[17] Not all [-control] predicates are stative though: a predicate like *blush*, for instance, describes an event but is [-control]. Thus, although most states are [-control], there are [+control] states and [-control] events. Interestingly, when such a predicate is selected by the auxiliary *moeten* 'must', a third kind of modality (next to deontic and epistemic) surfaces. In (i), there is no subject-external source that requires an event or state, as with deontic modality (and with sentences that express through imperative mood, a command), but there is a subject-internal source. Barbiers (1995) called this kind of modality 'dispositional'. As expected (because of the negative specification for control), the imperative command is semantically not ill-formed.

(i)   De onderwijzer moet blozen
        the teacher must blush
        'It happened to the teacher that he had to blush'

(ii)   * Bloos nu!
        blush-*imp* now
        'Blush now!'

The verb of possession, *hebben* 'have', does not have a [+control] variant; hence, it is absent in commanding RIs. In their SFs, the children use various states that are not used in commanding RIs: mental verbs, modals like *kunnen* 'can', *moeten* 'must', *mogen* 'may' or *willen* 'want' and the copula *zijn* 'be'. These verbs are [-control] and therefore unexpected in commanding RIs. In sum, I conclude that like children's use of states in volitional RIs, the absence of states in commanding RIs is fully in accordance with properties of adult Dutch. It follows from the selection restrictions on commands that are in force for child as well as adult language.

The explanation I gave is inspired by, and therefore resembles, previous claims made by Ferdinand (1996) and Hoekstra & Hyams (1998). It also differs from these claims, however. Ferdinand and Hoekstra & Hyams do not consider the feature [±control] and the incompatibility of deontic RIs and stative predicates that I analysed as the effect of this feature. The gist of the explanation given by Ferdinand and Hoekstra & Hyams is that the deontic modal reading "switches" to an *epistemic* modal reading when a stative predicate appears in a modal RI. As young children cannot yet use epistemic modality, stative modal RIs do not occur. Ferdinand as well as Hoekstra & Hyams do not go into the question as to why epistemic modality is unavailable to children during the RI-period. As will be shown in section 4.5, there is empirical evidence from the field of developmental psychology in support of this claim. Hence, in this respect the proposal is tenable. To illustrate the relation between eventivity and deontic modality, on the one hand, and stativity and epistemic modality, on the other hand, Hoekstra & Hyams take a (more or less) minimal pair such as *leren* 'learn'-*weten* 'know'. These two verbs contrast in dynamicity (*weten* 'know' is stative), but they are, for the rest, semantically quite similar. The example in (12b) suggests that stativity of the selected complement triggers an epistemic modal reading of *moeten* 'must':

(12) a.   Jij moet Chinees leren   DEONTIC
            you must Chinese learn
            'It is required that you learn Chinese
      b.   Jij moet het antwoord weten   EPISTEMIC/DEONTIC
            you must the answer know
            'Based on everything I know, I conclude that you know the
            answer'
            'It is required that you know the answer'

Example (12b) shows that the deontic necessity reading is still accessible
when a state verb is selected; this is illustrated by the second paraphrase.
Thus, the explanation given by Ferdinand (1996) and Hoekstra & Hyams
(1998) does not really account for the absence of stative RIs. The analysis I
have given shows that stative predicates and commanding sentences are
semantically incompatible and ruled out, because states are specified as
[-control], whereas commands select for [+control] predicates.

Before discussing the last type of modal use, i.e. intentions, I will briefly
illustrate that opposite conclusions can be drawn on the basis of
explanations - mine, on the one hand, and the proposals from Ferdinand,
1996, and Hoekstra & Hyams, 1998, on the other hand - for the same set of
observations in child language, because these explanations are based on only
slightly different examples from adult language. Recall that my claim was
that children, by ruling out states in their commanding RIs, behave like
adults. This claim stresses that young children and adult speakers are
sensitive to the same selection restrictions. Ferdinand and Hoekstra &
Hyams explain children's ruling out of states in commanding RIs as an
effect of the absence of epistemic modality in early child language. The
implication of this analysis is that children and adults differ, because
children do not use epistemic modality (for reasons that are not further
specified in the two studies), whereas adults do. I supported my claim with
imperative sentences from adult Dutch; these sentences show that
commands cannot select for a predicate that is [-control]. Ferdinand and
Hoekstra & Hyams used *declarative* sentences with deontic modal auxiliaries
to illustrate their viewpoint. In (13), the declarative counterparts of my
imperative examples in (11) are given. In (11b), [-control] states are simply
ruled out, whereas in (13b), [-control] are allowed but with an epistemic
modal reading:

(13) a. De kinderen moeten stil zijn   DEONTIC (*declarative*)
          the children must silent be
          'The children have to be silent'
          'It must be the case that the children are silent'[18]
      b. De vrouw van de onderwijzer moet zwanger zijn

                                          EPISTEMIC (*declarative*)

          the wife of the teacher must pregnant be
          'The teacher's wife must be pregnant'

In order to explain the very marginal appearance of states in modal RIs, Ferdinand and Hoekstra & Hyams assume that the epistemic reading is not available to children (it is to adults, though). This explanation is superfluous when examples like in (11) are used to paraphrase the modal denotation of RIs. Note that Ferdinand *has to* use the declarative examples with modal auxiliaries, since she claims that RIs contain an underlying silent modal auxiliary. Hoekstra & Hyams are not forced to use this particular example, however.

4.4.4   Intended states

Up to now, children's use and non-use of states in RIs is fully in accordance with what adult speakers do in utterances that are semantically comparable to child RIs: it is expected that volitional RIs contain states, whereas commanding RIs do not. I will now turn to the third kind of modality that Dutch children use in their RIs, that is, intentional modality. This denotation is illustrated in (14):

---

[18] In Dutch, an epistemic reading is reinforced if the particles *haast wel* are added. These particles are hardly translatable but they stress the speaker-orientation of epistemic modality and trigger the epistemic 'It must be the case' reading:

(i)   De kinderen moeten haast wel stil zijn
          the children must *particle particle* silent be

(14)   a.   Ik ga Chinees leren   FUTURE/INTENTION
            I go Chinese learn
            'I am going to learn Chinese'
       b.   Ik ga het antwoord weten   FUTURE/EPISTEMIC
            I go the answer know
            'I shall know the answer'

In Chapter 3 (§ 3.2.2) epistemic modality was described as speaker-oriented modality that refers to a speaker's beliefs or judgements (Lyons, 1977; Palmer, 1986). Given this definition, (14b) denotes epistemic modality: the speaker expresses a strong belief (*I am sure about it/I reassure you*) and emphasises his certainty about the possible truth of the proposition KNOW(I, the answer). The epistemic modal reading is not necessarily present in (14a), although this reading can be triggered by stressing *ga* 'go'. The feature [±control] does not play a role: the examples in (15) contain a [+control] and a [-control] state, respectively, and both receive the same epistemic modal interpretation:

(15)   a.   De kinderen gaan stil zijn
            the children go silent be-*inf*
            'The children are going to be silent'
       b.   De vrouw van de onderwijzer gaat zwanger zijn
            the wife of the teacher go pregnant be-*inf*
            'The teacher's wife is going to be pregnant'

If the claim that young children (that is, children during the RI-period) cannot use epistemic modality is true, RIs that denote intended states are not expected to occur. Young children's knowledge of epistemic modality will be examined in section 4.5. Taking the conclusions from this section for granted, I assume that epistemic modality is indeed not available to children in the RI-age. The implication of the analysis I have given is that the explanation proposed by Ferdinand and Hoekstra & Hyams does apply to intentional RIs, but not to commanding RIs.

4.4.5 Summary

The central tenet of the Modal Approach is that modal meanings and aspectual properties of the verbal complement interact. More specifically, the claim is made that stative predicates trigger modal meanings that cannot yet be used by children in the 'RI-age'. In this section, I studied this interaction in order to investigate whether or not the Modal Approach can capture the observation that the few stative RIs that are used in modal RIs, are restricted to volitional RIs. Table 4.10 summarises the observations. The bold-faced interpretations refer to the three modalities that appear in the RIs of Dutch children. When the interpretation is preceded by an asterisk (*), this means that the interpretation is ruled out given the semantic specification of the selected verb type. States are [-dynamic] and, almost always, [-control]. The selecting verb is the auxiliary that Dutch adults use to paraphrase the modal meanings expressed in RIs.

Table 4.10: Relations between modal use and restrictions on the verb type of the selected verbal predicate

Auxiliary	Verbal complement	Resulting interpretation
Moet 'must'	[-control]	Epistemic necessity (speaker certainty) *** Command (requirement, deontic necessity)**
Ga 'is going to'	[-dynamic]	Epistemic necessity (speaker certainty) *** Intention**
Wil 'want'	[-control]/ [-dynamic]	**Wishes and desires (volition, dynamic necessity)**

On the basis of this set of observations, the conclusion can be drawn that the use of states in volitional RIs is fully expected given the MA, that is, the Modal Approach to the eventivity of RIs. It is also expected that states do not occur in commanding RIs. Finally, provided that children in the RI-age

are unable to use epistemic modality, it is expected that states do not appear in RIs that denote intentions.

4.5   The effect of cognitive immaturity

It could be that the absence of epistemic modality in child RIs is accidental and, therefore, meaningless. However, there are a number of reasons to believe that this is not the case and that the cognitive immaturity of children in the RI-age prevents them from using epistemic modality. I will show that epistemic modal utterances are absent from young children's production, as well as comprehension data, across the board, and that there is, in this respect, an asymmetry between dynamic and deontic modality, on the one hand, and epistemic modality, on the other. Dynamic and deontic modality appear before the age of three, epistemic modality appears only after the age of three.

Various studies on English child language report that modal verbs in English child language come in between the ages of two and three (Kuczaj & Maratsos, 1975; Shepard, 1982; Stephany, 1986; Gerhardt, 1991). Stephany (1986) reported the forms *can* and *cannot* as the first to appear. Hoekstra & Jordens (1994) observed the modal forms *magnie* 'may not', *kannie* 'cannot', *nee* 'want not', *minne* 'want' and *wil* 'want' in the data of a Dutch-speaking girl around the age of two (1;10 and 1;11). These early 'can' and 'want' modals can be classified as dynamic (referring to a subject internal possibility or necessity), whereas the early 'may' modal is deontic (referring to a subject external possibility, a permission from 'outside'). Epistemic modality is not yet used. Schatz, Wellman & Silber (1983), Bartsch & Wellman (cited in Wellman, 1990) and Becker (1998) discovered that children do not refer to epistemic modal concepts such as belief or knowledge before their third birthday. Kuczaj & Maratsos (1975) reported that the first epistemic usages of modals in child English appear during the fourth year. Comprehension tasks carried out by Hirst & Weil (1982) showed that the understanding of epistemic certainty by English-speaking children also begins during the fourth year. These expressions begin to appear when the children are between three and four years old.

In an attempt to investigate the relation between preschoolers' under-
standing of modal verb meaning with their non-linguistic understanding of
modality, Gonsalves (1998) compared the production and comprehension
of preschoolers' modal verbs to their scores on a non-linguistic search task.
Gonsalves included children between 2;6 and 5 years and divided this
population into three age groups. The comprehension test, which was a
truth value judgement task, probed the preschoolers' understanding of *can*
(possibility), *cannot* (impossibility), *has to* (necessity) and *doesn't have to* (non-
necessity) across three context domains: deontic, dynamic and epistemic
modality. Gonsalves collected speech samples from child-caregiver
conversation. She examined whether or not, and with which frequency, the
children used deontic, dynamic and epistemic modals. The results from
these two linguistic tasks were compared to the results of a non-linguistic
task with possible and necessary hiding places for a toy. In the test
condition, there was either sufficient evidence for the hiding place (necessity
or impossibility) or insufficient evidence (possibility or non-necessity). The
aim of this non-linguistic task was to find out whether or not the child had
the cognitive abilities to differentiate between determinate or indeterminate
conclusions, hence to find out if the child has a non-linguistic grasp of
modal concepts. Gonsalves found that not a single child that was able to
distinguish epistemic 'can be true' from 'has to be true' in the
comprehension test, failed on the non-linguistic task. At the same time, all
children that failed on the non-linguistic test and were not able to
distinguish possibility from necessity also failed on the epistemic
comprehension test. However, these children did not fail on the deontic and
dynamic condition in the comprehension test and deontic and dynamic
modals were found in production data of these children. The conclusion can
be drawn that epistemic modality requires a full-fledged grasp of the modal
concepts possibility, necessity, impossibility and non-necessity, whereas
dynamic and deontic modality do not.

All the above results show that epistemic modality is systematically absent in
language production as well as comprehension of children in the RI-age. In
studies that concentrate on children's cognitive development, cognitive
immaturity is explained by the absence of a Theory of Mind (cf. Wimmer &
Perner, 1983; Carey, 1985; Wellman, 1990). Because young children lack a
concept of the mind, they neither talk about or understand their own mental
states nor talk about or understand those of others. As epistemic modality

concerns knowledge of beliefs, hence mental states, children will not use epistemic modality before they have a Theory of Mind. In the literature on this topic, the emergence of a Theory of Mind is described as a sudden change that takes place halfway during the third year. This claim is substantiated by responses of three and four year olds on the so-called false belief task.[19] Note that mental states like beliefs are distinct from mental states such as intentions and desires. In order to describe the differences between beliefs and intentions/desires Gopnik (1993) makes use of Searle's (1983) distinction between a "mind-to-world" and a "world-to-mind direction-to-fit" in the case of a "mind-to-world direction-to-fit", the mind is altered to fit the world, whereas in case of a "world-to-mind direction-to-fit", the world is altered to fit the mind. Beliefs are examples of the first, whereas intentions/desires exemplify the latter. There is evidence that young children perform better on tasks that involve "world-to-mind directions-to-fit" such as intentions and desires than on experimental tasks that require a concept of "mind-to-world directions-to-fit" such as belief (Flavell, Flavell, Green & Moses, 1990; Wellman & Woolley; 1990; Astington & Gopnik, 1991). Flavell et. al (1990) found for instance that three-year-olds perform better on desire tasks than on belief-tasks, but the proportions of errors that occurred when talking about the desires of others were still considerable: between 30 % and 40 %. According to Gopnik (1993), children between two and three do not have a full grasp of intentions and desires and their concept of these mental states is simple and non-representational. As young children lack internal representations of intentions and desires, it is expected that when these children talk about desires and intentions, the desired object or action is present in the speech situation (and triggered the intention or desire). This property of early desires and intentions is reflected in data from Abel, Daan, Josse, Laura,

---

[19] A simple false belief task as developed by Baron - Cohen (1985) works as follows. There are two puppets, Sally and Anne. Sally has a marble, which she keeps in a basket. Then Sally leaves the room, and while she is away Anne takes the marble out of the basket and hides it in the box. Sally comes back into the room. The child subject is then asked the question: "Where will Sally look for her marble?" Older children say that she will look in the basket, because although they know the marble is in the box, they know that Sally doesn't know it has been moved from the basket, and they can distinguish Sally's (false) belief from their own (true) belief. Younger children do not distinguish between the two and simply say that Sally will look in the box.

Matthijs and Peter in so far that the children almost always use the desiderative predicate *willen* to talk about their own desires in the here-and-now. Hardly any reference is made to the desires of others or to past desires.

Table 4.11: *Willen* 'want', first person singular subjects and temporal reference, data from all six children

	N_{WILLEN}	$N_{1ST PERS. SG.}$ (%)		$N_{PRESENT}$ (%)	
Abel	58	49	(84)	58	(100)
Daan	32	31	(97)	32	(100)
Josse	33	33	(100)	33	(100)
Laura	86	83	(97)	85	(99)
Matthijs	13	9	(69)	13	(100)
Peter	6	5	(100)	5	(83)

Astington & Gopnik (1991) furthermore found that three-year-olds had difficulty in understanding their own past desires. In an experimental setting, they presented children with a situation in which their desires were satiated over time. The aim of the task was to cause a change in the children's desires. One third of the three-year-olds, in response to a question, did not reflect on this change and they answered that they had been in their final (satiated) state all along. Bringing this discussion on the impact of children's cognitive immaturity to an end, I will close with an observation on the basis of the data from the six children. The asymmetry between epistemic modality (beliefs) and intentions/desires can be observed in the data of Abel, Daan, Josse, Laura, Matthijs and Peter: their modal RIs are never epistemic, whereas they are often volitional and used to express intentions and desires. [20]

---

[20] Interestingly, possibility readings are virtually absent from modal RIs, although verbs like *kan* 'can' and *mag* 'may' appear very early in SFs alongside RIs and appear some months after the first RIs in the first PVs (*kan dansen* 'can dance', *mag dansen* 'may dance'); see Chapter 5 for the relevant data. It is, as far as I know, still an open question why the possibility readings are systematically lacking from RIs.

4.6 Verb type in English RIs

By proposing the Modal Approach to account for the EC, I followed
Ferdinand (1996) and Hoekstra & Hyams (1998). As discussed in the
previous chapter, Hoekstra & Hyams propose the Modal Approach not only
to explain the EC in Dutch RIs, but also to account for differences between
Dutch and English RIs with regard to verb type selection. However, even
though I share the opinion with Hoekstra & Hyams that there is a close
relation between modality and verb type, I do not follow their explanation
for semantic asymmetries between Dutch and English RIs. In the light of
the present chapter, I briefly summarise my analysis for cross-linguistic
differences with regard to verb types that appear in RIs.

Hoekstra & Hyams claimed that the modal interpretation of Dutch RIs is
carried by the infinitival morpheme [-*en*]. Verb forms in English RIs do not
possess this morpheme, and hence, they are not restricted to a modal
interpretation, according to Hoekstra & Hyams. On the basis of the
observation that about one quarter of the Dutch RIs was non-modal, I
concluded in Chapter 3 that the denotation of RIs is not fixed and hence,
that infinitival morphology is not reponsible for the interpretation assigned
to RIs. This conclusion implied that morphological differences in verb form
do not explain differences in meaning between Dutch and English RIs. I
argued that the relative frequent use of RIs with non-modal (e.g. present
tense and, to a lesser degree, past tense) interpretations by English children
is captured by the Heterogeneous Set Effect (HSE): the set of English RIs,
as opposed to the set of Dutch RIs, contains finite sentences with dropped
inflection. Given the HSE, the observed asymmetry in meaning between
Dutch and English RIs is to be expected. In this chapter, we have seen that
finite verbs in early Dutch child language are more often stative than
infinitives. Expanding this observation to English child language, it can be
concluded that the HSE accounts for the observed asymmetry between the
verb types in Dutch and English RIs as well: since the set of English RIs
contains finite verbs, it is expected that English RIs contain relatively many
state-denoting predicates.

Both the explanation defended in this thesis as well as the explanation given
by Hoekstra & Hyams, relate semantic differences (i.e. temporal/modal
denotation and aspectual type of the verb) to morphological properties of

the verb form used in RIs in the two languages. However, my analysis includes a second difference, namely verb placement. In English, the HSE occurs, and finite forms with dropped inflection are indistinguishable from non-finite RIs, because of: (i) English does not have a true infinitive, and (ii) finite verb forms and non-finite verb forms are placed in the same position in English. Both properties do not apply to Dutch. By implication, the HSE does not occur in Dutch and RIs in Dutch child language are not only more often modal than RIs in English children, but also more frequently eventive.

4.7 The role of the input

I argued that the EC is due to two factors: children's early knowledge of selection restrictions explains the absence of stative predicates in commanding RIs and children's cognitive immaturity accounts for the absence of stative predicates in RIs that denote intentions. There is a third factor involved in the EC in early child Dutch: patterns in the input.

In section 4.3.3, distributions of eventive and stative predicates in adult Dutch were discussed. It was shown that patterns in child and adult Dutch resemble each other, in the sense that both display an Eventivity Constraint (EC). Given that the language children hear is one of the factors that determine patterns in language children produce, it is expected that the EC in adult Dutch also has effect on child Dutch. More specifically, it could be argued that the EC in the data from Abel, Daan, Josse, Laura, Matthijs and Peter is not only similar to the EC in the language of the children's caregivers, but also that it *follows* from the EC in the caregivers' speech.[21]

The six children hear hardly any states in infinitival form, whereas they hear numerous states in finite form. Given this observation, the chance that the children pick up states in infinitival form must be fairly low. The only way in which the states-in-infinitival-form can become available to the children is by deriving them from the finite forms they know. However, the children

---

[21] Similar claims are made by Schlichting (1996) for Dutch and by Pine, Lieven & Rowland (1998) for English.

can only do this when they know the rules for segmentation and suffixation. In Chapter 5, I will go over a number of indicators for the acquisition of segmentation and suffixation. It will be concluded that Dutch children acquire this capacity only about halfway through the RI-period. Thus, prior to this moment, children are fully dependent on the forms they pick up from the input. This implies that infinitival forms are predominantly eventive, whereas finite forms are relatively often stative verbs.

4.8 Summary

De Haan (1987) observed that children tend to use different verb types in RIs and finite sentences. In the literature, this observation has been explained in various ways. Evaluating these hypotheses, I examined the data from Abel, Daan, Josse, Laura, Matthijs and Peter and concluded that their RIs obey the Eventivity Constraint (EC). However, after considering the state-denoting predicates that the children also used in their RIs, this EC was found to be violable. I argued that there were three factors that cause the EC in early child Dutch. In the first stages of children's language development, the EC is the effect of patterns in the input. A second factor that contributes to the EC is that children are unable to use epistemic modality, because of their cognitive immaturity. This leads to the absence of RIs that denote intended states. The effect of this second factor will also disappear when the children grow older. By then, the RI-period is expected to have ended, however. The third factor that causes the EC will not disappear as a function of grammatical development or cognitive maturation: the inappropriateness of stative predicates in commanding sentences leads to the EC in Dutch child language as well as Dutch adult language. I argued that the violable character of the EC provided support for the Modal Approach to the EC, which is constituted by the second and third factor. The Modal Approach correctly predicts that a combination of (i) children's knowledge of modality (i.e. deontic modality and volition), and (ii) deficits in their knowledge of modality (i.e. epistemic modality) leads to the restriction of stative predicates to volitional RIs.

CHAPTER 5

Changes over Time

In the previous two chapters, data that were collected over a longer period of time were collapsed. By doing this, I followed the common method in RI research and abstracted away from developments that take place during the period in which children use RIs more frequently than adults do. In this chapter, I will look at the data from Abel, Daan, Josse, Laura, Matthijs and Peter from a different viewpoint: the RI-period of the six children will be sliced into smaller chunks, more specifically, into four stages, in order to give a precise description of the changes over time. It will be shown that Dutch children acquire finite sentences (FINs) step-by-step: absence of finiteness is followed by a stage in which finiteness is a lexical property. The lexical stage is followed by a grammatical stage in which the morpho-syntactic properties of finiteness become available. I argue that the acquisition of FINs has an impact on children's semantic, morphological and syntactic development. Finite forms pave the way for verb movement and rule-governed subject use. The acquisition of FINs has an effect on the semantic and structural properties of RIs. In this chapter, an alternative for the modal hypotheses discussed in Chapter 3 will be proposed. According to this proposal the Modal Reference Effect does not follow from a silent modal auxiliary or the meaning of infinitival morphology, but results from the rise of FINs. The data presented in this fifth chapter support the idea that children have to unravel the information that guides them to grammatical knowledge. They do this in a piece-meal fashion through analysis and re-analysis of words and phrases on the basis of newly acquired lexical information.

5.1   Outline of this chapter

Although Dutch children use infinitives very early, they do not learn the properties of infinitives until they have access to contrasting finite verb forms. This hypothesis enables us to make predictions for changes in RIs based on the development of FINs. The presumed correlation between FINs and RIs implies that there are (at least) two ways to organise this chapter. One way is to start by looking for changes in RIs that happen over time and then make predictions for the development of FINs that are based on the observations within RIs. Another way is to begin by describing the rise of FINs and derive predictions for RIs on the basis of the developmental patterning of FINs. Since the final conclusion is that changes in FINs cause changes in RIs, the order seemed the more logical of the two. Therefore, I will start with the rise of FINs. The acquisition of FINs will be discussed in section 5.2. Section 5.3 deals with the effect of the acquisition of FINs on the meaning assigned to RIs. Section 5.4 is about the effects of the acquisition of FINs on subject use. Theoretically, changes in meaning as well as in subject use in RIs can be explained from a full competence perspective (that is, the Agreement and Tense Omission Model that was proposed by Wexler & Schütze, 1996 and Schütze, 1997). In section 5.5, I will compare the latter explanation to the interpretation of the changes that I propose. Section 5.6 contains a note on the methodology that is applied here, whereas 5.7 summarises the findings and conclusions of this chapter.

5.2   The rise of FINs

By investigating changes over time, I focus on a topic that has been ignored by most studies on RIs, namely the actual development from RI-stage to adult finite stage. Taking a brief look at the hypotheses that deal with the absence of finiteness in child language - more details can be found in Chapter 1, section 1.1 -, it turns out that strikingly little is said about the developmental path of the acquisition of finiteness. According to maturational version of the RCH, children's use of RIs reveals absence of functional categories (Lebeaux, 1988; Radford, 1988, 1990). Children are not immediately ready to project, for instance, the functional projection FinP (or

TP, under the assumption that finiteness is tense). Instead, FinP matures. According to the FCH, children have abstract knowledge of finiteness from early on, but this knowledge does not surface when children do not know the instantiations of finiteness (Poeppel & Wexler, 1993).[1] By concentrating on the abstract level, assuming a process like maturation or by making a distinction between instantiations and abstract representations, the actual development from RIs to FINs becomes unessential for both the maturational RCH and the FCH.[2]

This section will show that the examination of the small steps that children take is a fruitful research strategy. It will be argued that the 'transition' from a system that produces RIs to a system that produces finite sentences, consists of a series of smaller steps, and, hence that there is no all at once transition. Via these smaller steps, children acquire the abstract grammatical finiteness features of Dutch: inflection and, ultimately, verb movement. To argue for this claim, I will first discuss the different properties of FINs in adult Dutch (§ 5.2.1), then I turn to the sequence of development of FINs in early child Dutch (§ 5.2.2 -§ 5.2.3), followed by an examination of the data of Abel, Daan, Josse, Laura, Matthijs and Peter (§ 5.2.4 - § 5.2.10).

5.2.1 Properties of FINs in the target language

FINs comprise different kinds of sentences, but I will restrict myself to the development of FINs that stand in relation with RIs (and that are expected to affect children's use of RIs, given that RIs are the focus in the present

---

[1] When the maturational version of the RCH and the FCH are compared on a very basic level, it quickly becomes obvious that in the studies that defend the former hypothesis, RIs are used to illustrate that children do not have Fin_{AL}, where 'Fin' stands for the functional finiteness category and the index $_{AL}$ stands for adult language. The FCH studies concentrate on syntactic properties of the earliest FINs used alongside RIs and on syntactic differences between FINs and RIs in order to show that $Fin_{CL}= Fin_{AL}$ - where the index $_{CL}$ stands for child language - even when Fin_{CL} does not surface. Considering the stages I-IV as observed by Wijnen (2000) and in the present study, it also becomes clear that the maturational version of the RCH and the FCH emphasise different stages.

[2] Either of the two models may fit developmental findings, but the models do not make any predictions for the sequence of development themselves (apart from some major transitions from child RIs stage to adult finite stage or child OI stage to adult finite stage).

work). These are finite alternatives to RIs, which are PVs with infinitives and SFs. This set of FINs, which excludes PVs with participles, express the same modal, temporal and aspectual meanings that are expressed with RIs, that is, all tenses, modal as well as non-modal meanings and imperfective aspect. PVs with a participle are excluded because they are not an alternative for RIs, as RIs do not denote completed aspect. For reasons given in Chapter 3, § 3.3.6, no distinction is made between genuine PVs and pre-PICs, i.e. children's precursors of prepositional infinitival complements.

The finiteness markers in Dutch children's FINs are either attached to auxiliaries (in PVs) or inflected main verbs (in SFs). In Chapter 3, I explained various properties of the Dutch auxiliary and inflectional system. In this section, I summarise the main properties of these systems. SFs and PVs differ syntactically and semantically from each other. In SFs, the main verb carries inflectional morphology and is moved out of its sentence-final base position to first or second position. This is illustrated in (1):[3]

(1)   [CP Jan$_i$ loopt$_j$ [IP t$_i$ t$_j$ [VP de Vierdaagse van Nijmegen t$_j$]]]
        Jan walk-*fin* the Vierdaagse van Nijmegen
        'John is walking the Vierdaagse van Nijmegen'[4]

In PVs, the lexical verb is still in its base position and the auxiliary takes first or second position:

(2)   a.   [CP Jan$_i$ **moet**$_j$ [IP t$_i$ t$_j$ [VP de Vierdaagse van Nijmegen **lopen**]]]
                Jan must the Vierdaagse van Nijmegen walk-*inf*
                'John must walk the Vierdaagse van Nijmegen'
        b.   [CP Jan$_i$ **is**$_j$ [IP t$_i$ t$_j$ [VP de Vierdaagse van Nijmegen **aan het lopen**]]]
                Jan is the Vierdaagse van Nijmegen on the walk-*inf*
                'John is walking the Vierdaagse van Nijmegen'

---

[3] I have given here a more or less traditional representation that does not contain a split IP. With this representation, I do not want to make any claims about the exact representation. My aim is to show that the main verb moves.

[4] This *Vierdaagse* 'four-day-' is a yearly event in The Netherlands at which lots of people walk a certain distance for four subsequent days. Particularly popular is the *Vierdaagse* that takes place in the area surrounding the city of Nijmegen.

Morphologically, there are differences between (1) and (2) as well. The main verb in (1) has a finite suffix that marks tense and agreement, whereas the main verb in (2) has an infinitival morpheme with no tense or agreement specifications. The inflectional markings of the finite verbs in (1) and (2) are different because the auxiliaries used in PVs usually have impoverished or irregular inflection as compared to main verbs used in finite SFs. Apart from agreement, regular inflection in Dutch marks a past-present tense distinction. As described in Chapter 3, tense inflection has two forms in Dutch. The form that is structurally represented by a zero or null suffix [*-ø*] primarily denotes present tense but can also be used with a future reference.[5] Regular past tense is marked on the verbal stem through a suffix [*-de*] or [*-te*]. The present tense agreement paradigm encodes a number distinction (singular vs. plural). In the singular forms, a distinction is made between first person singular, on the one hand, and second/third person singular, on the other hand. The first person singular is similar to the bare verbal stem. Here, I focus on morphological differences between regular inflection and auxiliaries. In § 5.4.5, the properties of agreement in Dutch will be discussed in more detail. In (3), the present tense paradigm of the verb *lopen* 'walk' is given:

(3) | | SG | PL |
|---|---|---|
| 1st | loop**ø** | lop**en** |
| 2nd | loop**t** | lop**en** |
| 3rd | loop**t** | lop**en** |

With respect to auxiliaries, it can be observed that many auxiliaries lack the overt distinction between first person singular and second/third person singular. This is illustrated in (4) with the present tense paradigm of the modal auxiliary *moeten* 'must':

---

[5] In context, the present tense form can also be used with a future tense reference:

(i)   Over dertig jaar loop ik naar Santiago de Compostella
        over thirty years walk I to Santiago de Compostella
        'In about thirty years, I will walk to Santiago de Compostella'

(4)   SG   PL
    1st   moet   moet**en**
    2nd   moet   moet**en**
    3rd   moet   moet**en**

Moeten 'must' is not an exception: the modal auxiliaries *kunnen* 'can/be able', *mogen* 'may/be allowed', *willen* 'want', *zullen* 'will' show a similar pattern. *Zijn* 'be' and *hebben* 'have' have their own, idiosyncratic, irregular paradigms. The auxiliaries *hoeven (niet)* 'must (not)' and *gaan* 'are going to' are exceptional as their agreement paradigms are regular.

Table 5.1 provides a summary of the properties of SFs and PVs with infinitives:

Table 5.1: Syntactic, morphological, and semantic, properties of SFs and of PVs in adult Dutch

	Placement of main verb	Morphology of main verb	Meaning
SFs	initial ('moved')	Finite	past, present, future tense
PVs	final ('unmoved')	Infinitival	modality, imperfective aspect

The apparent differences between SFs and PVs in adult Dutch lead to specific expectations with regard to the sequence of development of FINs. In the next section, I describe the acquisition of FINs as the result of an incremental learning procedure. The hypothesis that will be proposed is based on a number of findings reported in previous studies. It is in line with ideas expressed by several other scholars (cf. De Haan, 1987; Hoekstra & Jordens, 1994; Van Kampen, 1997; Wijnen & Elbers, 1998; Wijnen, 1999, 2000).

5.2.2 The sequence of development

From the studies carried out by Wijnen & Bol (1993) and Haegeman (1995), we know that Dutch-speaking children start out their 'verbal career' with RIs and that FINs (i.e. SFs and PVs) come in later. This means that the 'Optional Infinitive stage', that is, the stage in which children use FINs and RIs at the same time is preceded by a stage in which infinitives are not optional, but obligatory. The absence of FINs and the sole use of RIs in the earliest stages suggest that the knowledge to use finite sentences is lacking entirely in these stages. The absence of finiteness in RIs is marked by morphological as well as syntactic properties: the verb carries infinitival morphology and is not moved. Findings from § 3.3 and § 3.4 indicate that the semantics of RIs are compatible with their structural properties, since they have no fixed temporal reference. The evolvement of the semantic properties of FINs will be discussed in § 5.3. The present section focuses on the sequence of development of morpho-syntactic properties of finiteness.

This sequence of development begins with RIs in which the verb is non-finite and unmoved and ends when children show productive use of SFs, i.e. sentences that contain a finite moved verb. Begin and end are exemplified in (5) and (6), respectively:

(5)   Ik de toren **bouwen**
        I the tower build-*inf*

(6)   Ik **bouw** de toren
        I build-*fin* the tower

The various differences between these two sentences suggest that the development from a system that generates sentences such as (5) to a system generating sentences such as (6) involves more than one step: the main verb is placed in different positions (sentence-finally versus sentence-initially), the morphology of the main verb differs ([*-en*] versus [*-ø*]) and the meaning of (5) and (6) is different (untensed versus specified for present tense).

A number of observations from early child Dutch (that will be discussed below) are consistent with the hypothesis that knowledge from the morphological and syntactic properties of finite verbs is gradually build up

(De Haan, 1987; Jordens, 1990; Van Kampen, 1997; Wijnen, 1999, 2000). More specifically, the observations suggest that Dutch children first pass a lexical-finiteness-stage before they access the adult-like grammatical-finiteness-stage. This sequence of development of finiteness is stated in (7) as the Grammaticalisation of Finiteness (GoF) Hypothesis:[6]

(7)   *Grammaticalisation of Finiteness (GoF) Hypothesis*
        In early child Dutch, the grammatical marking of finiteness by means of inflection and verb movement is preceded by a lexical-finiteness stage.

The observations furthermore indicate that the lexical-finiteness-stage consists of two substages: a stage in which (i) lexical finiteness markers come in, and (ii) lexical finiteness markers are combined with infinitives. The stages are exemplified in (8) and (9), respectively.

(8)   a.   Ik **wil** een toren
               I want-*fin* a tower
        b.   Dat **is** een toren
               That is-*fin* a tower

(9)   a.   Ik **wil** een toren **bouwen**
               I want-*fin* a tower build-*inf*
        b.   Ik **ga** een toren **bouwen**
               I go-*fin* a tower build-*inf*

These observations lead to further specification of (7). The revised GoF Hypothesis, that includes the two substages of the lexical-finiteness-stage, is given in (10):

(10)   *Grammaticalisation of Finiteness (GoF) Hypothesis (final)*
        (i)   In early child Dutch, the grammatical marking of finiteness by means of inflection and verb movement is preceded by a lexical-finiteness-stage.

---

[6] The GoF-Hypothesis is in line with ideas expressed previously by De Haan (1987) and Wijnen (1999, 2000). The idea that a lexical stage precedes a grammatical stage can also be found in the work of Roeper (1996) and Powers (1998).

(ii) In the lexical-finiteness-stage, simple lexical finiteness markers precede complex lexical finiteness markers.

Support for the existence of a lexical-finiteness-stage comes from the observation that the (earliest) SFs that appear alongside RIs contain verbs that are fundamentally different to RIs. More specifically, the earliest finite verbs are auxiliary-like forms that denote tense and modality (De Haan, 1987; Van Kampen, 1997; Wijnen, 2000). These are illustrated in (8). Wijnen & Elbers (1998) pointed out that the early no-overlap stage (which I interpreted as the first substage of the lexical-finiteness-stage) precedes the appearance of PVs (which I argued to be second substage of the lexical-finiteness-stage), illustrated in (9).[7] Wijnen & Elbers take the acquisition of these early SFs as a step towards the acquisition of PVs:

"[…] at this point in the developmental sequence, their [i.e. periphrastic verbs, EB] perceptual salience and segmentability is boosted by the fact that they are now known and used as independent predicates. The result is that the periphrastic predicate structure is added to the child's repertoire" (p. 19)

In the citation above, Wijnen & Elbers refer to the notion of *segmentability*. In the next section, I will emphasise the importance of segmentability for the acquisition of grammatical finiteness.

Van Kampen (1997) emphasised the *overuse* of PVs by Dutch children. This overuse shows in the way in which children use certain PVs. The PVs contain a finite *doen* 'do' or *gaan* that selects an infinitive; some examples are given in (11):

---

[7] This is a re-interpretation of De Haan (1987). De Haan did not distinguish between different stages within the RI-period, so, implicitly, he seemed to assume that the no-overlap generalisation applies to the entire RI-period. Unlike De Haan, Wijnen & Elbers argue that the no-overlap generalisation applies to a stage within the entire RI-period, however.

(11)   a.   Lucy **doet** een toren **bouwen**
             Lucy does-*fin* a tower build-*inf*
             'Lucy is building a tower'
       b.   Andy **gaat** een toren **bouwen**
             Andy goes-*fin* a tower build-*inf*
             'Andy is building a tower'

Formally, these PVs look like 'normal' adult PVs, but semantically they are different from PVs in adult Dutch in the sense that they denote present tense like SFs do.[8] According to Van Kampen (1997), children use these PVs as a strategy to avoid verb movement. The motivation for this avoidance is that it is 'cheaper' for a child to insert a finiteness marker than to move the verb in order to mark finiteness, according to Van Kampen. My proposal is different. I will argue in § 5.2.3 that the overuse of PVs is a step that *mediates* between the exclusive use of RIs and the productive use of SFs. Children overuse PVs as long as they lack knowledge of the grammatical marking of finiteness by verbal inflections and, consequently, verb movement.[9]

In this section, I hypothesised that Dutch children learning finiteness, develop from a stage that is characterised by the absence of finiteness (RI) via a lexical-finiteness-stage (SF$_{lex}$ and PV$_{lex}$, respectively) to a grammatical-finiteness-stage (SF$_{gram}$). Further motivation for the GoF Hypothesis will be given in the following section. Wijnen, Kempen & Gillis (2001) already discussed the reasons for Dutch children to begin with infinitives. They point out that a complex of factors - frequency, saliency, information load and semantic transparency - lets Dutch children pick up infinitives early (see Chapter 1, section 1.7). I focus on the following two (related) questions: (i) Why does the lexical-finiteness-stage precede the grammatical-finiteness-stage in early child Dutch, and (ii) Why does it take Dutch children relatively long to learn inflection?

---

[8] Similar observations for Dutch child language are reported by Schaerlakens & Gillis, 1987; Verhulst-Schlichting, 1985; Jordens, 1990; Evers & Van Kampen, 1995; Hollebrandse & Roeper, 1996; Zuckerman 2001).

[9] Van Kampen's analysis, however, could apply to a somewhat later stage than the stage I discuss, that is, when children do have the knowledge to use the verb but still display overuse of PVs.

5.2.3 Motivation for the GoF (Grammaticalisation of Finiteness Hypothesis)

In order to answer the questions above, I begin by determining the knowledge that is required in order to be able to move the verb. Since verbal inflections are identified as the trigger for verb movement, I will, subsequently, discuss the knowledge children require for learning verbal inflections.

In generative theory, verb movement is analysed in different ways. In minimalist terms, the verb has to move in order to check its finiteness features to the corresponding functional head (cf. Chomsky, 1993; Zwart, 1993, 1997 for verb movement in Dutch). The required functional structure is simply there in order to provide an appropriate checking environment. There is another approach to the movement operation that takes an economical view on representations and is compatible with the idea that children build up grammatical representations, i.e. the structure building approach (Grimshaw, 1994).[10] According to this view on verb movement, the inflected verb is inserted in V and has to move outside the VP in order to *project (functional) structure*. Movement is motivated as a structure building operation so that the finite (more specifically, tensed) verb can be in a position where it has scope over the entire proposition, including the subject. An analysis of verb movement along these lines can be found in Koeneman (2000).[11]

---

[10] This is in line with the Lexical Learning Hypothesis (Clahsen, 1990; Clahsen & Penke, 1992; Clahsen, Eissenbeiss & Penke, 1993) in the sense that the acquisition of inflection triggers the appearance of functional structure. Clahsen, Eissenbeiss & Penke (1993: 133) wrote:

"The idea is that functional categories such as IP, AgrP, etc. or syntactic features may come into the child's phrase-structure representations as a consequence of the child's learning a regular inflectional paradigm of distinct inflectional affixes."

In addition, however, inflection also introduces verb movement.

[11] Koeneman (2000) argues that verb movement is triggered by tense and/or agreement features. The claim is that verb movement for reasons of tense is universally triggered, but verb movement for reasons of agreement is dependent on richness of the agreement paradigm. Only when agreement is rich, agreement can function as a full DP argument and

Adopting the structure building view on verb movement, I conclude the there are three prerequisites that must be fulfilled before children move the verb: children have to know that (i) finite verbs and infinitives belong to the same category V, (ii) that finiteness is a property that can be added to items belonging to this category via inflection, and (iii) that finiteness is related to a syntactic position high up in the tree. Like Roeper (1996) and Van Kampen (1997, 2001) who claimed that children are ready to project the sentence skeleton of CP-IP-VP from onset on, I assume that children know about dominance relations in a sentence from early on. One could argue that the basic hierarchy has an inherent logic. The features are projected in a strict hierarchical pre-determined order: first argumental features (resulting in VP and creating structure for the verbs internal arguments), then predicational features (leading to IP and creating a position for the verb's external argument) and finally illocutionary features (leading to CP and providing the structure that enables binding to a wider discourse). Note that Finiteness (or Tense; I do not make a distinction, see Chapter 3 section 3.2.1, for the relation between finiteness and Tense) is a property of the C-domain, as it scopes over the entire proposition expressed in a sentence and is bound to an antecedent in the preceding discourse. Thus, I assume that the third prerequisite I formulated for the acquisition of verb movement is fulfilled because the knowledge of the positioning of finiteness features in a sentence belongs to the a priori knowledge of language that children have.[12]

To meet the other two prerequisites, children have to learn inflection. The acquisition of inflection leads (i) to a set of inflectional endings that can be attached to the verbal stem, and (ii) to a generalisation over inflected finite verbs and non-inflected infinitives. Segmentability is children's ability to segment a phonological form into smaller chunks (Peters, 1982). Earlier this chapter, I cited Wijnen & Elbers (1998), who mention the role of segmentability - the ability to segment - for learning verb forms.

---

the inflected verb moves in order to project agreement features and create a position for the verb's external argument (that is, outside the VP, traditionally this is V-to-I movement). In Dutch, agreement is poor. Thus, the verb only moves for tense and there is no I or Agr projection. See section 5.4.5 for more information about agreement in Dutch.

[12] This is confirmed by the child data in so far as Dutch children (and also the children in my sample) tend to place the finite item, either lexical or grammatical, in sentence-initial position from the moment they come in.

Segmentability is crucial for the acquisition of finite verb forms: children have to learn that finite verbs consist of stem + suffix or of an auxiliary + infinitive. A number of cues guide them to the discovery that verb forms are segmented and consist of stem + suffix. When children know different forms from one and the same paradigm (i.e. paradigmatic variation), for instance *loopt* 'walks' and *lopen* 'walk', the children are, theoretically, able to analyse verb forms and recognise what is the stem and what the suffix. Likewise, knowledge of finite forms from different verbs, e.g. *loopt* 'walks' versus *kust* 'kisses', provide the possibility to segment as well (i.e. lexical variation). Lexical overlap between finite forms and infinitives is a special case of paradigmatic variation: it is the variation between paradigmatic forms and non-paradigmatic verb forms. Paradigmatic and lexical variation is knowledge that children in all languages need in order to build up inflectional paradigms. Wijnen & Elbers' (1998) claim is that the independent appearance of various auxiliaries (e.g. the modal auxiliaries) as main verbs next to the independent use of infinitives in RIs is helpful for detecting the borders of auxiliary and infinitive in PVs and hence, for the analysis of verb forms into auxiliary + infinitive. As it is precisely auxiliaries that Dutch children acquire as their earliest finite forms, it is expected that PVs can be segmented, and hence will be acquired early on in Dutch child language.[13] I claim that inflection, on the other hand, comes in fairly late in child Dutch.

On the basis of findings reported in Chapter 4, I conclude that both the early appearance of PVs as well as the late acquisition of inflection are effects from one and the same factor, namely patterns in the input. In section 4.7, I examined the types of verbs that Abel, Daan, Josse, Laura, Matthijs and Peter hear as finite forms and as infinitives in the input. Partitioning of the verb types into state-denoting and event-denoting predicates led to the observation that 89 % of the finite verbs in the caregivers' speech data were stative, showing a major contrast with the 3 % stative infinitives. The finite forms in the input are predominantly auxiliaries, modals and copula (N = 5920), and a minority of the finite forms in the

---

[13] In this respect, the fact that Dutch is a SOV language with Verb Second in main clauses may prove useful: auxiliary and infinitive are divided by an object. This disjunction may be helpful to the child to tease auxiliary and infinitive apart. English, as an SVO language, does not have this advantage.

input are thematic (N = 2002). The thematic finite predicates are most often verbs of position (*liggen* 'lie', *zitten* 'sit', *zijn* 'be')[14], possession (*hebben* 'have'), and mental states *(denken* 'think', *vinden* 'find/think', *weten* 'know', *bedoelen* 'mean', etc.).[15] Thus, the early emergence of finite auxiliaries in early child Dutch seems to be directly linked to the frequent use of finite auxiliaries in the input. Since lexical variation is one of the cues for children to segment verb forms into stem and suffix, I conclude that the observed marginal lexical variation within the set of finite forms in the input has a negative effect on the acquisition of inflection by Dutch children. Moreover, the Dutch inflectional paradigm contains a number of known obstacles for acquisition of inflection such as zero-morphemes and homomorphemes (Slobin, 1981). Homomorphemes, in particular, delay segmentation. Note that the segmentation of the verbs in SFs into stem + suffix does not profit from the early acquisition of lexical finiteness markers (unlike PVs).

The implication of the late acquisition of inflection is that the earliest finite forms (i.e. the finite verbs that are used before inflection is acquired), are unanalysed. This means that finiteness is already present before the children can make use of grammatical devices for finiteness marking (inflection followed by verb movement). Hence, finiteness starts out as being (part of) the lexical content of 'verbs'. This is confirmed by the observation that lexical-finiteness-markers are auxiliary-like predicates that denote primarily temporal and modal meanings (De Haan, 1987). I placed 'verbs' in quotes, because it is questionable whether children classify these early finite forms as verbs (V). Early in the developmental sequence, children treat finite forms and infinitives as disjoint classes. The two forms differ lexically, morphologically and syntactically; there are no indications whatsoever that the two types of predicates are part of the same class in early child language. Rather, the disjunction indicates that the opposite is true. Syntactically, early finite verbs could be finiteness markers that are simply adjoined to an XP

---

[14] I distinguished the copula *be* ('he is mad') from the verb of location *be* ('he is at home').

[15] See for similar observations about the uneven distribution of verb types of finite forms and infinitives in adult Dutch and Dutch child-directed speech Schlichting (1996) and Wijnen, Kempen & Gillis (2001).

high up in the sentence-representation, say the CP, where they modify the proposition.[16] This is illustrated in (12):

(12)
```
                    CP
                  /    \
              MOD        VP
               |          |
             gaat         V
               |          |
           [ +fin ]     lopen
                          |
                       [ -fin ]
```
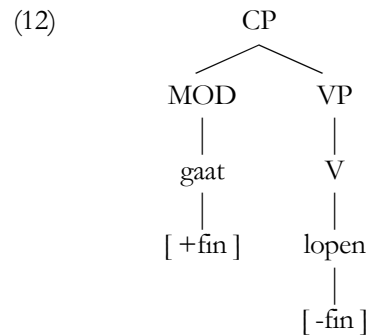
The acquisition of inflection leads to a change. Through inflection, children generalise over infinitives and finite forms, as predicates in the 'finite class' as well as in the 'non-finiteness class' possess the ability to inflect. From now on, the two forms belong to the same lexical category (V). In other words, via the acquisition of inflection, finiteness is re-analysed as a property of V.[17] The effect of the acquisition of inflection is that the verb enters the derivation as a fully inflected finite form and not as an infinitive. The syntactic effect is that the item carrying finiteness features (i.e. the inflected verb) is not simply 'stuck in' a high sentence position through adjunction, but moves to this position, which I call FinP in the representation given below in (13), in order to create a position for the finite verb.

---

[16] According to De Haan (1987) and Van Kampen (1997) the early finiteness markers are located in C. Hoekstra & Jordens (1994) argue that they are VP adjuncts. Given my assumption that children know from early on that finiteness or tense takes scope over the entire proposition, the lexical-finiteness-marker is a CP-adjunct and does not adjoin to the VP.

[17] Note that one could say that V is also re-analysed. Following Pinker (1984), children initially link V to a semantic class of words that denote activities. However, in order to arrive at an adult category V, morphology is a necessary cue: not all verbs denote activities but all verbs can be inflected.

(13)   FinP

                         Fin

            loopt$_i$ VP

            [+fin]   V

                                 t$_i$
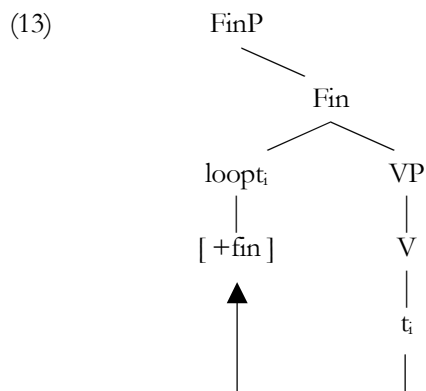
In sum, on the basis of a number of observations in early child Dutch and properties of the target/input language of Dutch, I have described the rise of FINs in the production of children acquiring Dutch in terms of the GoF (Grammaticalisation of Finiteness hypothesis), according to which the grammatical marking of finiteness is acquired incrementally. The GoF states that Dutch children take the following developmental path: RI-SF$_{lex}$-PV$_{lex}$-SF$_{gram}$. The initial stage that is characterised by the absence of finiteness marking is followed by a stage in which finiteness is a lexical property (SF$_{lex}$-PV$_{lex}$). Via the lexical stage, the children bridge a gap that exists between the forms with which they start out, RIs containing an unmoved infinitival verb form, and the final grammatical stage that is characterised by inflected moved verbs (SF$_{gram}$). The distance between RIs and SFs is emphasised by the input of Dutch children, as, in the input, there is marginal lexical overlap between finite verbs and infinitives. Additionally, there is little lexical variation within the set of finite verbs. Considering the three factors, marginal lexical overlap, little lexical variation, and poverty of paradigmatic variation of the Dutch inflectional paradigm, it is expected that inflection is learned relatively late by Dutch children. This does not imply that finiteness *marking* comes in late. Distributions in the input may have a delaying effect on the acquisition of the grammatical marking of finiteness, but they reinforce the acquisition of lexical-finiteness-markers. As an effect, the grammatical-finiteness-stage in Dutch child language is preceded by a lexical-finiteness-stage. In comparison to the RI-stage, the lexical-finiteness stage implies linguistic progress for the children: they can now add meaning to their sentences by using a finiteness marker. Structurally, the sentences that characterise the lexical-finiteness stage are close to the sentences produced

in the preceding RI stage because position as well as morphology of the main verb remains unchanged. Thus, by using lexical finiteness markers, children produce finite sentences that are close to RIs and that do not require knowledge of inflection and movement.[18] In order to look beyond the misleading patterns in the input and leave the lexical-finiteness-stage behind, Dutch children need a certain amount of lexical knowledge to learn inflection and move the verb.

5.2.4   Testing the GoF

Recall that the data I selected from Abel, Daan, Josse, Laura, Matthijs and Peter represent four stages that are characterised by a decreasing number of RIs and an increasing number of FINs (cf. § 2.1). I selected the files relatively roughly on the basis of frequencies of FINs. In the selection, I did not consider qualitative properties of verbs other than a simple finite/non-finiteness distinction. This distinction was made on the basis of morphology and verb placement. In the first stage, children use RIs only and no FINs (that is, the few FINs that are found, seem to be accidental) In order to select files that represent stage II, i.e. the stage in which the first FINs appear, I searched the children's corpora for the file in which the first FINs appeared besides RIs, selected this file and the files that immediately followed. Thus, the second stage represents the stage in which the first FINs come in. In the third stage, the number of FINs has reached a frequency that equals the number of RIs, whereas in the fourth stage, FINs are predominant.

In following sections, I test whether or not the Dutch data support the GoF, repeated in (14) below:

---

[18] Note that this avoidance strategy can be applied for two reasons: (i) because it is necessary, as knowledge of inflection and verb movement are absent, and (ii) because it is less effortful, as the two grammatical operations (inflect and move) are not yet automatise processes for the child (this claim can be found in the work of Philips, 1995, for instance).

(14)   *Grammaticalisation of Finiteness (GoF) Hypothesis (final)*

    (i)   In early child Dutch, the grammatical marking of finiteness by means of inflection and verb movement is preceded by a lexical-finiteness-stage.

    (ii)   In the lexical-finiteness-stage, simple lexical finiteness markers precede complex lexical finiteness markers.

Thus, we have to determine whether FINs appear before inflection is learned. I begin by investigating at which point of time Abel, Daan, Josse, Laura, Matthijs and Peter acquire inflection. The cues for the acquisition of inflection that will be examined are lexical overlap (§ 5.2.5), lexical variation (§ 5.2.6), paradigmatic variation (§ 5.2.7) and errors with inflection (§ 5.2.8). Inflection errors, i.e. non-adult use of inflectional endings, provide a classical test for determining when the children segment finite forms. The errors indicate that the children have some but not full knowledge of the inflectional system. On the basis of the outcome of the preceding four sections, I will examine in § 5.2.9 whether there I evidence for the claim that (i) a lexical-finiteness-stage precedes a grammatical-finiteness-stage, and (ii) the lexical-finiteness-stage consists of two substages. The findings and conclusions are summarised in § 5.2.10.

5.2.5   Lexical overlap

In Chapter 4, I examined lexical overlap between the verb types used in SFs and RIs and concluded that, contrary to claims made by for instance De Haan (1987) and Ferdinand (1996), lexical overlap occurred and was not even marginal during the RI-period. Further examination of these data reveals a time effect: lexical overlap emerges at stage III and grows over time.

Table 5.2: The development (accumulation) of lexical overlap between the verbal predicates in RIs and SFs, data from all six children

	I/II $N_{OVERLAP}$ (N_{INF}/N_{FIN})		III $N_{OVERLAP}$ (N_{INF}/N_{FIN})		IV $N_{OVERLAP}$ (N_{INF}/N_{FIN})	
Abel	0	(13/2)	3	(34/15)	10	(46/24)
Daan	2	(24/9)	6	(36/21)	13	(53/35)
Josse	0	(20/5)	1	(38/11)	8	(53/21)
Laura	2	(14/8)	6	(45/17)	8	(55/24)
Matthijs	0	(36/2)	7	(63/14)	16	(73/35)
Peter	1	(21/3)	9	(31/18)	18	(33/33)

The steady growth of lexical overlap from almost no overlapping items to 4.3 overlapping items in stage III (i.e. the average of the six children) and 12.2 in stage IV shows that the collapsed data from Chapter 4 obscure changes over time. Detailed information about the data in Table 5.2 can be found in Appendix 5.1 (p. 253).

5.2.6   Lexical variation

The investigation of the development of lexical variation of finite verbs serves two purposes. It provides further insight in the lexical knowledge of the children in the four stages (besides lexical overlap and paradigmatic variation). It may also be expected that when the children learn to inflect, a sudden growth is noticeable in the number of different lexical items that get inflected. In Table 5.3, I summarised the lexical growth of SFs. This table gives the lexical variation of finite forms in the different stages. 'N_{ACC}' gives the total numbers of the elements in set of finite verbs in a stage (i.e. the accumulated set), whereas '$N_{INCREASE}$' gives the number of new verb types from one stage to the next stage. Detailed information can be found in Appendix 5.1 (p. 253).

Table 5.3: Accumulation of finite verb types (i.e. growth of lexical variation) in the stages I, II, III and IV, data from all six children

	I/II	III	IV
	N_{ACC} ($N_{INCREASE}$)	N_{ACC} ($N_{INCREASE}$)	N_{ACC} ($N_{INCREASE}$)
Abel	2 (2)	13 (15)	9 (24)
Daan	9 (9)	21 (12)	35 (14)
Josse	5 (5)	11 (6)	21 (10)
Laura	8 (8)	17 (9)	23 (6)
Matthijs	2 (2)	14 (12)	35 (21)
Peter	3 (3)	18 (15)	33 (15)

The results in Table 5.3 show that lexical variation of finite forms undergoes a steady growth from stage to stage. There is no sudden increase pointing to the acquisition of a morphological rule. However, in this respect it is important to be aware of the lexical restrictions that adult speakers of Dutch (or, at least, the children's caregivers) show in their finite verbs. Given this observation, a sudden growth in lexical variation in children's finite verbs may be concealed by limitations on the use of finite forms, other than the absence of knowledge of inflection.[19]

5.2.7 Paradigmatic variation

If we collapse the data of the six children, there are a total of 116 SFs in stage II. They have either first person singular subjects or third person singular subjects. Only 8 forms carry an inflectional suffix and the rest of the forms are similar to the stem or are part of an irregular paradigm (and hence, do not carry inflection). Thus, the earliest finite verbs that children use carry hardly any overt signs of inflection. One could think that the

---

[19] It is unclear where this pattern (or restriction) in adult Dutch comes from. Hence, I cannot determine if it is likely that the children have this knowledge as well.

absence of inflection shows omission of inflection, and hence, can be interpreted as an error that points to early segmentation. This is *not* the case, however: the earliest finite forms are predominantly auxiliary-like items with impoverished inflectional paradigms, which implies that they often do not have overt inflectional endings (see: § 5.2.1). In addition, the items with first person singular subjects are not expected to carry inflectional suffixes.

From stage III onwards, all children show some paradigmatic variation; they start to use different inflectional endings with the same verbal stem. Note that here the lexical growth, which was described in the previous section, surfaces: obviously, the children do not only use verbs with impoverished inflectional paradigms anymore. From stage III onwards, the children vary number and distinguish between first and second/third person (like Dutch adults). In Appendix 5.2 (p. 261), the paradigmatic variation in the inflected forms that are used by Abel, Daan, Josse, Laura, Matthijs and Peter in stages III and IV is listed. This list does not contain verb types from which only one paradigmatic form is used. Paradigmatic variation in these stages concerns agreement rather than tense. The (infrequent) past tense forms that are used in stages III and IV are irregular: they are not formed by suffixation of the past tense suffix but they show vowel change in the stem. The past tense forms I found are highly similar for all children: *wou* 'want-SG-past', *was* 'be-SG-past', *deed* 'do-SG-past', *had* 'have-SG-past', *ging* 'go-SG-past' and *zat* 'sit-SG-past', which are the past tense forms of respectively *willen* 'want', *zijn* 'be', *doen* 'do', *hebben* 'have', *gaan* 'go' and *zitten* 'sit'. These are all highly frequently used verbs. This observation raises the question whether the early past tenses are solely a frequency effect or whether the early acquisition of irregular past tense forms reveals that irregular past tenses are easier to pick up or easier to use than regular forms. The rationale is that the clear phonological contrast between present and past tense in the irregular cases may enlarge the saliency of the forms, and hence, facilitate their acquisition. Another, processing-related account could provide an explanation as well. If the irregular forms remain stored as unanalysed chunks, the irregular forms have the advantage that the operation of inflection does not have to be carried out. The direct retrieval of past tense forms out of the lexicon may put less strain on the children's processing abilities than when a past tense form has to be composed. Because of limitations on their processing abilities, children may fall back on RIs in the regular cases.

There are various possible explanations for children's use of past tense during the RI-period and it is, at this point, unclear which of these explanations is most successful. Nevertheless, we can draw some conclusions based on the empirical observations in this section. In stages III and IV, all children show some paradigmatic variation. They use different forms from the agreement paradigm. Tense inflection appears after stage IV. With regard to segmentation, the results suggest that the children are not able to segment verb forms earlier than stage III, simply because they lack the lexical knowledge that enables them to learn how to segment.

5.2.8   Errors

A test for the acquisition of the analysis of verb forms is the emergence of inflection errors. Errors indicate that children segment the unanalysed chunks they have extracted from the input; they attach inflectional endings in a way that is different from what they could have heard in the input. In particular omissions of inflection, resulting in bare stems, suggest that the children are segmenting. In Table 5.4, the errors in the SF_θ of the six children are summarised. I only included SF_θ, as non-thematic verbs are hardly inflected. Note that the errors concern errors with agreement, as there is no tense inflection yet. Nearly all errors are made in stage III and IV. The errors in Table 5.4 are divided into 'overused bare stems' and 'wrong inflection'. Items with wrong inflection are errors with person agreement that show overgeneralisation of the second or third person singular suffix [-t] or errors with number agreement that show overgeneralisation of the plural ending [-en]. As mentioned in Chapter 3, bare stems can be items with dropped inflection or agreement errors (that is, overuse of first person singular forms). As this distinction cannot be made between the tokens in the data, I listed them simply as 'overused bare stems (BS)'. Unlike the bare stem forms listed in Chapter 3, I have now included the items with dropped infinitival or participial morphology:[20]

---

[20] I did not include forms like *vin* 'find' or *lus* 'taste' in which the children have omitted the final [-t]. This is omission of part of the stem rather than omission of inflection. In colloquial

Table 5.4: Table 5.4: Inflection (agreement) errors, data from all six children

	$N_{SF\theta}$	Overused BS[21]	Wrong Inflection
Abel	119	11	4
Daan	289	16	2
Josse	137	22	2
Laura	157	29	2
Matthijs	95	8	5
Peter	424	21	2
Total	*1221*	*107*	*17*

Appendix 3.3 (p. 245) contains a list of the overused bare stem forms. The unambiguous agreement errors (overgeneralisation of second/third person singular and wrong number) are listed in Appendix 5.3 (p. 262). Omission of inflection occurs as early as stage II, but the unambiguous errors appear in stages III and IV. The errors are most frequently examples of overgeneralisation of second/third person singular forms. This is illustrated in (15):

---

speech, adult speakers of Dutch also tend to omit this ending in sentences such as *lus je geen erwtensoep?* instead of *lust je geen erwtensoep?* 'don't you like pea soup?'.

[21] These do not contain utterances with first person singular subjects as in Dutch these forms cannot be distinguished from bare forms with dropped inflection.

(15) a.   ik ook ziet ook niet   Abel 2;03.02
            I also see-*2/3SG* also not
            'I do not see it either'

      b.   ik doet weer   Josse 2;03.28
            I do-*2/3SG* again
            'I do it again'

      c.   ik gaat me niet   Laura 3;03
            I go-*2/3SG* me not
            'I am not going'

      d.   is ik weer   Matthijs 2;10.22
            be-*2/3SG* I again
            'I am again'

      e.   ik heeft hem   Peter 2;03.21
            I have-*2/3SG* him
            'I have him'

In a few nice cases, children regularise. In the case of the plural forms *he**vv**en* instead of *he**bb**en 'have'* (Daan 2;09.10) and *k**ann**en* instead of *k**unn**en 'can'* (Josse 2;07.20), the child based the plural on the singular forms *heeft* and *kan*, segmented the forms and added the plural suffix [*-en*]. In adult Dutch, these forms are irregular, however. The data in Table 5.4 show that the children make quite a number of agreement errors: if the data of the six children are collapsed, and the overuse of bare stems (which is either drop of inflection or incorrect use of first person singular forms) and wrong inflection are included in the count, the percentage of errors is approximately 10 %.[22] Most errors can be characterised as overuse of bare stems.

A. de Haan (1996) carried out a study on errors that Dutch children make with inflection. She reports data that are invaluable for the present study as they provide extra insight in the acquisition of inflection by Abel, Daan, Josse and Matthijs, more specifically in the errors that these children make. As A. de Haan did not select files out of the entire corpora of these children but examined all files and therefore made use of a dataset with greater

---

[22] This contrasts with findings reported for German by Clahsen & Penke (1990) and Poeppel & Wexler (1993) and Guasti's (1994) findings for Italian. Both report hardly any agreement errors. They have to carried out analyses over time, though.

density, she has been able to make more precise longitudinal observations about errors than I did. In addition, she followed these four children until an older age than I did; see Table 5.5:

Table 5.5: Data studied by A. de Haan (1996)

Child	Age range
Abel	1;10.30 - 3;04.01
Daan	1;08.21 - 3;03.30
Josse	2;00.21 - 3;04.17
Matthijs	1;10.13 - 3;07.02

Similar to what is reported here, A. de Haan did not find tense inflection. This agrees with the conclusion that tense inflection is acquired after agreement inflection.[23] With regard to agreement, A. de Haan distinguished between person and number agreement. She observed that number errors start to appear at the age of 2;5 while person errors start to emerge around the age of 2;6. The older the children got, the worse they performed.[24] It can be concluded from this that segmentation/inflection generally comes in around the age of 2;5/2;6.[25] A comparison with my data shows that these ages are exactly between stages III and IV in the data I examined of these four children: the average age of Abel, Daan, Josse and Matthijs in stage III is 2;4 and in stage IV is 2;9. The conclusion is that the children start to segment finite verb forms. The onset of the acquisition of inflection lies between stages III and IV. This is compatible with earlier findings on lexical

---

[23] This does not mean that the children do not have any knowledge of the distinction between present and past. Recall that I found a few lexically marked tensed forms; this suggests that the children know that past tense can be marked with a verb. The grammatical means to do this comes in late, however.

[24] Only up to a certain age of course. When the children master inflection, the error rate is expected to decrease rapidly.

[25] I examined data from Laura as well as Peter. These two children may deviate form the other four children in two different ways. Laura has a delayed development, and therefore, may start to acquire inflection relatively late. Peter is a quick learner, and hence, may be earlier than the other children.

overlap, lexical variation and paradigmatic variation. These indicate that the children are, theoretically, able to segment finite verbs from stage III onwards but not earlier.

5.2.9   The first occurrences of FINs

Given the GoF, it is expected that FINs come in prior to the stage in which inflection is acquired, that is, prior to stage IV. From the previous sections, we already know that this prediction is borne out and that FINs appear already in the stages II and III. Thus, Dutch children use finiteness before they have knowledge of the grammatical marking of finiteness by inflection. In this section, I zoom in on the lexical-finiteness-stage, that precedes the grammatical-finiteness-stage in early child Dutch. Does the lexical-finiteness-stage fall apart into two substages indicating the application of an incremental learning procedure? More specifically, do the early SFs contain a restricted set of auxiliary-like predicates that, one stage later, appear as auxiliaries in PVs?

Tables 5.6 - 5.8 give the frequencies of RIs, SFs and PV in stages I, II, III and IV. The numbers in a stage are bold-faced when a verb form shows a level of productivity that exceeds incidental use. If there were 5 or more occurrences of a verb form I considered those verb forms to be productive. Although the size of incidental use is quite arbitrary and the data below show considerable differences between the children, the general trend across these six children is similar: the first FINs are SFs (stage II), whereas PVs appear in stage III:

Table 5.6: Number of RIs in stages I, II, III and IV, data from all six children

	I	II	III	IV
Abel	n.a	33	84	42
Daan	**5**	54	66	71
Josse	n.a	99	76	64
Laura	**18**	56	276	65
Matthijs	**40**	127	112	51
Peter	**26**	114	57	29

Table 5.7: Number of SFs in stages I, II, III and IV, data from all six children

	I	II	III	IV
Abel	n.a	**9**	172	193
Daan	4	**47**	223	505
Josse	n.a.	**13**	93	325
Laura	4	**21**	282	272
Matthijs	8[26]	**5**	56	384
Peter	0	**5**	79	542

---

[26] This is an unexpected high number of SFs in a very early stage, but as this includes a repetition of a form that appears in one early file and re-appears only a couple of months later (*mag* 'may'), it seems justified to interpret this as 'incidental use'.

Table 5.8: Number of PVs in stages I, II, III and IV, data from all six children

	I	II	III	IV
Abel	n.a	0	**26**	82
Daan	0	2	**23**	116
Josse	n.a.	0	**5**	114
Laura	0	5	**66**	100
Matthijs	0	1	**29**	140
Peter	0	0	**16**	85

Can the emergence of PVs in stage III be seen as the result of the preceding two stages, in the sense that the PVs are combinations of the (stage II) SFs with the earlier (stage I) infinitives? In order to find this out, the finite forms from stage II have to be compared to the auxiliaries that appear in stage III in PVs. The PVs in stage III contain a variety of auxiliaries. Modal auxiliaries are predominant (*kan* 'can', *wil* 'want', *moet* 'must or *mag* 'may'), aspectual auxiliaries like *gaat* 'goes' and *kom(t)* 'comes' or dummy tense markers like *doe* 'do', *gaat* 'goes' and *is* 'is' are used by most children. Incidentally, a form like *lig* 'lie' is used in a PV in stage III.

If this set of auxiliaries is compared to the verbs that appear in SFs in stage II, there is a considerable amount of lexical overlap: 75 out of the 116 SFs in stage II contained modals, aspectuals like *gaat* 'goes' or *komt* 'comes' or the copula *is* 'is' (with a preponderance of modals). These are all auxiliary-like predicates and are similar to the finite verb in stage-III-PVs in form as well as meaning. Additionally, there were some verbs of position like *zit* 'sit' or *lig* 'lie', some of which were used to describe a position (SF$_{AUX}$), and others to describe an action (SF$_\theta$).[27] An overview is given in Table 5.5:

---

[27] Verbs of position can select infinitival complements in Dutch. These constructions denote ongoing aspect:

Table 5.9: Distribution of SF_{AUX} and SF_θ over SFs in stage I/II, data from all six children

	SF_{AUX}		SF_θ	
Abel	9		-	
Daan	39		12 (past, 'fits', zie 'see', zingt 'sings')	
Josse	11		2 (rink 'jumps', heet 'is called')	
Laura	19		6 (hoor 'hear', valt 'falls', zit 'sits')	
Matthijs	6		7 (past 'fits')	
Peter	3		2 (eet 'eat', zit 'sits')	
Total	*87*	*(75 %)*	*29*	*(25 %)*

Nearly all verb types that appear in stage III in PVs as auxiliaries are used before in the preceding stage II as main verbs in SFs. In the SFs of stage II, SF_{AUX} dominate and SF_θ are relatively rare. This pattern, that is, the distribution of verb types over the earliest SFs, is very similar to the earliest finite forms found by others who examined production data from Dutch children (De Haan, 1987; Van Kampen, 1997; Wijnen & Elbers, 1998).

5.2.10 Conclusion

A comparison between the results reported in § 5.2.5 - § 5.2.8, on the one hand, and § 5.2.9, on the other, shows that the Grammaticalisation of Finiteness (GoF) Hypothesis that I formulated at the beginning of section 5.2, is confirmed by the Dutch data: Dutch children use finite sentences before they have access to the morphological rule to inflect verbs. Stage II is characterised by the appearance of SFs, whereas the appearance of PVs is a

---

    (i)   Jakob zit te tekenen
                Jakob sit-*fin* to draw
                'Jakob is drawing'

characteristic of stage III. Various factors (i.e. lexical overlap, lexical variation and paradigmatic variation) indicate that in stages II and III, the children do not know that verb forms consist of segments, that is, stem + suffix, however. By implication, the earliest finite sentences contain lexical-finiteness-markers. The conclusion that finiteness starts out as a lexical feature and is reanalysed as a grammatical feature is confirmed by (i) the emergence of inflection errors between stages III and IV (indicating that children do not segment verb forms prior to stage IV), and (ii) the dominance SF$_{AUX}$ predicates in the earliest SFs (which have hardly any other lexical content than temporal and modal meanings). As a result from lexical overlap, lexical variation and paradigmatic variation, finiteness will be re-analysed as a grammatical feature that can be added to items of that belong to the class of V by means of a morphological rule. The syntactic effect of the acquisition of inflection is that the operation of verb movement is introduced in the child grammar.

5.3   The modal shift in RIs

In this chapter, little attention has been paid to RIs so far. In this section, one of the effects of the rise of FINs on RIs will be discussed. Returning to an observation that was left unexplained in Chapter 3, I will argue that the Modal Reference Effect (MRE), i.e. the predominance of modal RIs in the set of RIs, has a developmental cause. Furthermore the semantic effects of the acquisition of finiteness are discussed.

I begin by showing that a breakdown of the results from Chapter 3 over time reveals a developmental effect: in stages III and IV, RIs are used significantly more often to denote modality than in stages I/II. In earlier work, I referred to this development as the *modal shift* (Blom & Wijnen, 2000; Blom, 2002). Figure 5.1 gives an impression of the modal shift: the proportion of modal RIs in stages I/II is compared to the proportion of modal RIs in stages III/IV. For all children the lower bar, which gives the proportion of modal RIs in stages III/IV as a percentage of the total number of interpretable RIs in these stages, is longer than the upper bar, which shows the proportion of modal RIs in stages I/II.
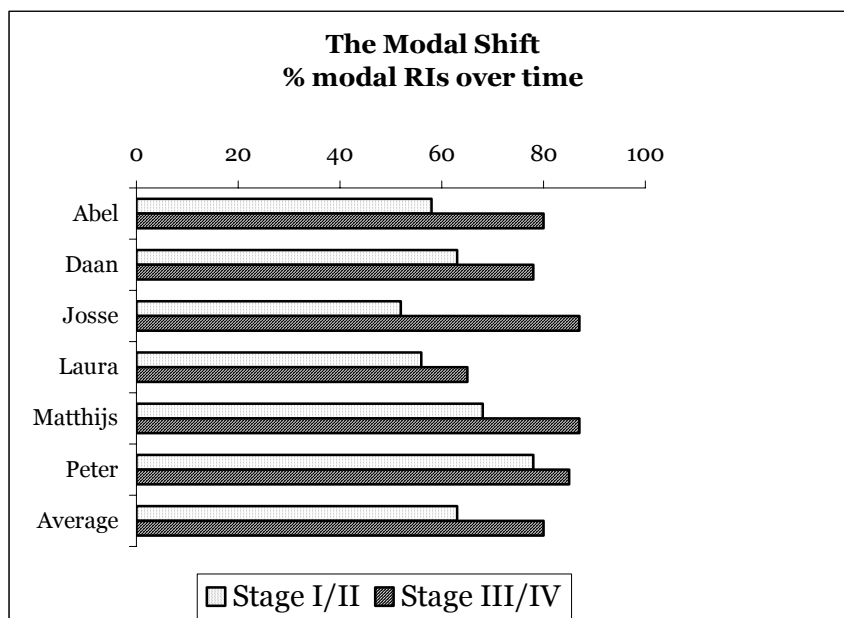
Figure 5.1: Modal RIs, average and standard deviations; a comparison between the stages I/II and III/IV, data from all six children

Table 5.10: Modal RIs, average and standard deviations; a comparison between stages I/II and III/IV, data from all six children

	Stages I/II		Stages III/IV	
	N_{RI}	% MODAL (N)	N_{RI}	% MODAL (N)
Abel	24	58 % (14)	96	80 % (77)
Daan	49	63 % (31)	108	78 % (84)
Josse	84	52 % (44)	122	87 % (106)
Laura	59	56 % (33)	255	65 % (167)
Matthijs	120	68 % (82)	134	87 % (117)
Peter	123	78 % (96)	74	82 % (61)
Average		*63 %*		*80 %*
SD		*9*		*8*

In order to calculate the statistical significance of the modal shift, the permutation test and Jack knife technique are applied (cf. Chapter 2). Table 5.11 gives the p-values that give an estimation of the probability that our observations can be found by chance.

Table 5.11: Probabilities that the observed average difference between modal use in stage II and modal use in stages III/IV is due to chance. The probabilities are calculated over the entire group of six children and over 6 groups of five children, with one specific child omitted at a time (Jack knife method)

	p-value
All children	**0.004**
Omitted	
Abel	0.007
Daan	0.014
Josse	0.018
Laura	0.003
Matthijs	0.005
Peter	0.01

The first value (p = 0.004) is the probability when the average from the total sample of six children is taken; the rest are the values when one of the children is left out from the statistical analysis. Given that the probability is below the critical value of 0.05 in all cases, we may be confident that the modal shift is a statistically reliable change over time that takes place between stages II and III.

Interestingly, the modal shift is not restricted to child Dutch. Behrens (1993) analysed data from a German-speaking girl, Simone, (between the ages of 1;9.11 and 2;7.19) and found that Simone used relatively more RIs to

talk about modal events as she got older.[28] Initially, only 28 % (N = 7) of Simone's RIs were modal, while in the final file that Behrens examined 85 % (N = 27) of Simone's RIs were modal. Armon-Lotem (1995) observed a modal shift in the RIs of children acquiring Hebrew. Although I will not go into the cross-linguistic aspects of the modal shift, I think the German and Hebrew findings are noteworthy because they corroborate the conclusion that the modal shift is not an accidental observation.

5.3.1   The cause of the Modal Shift

If the number of modal utterances in general increases during the RI period, we would have a very straightforward explanation for the modal shift in RIs. It would just be one of the various modal shifts.[29] This explanation does not hold for the data I examined, however. In fact, if we look at the total number of utterances that the children use to talk about events in all four stages (see the data in Table 5.12) and calculate the proportions of ongoing and modal events, it turns out that during the four stages, the amount of ongoing, non-modal, utterances remains rather stable in the first three stages and that the proportion of non-modal utterances even shows an increase between stages III and IV. The proportions are given below:

---

[28] Behrens (1993) subsumes intentions, commands, non-actuals and future events under the label non-ongoing. According to my definition, these interpretations are modal.

[29] Halliday (1975) carried out a detailed investigation into the early development of language functions in the spontaneous speech data of his son, Nigel. He pointed out that use of the modal functions, such as the regulating ('Do as I tell you' or 'You must') and instrumental ('I want') function, clearly increase over time. This increase, however, takes place *before* the general age in which RIs come in (namely when Nigel is between the 16½ and 18 months) and clearly before the modal shift takes place. Therefore, Halliday's observations provide no further insight into the modal shift. Apart from this study, I do not know of any studies examining the quantitative development of modal child language.

Table 5.12: Development of frequencies and proportions of ongoing and modal utterances that denote events

	$N_{EVENTIVE}$	$N_{ONGOING}$ (%)		N_{MODAL} (%)	
I	80	34	(43 %)	46	(57 %)
II	436	168	(38 %)	268	(62 %)
III	999	447	(45 %)	552	(55 %)
IV	2621	1854	(71 %)	767	(29 %)

In earlier work (Blom & Wijnen, 2000; Blom, 2002), I proposed a different kind of explanation embedded within Hoekstra & Hyams' (1998) model. Recall that according to this model, infinitives in Dutch receive a modal reading because they possess infinitival morphology carrying the feature [-realised]. I hypothesised that the modal shift reflects children's discovery of the infinitive, as an effect of the ability to analyse verb forms. I argued that this ability directly followed from the acquisition of contrasting finite forms and, more specifically, from the emergence of lexical overlap between infinitives and finite forms in stage III (see also the results in Table 5.2 on p. 177). In essence, the idea was that the knowledge of infinitival morphology leads to the mapping of the infinitival suffix to the feature [-realised]. In this section, a revised analysis will be given. I will not argue that specification of infinitival morphology leads to the modal shift but that the unspecified temporal/modal nature of the infinitive, in combination with the growth of specified tensed forms, causes a relative increase of modal RIs. This property does not only lead to an explanation of the observed modal shift in child Dutch, but ultimately to an explanation of the marginal and context-dependent allowance of RIs in adult Dutch. In this section, I will present a number of empirical and theoretical arguments in support of the revised hypothesis.

In brief, my claim is that the modal shift in RIs can be interpreted as the effect of RIs being 'pushed out' by sentences containing verb forms that are semantically more specified than infinitives. In § 5.3.3, I will give a more detailed description of this 'pushing out'. For now, I hope that the metaphor is clear enough. Basically, the idea is that FINs take over at the expense of RIs. According to this scenario, the modal shift takes place because present

tense FINs are earlier productive than modal FINs. By effect, non-modal present tense RIs are used relatively less often than modal RIs. Hence, RIs display a modal shift.

In order to find out whether or not this hypothesis is supported by the data, I examined the distributions of RIs and specified alternatives of RIs used by the children during the four stages to describe ongoing present tense events and modal events. First, I selected all utterances from the six children that denoted events and divided this set into utterances that denoted events that were ongoing at speech time and modal events. The children used three different sentence types in both contexts: RIs, SF_θ 's and PVs. I examined the relative frequencies of utterances containing these three forms in both contexts in the four developmental stages. The results are shown in Figures 5.2 and 5.3. Tables 5.13 and 5.14 give the corresponding numbers. The relative number of ongoing and modal events remains similar in all stages and, at the same time, the forms children use to express ongoing and modal events change from stage to stage. This change is dependent on the novel verb forms that children acquire and that become productive. Initially, RIs will be used for ongoing as well as modal events. Thus, the distributions of forms over the sets of ongoing and modal events are identical. The modal shift suggests that the distributions of forms over the two sets diverge over time, however. Based on this hypothesis, my expectation is that in stage III, that is, when the modal shift takes place, specified alternatives to RIs that describe ongoing events are more productive than specified modal alternatives.
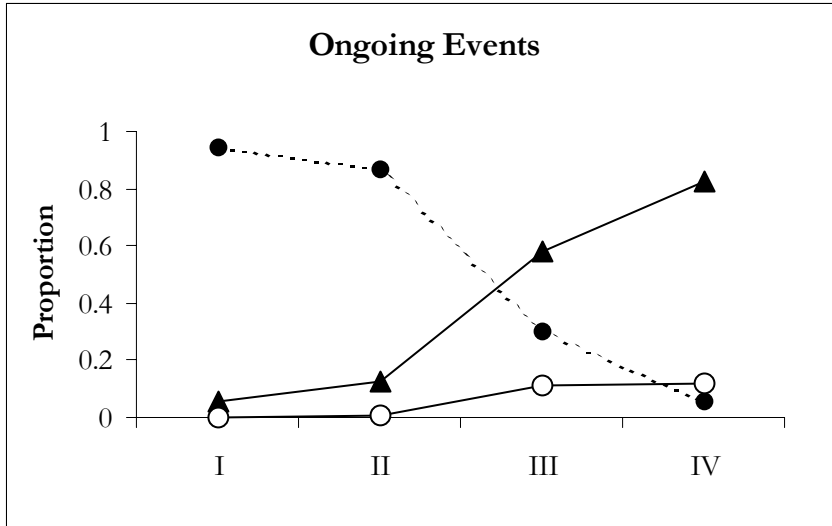
Figure 5.2: Percentages of RIs, SF_θ's and PVs used to describe ongoing events in four developmental stages, data form all six children collapsed
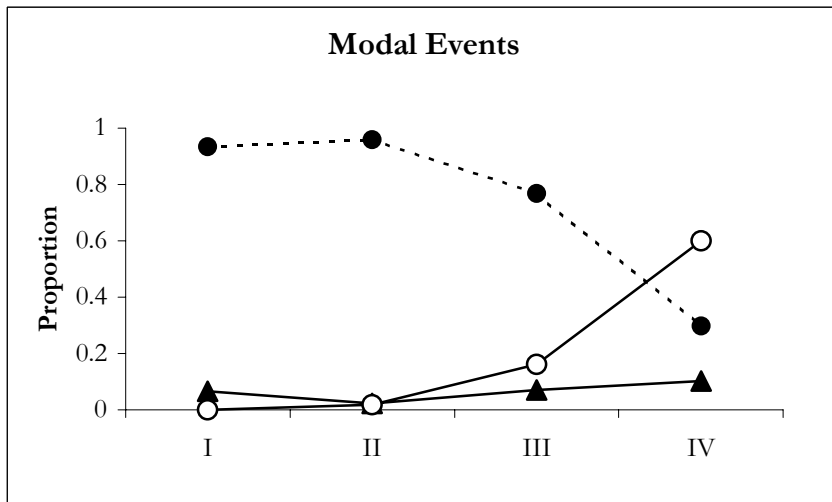
Figure 5.3: Percentages of RIs, SF_θ's and PVs used for modal events in four developmental stages, data from all six children collapsed
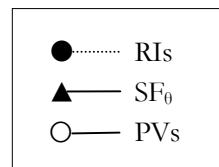
Table 5.13: Numbers (percentage) of RIs, SF_θ 's and PVs used to describe ongoing events in four developmental stages, collapsed data from all six children

Stage	N_{RI}	(%)	$N_{SF\theta}$	(%)	N_{PV}	(%)	N_{TOTAL}
I	32	(94 %)	2	(6 %)	0	(0 %)	34
II	127	(86 %)	19	(13 %)	1	(1 %)	147
III	116	(31 %)	212	(57 %)	43	(12 %)	371
IV	61	(6 %)	751	(79 %)	135	(14 %)	947

Table 5.14: Number (percentage) of RIs, SF_θ 's and PVs used for modal events in four developmental stages, data from all six children collapsed

Stage	N_{RI}	(%)	$N_{SF\theta}$	(%)	N_{PV}	(%)	N_{TOTAL}
I	43	(93 %)	3	(7 %)	0	(0 %)	46
II	256	(96 %)	6	(2 %)	5	(2 %)	267
III	402	(79 %)	22	(4 %)	85	(17 %)	509
IV	210	(30 %)	71	(10 %)	422	(60 %)	703

The data in Figures 5.2 and 5.3 suggest that the prediction is borne out: between stages II and III, the number of SF_θ that are used to describe events shows an increase. It takes place before modal PVs start to be used instead of modal RIs: modal PVs show a considerable increase between stages III and IV. A comparison between the rise of PVs in the two figures indicates that when PVs come in in stage III, they are used for modal as well as ongoing events. After stage III, PVs show a clear tendency towards modal use. Some examples of the stage III PVs are given in (16):

(16)   a.   kan niet zoeken   Abel 2;03.02
              can-*fin* not search-*inf*
              'I cannot get it'[30]

       b.   koe gaat rijden   Matthijs 2;04.24
              cow goes-*fin* drive-*inf*
              'The cow is driving'

       c.   wil tekening maken   Peter 2;00.28
              want-*fin* painting make-*inf*
              'I want to make a drawing'

       d.   sijn de soep e ete   Laura 2;04.01
              are the soup eat-*inf*
              'They are eating the soup'

In stage III, SF_θ 's and PVs together cover 69 % of the ongoing events.[31]
They cover only 21 % of the modal events in stage III. Thus, in stage III,
ongoing events are predominantly described through other forms than RIs
(that is, newly acquired forms), whereas in stage III modal events are still
predominantly described with RIs. This comparison suggests that the
hypothesis that the modal shift is an effect of the acquisition of specified
alternatives is on the right track.

Although the results are compatible my revised analysis of the modal shift, it
must be remarked that the test I carried out in this section gives only a
rough estimation of the way in which RIs are pushed out. I assumed that the
relative number, i.e. the quantity, of ongoing and modal utterances remains
rather similar throughout all stages. However, this does not mean that the
specific meanings, thus the qualitative properties of the sets of modal and
ongoing utterances, remain constant over time. If we zoom in on the
development of the specific modal forms, it becomes clear that we cannot
simply assume that the children use the newly acquired modal PVs instead
of RIs: with the acquisition of modal PVs, new modal meanings come in as

---

[30] Abel uses the verb *zoeken* 'search' also with the denotation 'get'.

[31] Note that this situation in which the cumulative effect of the acquisition of two
forms takes more effect than the acquisition of one form can only happen if the
development is gradual and children still use RIs alongside the other forms. If RIs
were immediately replaced across the board by a new form, there would be not
cumulative effect.

well. Modal RIs nearly always expressed intentions, dynamic necessity (wishes and desires) or deontic necessity (commands). With their modal PVs, all children introduce dynamic and deontic possibility, denoting abilities and permission, respectively. An example is given in (16a). Thus, only a part of the increasing number of modal PVs replaces RIs. The kind of expansion within the set of modal meanings does not take place in the case of ongoing/present tense meanings as (i) there are in the temporal domain not so many differentiated meanings as in the modal domain, and (ii) the forms that are used to mark the present-past tense distinction come in after stage IV. This does not mean that the rise of SF_θ 's is a case of simply replacing RIs, however: the rise of SF_θ 's is also evidence for lexical growth. With their newly acquired SF_θ 's, children denote ongoing activities that they did not denote before with their RIs.

Figures 5.2 and 5.3 do not only show that ongoing specified alternatives for RIs are productive earlier than modal specified alternatives; the developmental pattern of PVs in the two figures also reveals a developmental step that Dutch children take in their development of finiteness. This step is in the direction of the target language. By itself, however, it leads to non-adult forms, that is, the introduction of ongoing or present tense PVs. In stage III, the proportion of PVs that are used for ongoing events (12 %) is close to the proportion of PV, used for modal events (17 %). Between stages III and IV, however, the proportion of ongoing PVs stabilises (14 %) while the proportion of modal PVs goes up (60 %). The initial rise of non-modal PVs in stage III, as well as their stabilisation in stage IV, confirms the hypothesis that these PVs mark a developmental step that precedes the acquisition of the grammatical marking of finiteness by inflection and verb movement. The results are compatible with my conclusion that grammatical finiteness comes in around stage IV.

5.3.2   The driving force

To account for the modal shift, the 'pushing out'- metaphor was used. In this section, I will explain the process of replacing unspecific forms by specific forms can be captured in a more principled way. The idea is that children are forced to use the most specified form out of a set of related

forms (e.g. a inflectional paradigm) by an innate mechanism that works in the lexicon.

Kiparsky (1973), Anderson (1992) and Halle & Marantz (1993) proposed a selection mechanism that blocks the selection of underspecified forms: the *Elsewhere Condition*. Kiparsky (1973) introduced the Elsewhere Condition as a principle governing the application of morphological rules. The Elsewhere Condition states that if there are two rules, A and B, and (i) A (the specific) case includes B (the general case) and (ii) the application of rule A yields a distinct result from the application of rule B, A is applied first. If A takes effect, B is not applied. Anderson (1992) and Halle & Marantz (1993) applied the Elsewhere Conditions to related forms stored in the lexicon. In this framework, known as *Distributed Morphology*, lexical insertion takes place after syntax. The result of syntactic operations (movement and adjunction) is that a feature bundle consisting of morpho-syntactic features fills each syntactic position. For the phonological representation, the lexicon is searched for a form that matches this feature bundle. In this competition, the most specific form, that is the form that matches most features, wins. The less, or under-, specified form is used elsewhere, that is, where none of the more specified forms can be used.

Under the assumption that the Elsewhere Condition is an innate principle that works from early on, the driving force behind the modal shift follows from the elsewhere, or unspecified, character of infinitives. As soon as specified alternatives are acquired, RIs will not be used anymore. When the children have picked up the present tense form *zit* 'sit', they have to use (17) instead of (18):

(17)   mama zit
        mama sit-*fin*
        'mama is sitting'

(18)   mama zitten
        mama sit-*inf*
        'mama is sitting'

As there is lexical overlap between RIs and SFs at stage x, this hypothesis can, theoretically speaking, be tested: the prediction is that a certain lexical item used in an ongoing RI will not be used anymore in this way when this

item has appeared in an SF. In practice, however, the corpus data I examined provide too little information to test this specific prediction. Apart from the fact that the data are not dense enough, an additional problem is that non-modal RIs can still be used (in spite of the Elsewhere Principle) even though there are specified alternatives. They appear as so-called *presentationals* (*dit is* + *inf* 'this is + *inf*): SFs do not provide an alternative for this specific usage as the verb in SFs lacks the nominal properties of the infinitive. As pointed out in Chapter 3, the corpus data provide too little information to determine whether or not children use RIs to label actions.

5.3.3   Aspect, modality and tense

The Elsewhere analysis of RIs provides an explanation for a second observation that remained unexplained in Chapter 3. Recall that RIs do not only show predominance for modal interpretations, but they seem to be aspectually restricted in the sense that usage of RIs for completed events is excluded. In Lasser's (1997) terms, RIs obey a Non-Completedness Constraint (NCC). In this section, I will point out how the Elsewhere Hypothesis accounts for the NCC. There is, however, a second possible interpretation, which differs from the Elsewhere-explanation. This second explanation raises various questions with regard to the sequence of development of tense, modality and aspect in early child language and this may be a point of departure for further research. The difference between the two possible explanations hinges upon the nature of the infinitive. According to the first, the infinitive is aspectually, temporally, and modally an Elsewhere form. According to the second, the infinitive is only an Elsewhere form with respect to tense and modality.

Let me start with the first explanation, according to which RIs surface as long as children do not have access to specified modal, temporal and aspectual forms. According to the Elsewhere Hypothesis, RIs are vulnerable forms: their use is immediately affected by the acquisition of specified forms. Thus, if children learn [+completed] alternatives for RIs early, it is expected that completed RIs do not occur. Jordens (1990) reported that participles appear very early in Dutch child language. In the same period that RIs are used, children also use 'root participles', i.e. non-finite clauses

like RIs that contain a sentence-final participle instead of a sentence-final infinitive. Some examples of Dutch RPs are given in (19):

(19) a.   jij ook maakt   Abel 2;02.19
           you also make-*part*
           'You have made it as well'
      b.   afvalt   Daan 2;04.14
           off-fall-*part*
           'It has fallen off x'
      c.   ape poept   Josse 2;00.21
           monkey defecate-*part*
           'The monkey has defecated'
      d.   die ook in bad wees   Laura 2;05
           that also in bath be-*part*
           'That one has also been in bath'
      e.   boer daan   Matthijs 1;11.24
           burp do-*part*
           'I have done a burp'
      f.   Peter emmer daan   Peter 1;10.03
           Peter basket do-*part*
           'Peter has done it in the basket'

In my data, the early appearance of RPs is also found: the first unambiguous examples appear in stage II. When RPs were used, they had a very specific denotation (unlike RIs): all 102 interpretable RPs that I found had a completed aspect denotation. Though the overall number is low, this number must be seen in relation to the number of times that children talk about completed events in general. Given that there are no completed RIs, that SFs are nearly always used for present tense and that PVs (that is, the PVs that I concentrated on, namely those with an infinitival complement) are either modal or ongoing, it seems that all completed events in the earliest stages are denoted by RPs and later on, via PVs with participial complements. In sum, these observations make the assumption that children specify the infinitive as [-completed] superfluous: the NCC follows from the early availability of participles.

As mentioned in the introduction, a second interpretation of the NCC and the early appearance of RPs is possible, according to which the contrast

between [-completed] RIs and [+completed] RPs is more meaningful than is assumed in the preceding explanation. It can be argued that children learning new forms will always try to map these to a meaning and that different forms will always be mapped onto different meanings.[32] In this case, the children make a very early aspectual distinction with RIs and RPs and formally partition the set of events into completed and uncompleted events: completed events are described with RPs that contain the specified [+completed] participle form whereas uncompleted events are denoted by RIs containing the [-completed] infinitival form.[33] Considering the results I presented earlier in this chapter, the early aspectual distinction is followed by a mood/modal distinction between realis and irrealis (in the sense of Hoekstra & Hyams, 1998); the forms children use to make this distinction are SFs and PVs, denoting non-modal ongoing events and possible or necessary future events, respectively.[34] This distinction surfaces around stage III and, hence, after the aspectual distinction marked by RIs and RPs. According to this scenario, the tense distinction between present and past, systematically marked through inflection on the verb is the last to appear. Even in the fourth stage that I examined, Abel, Daan, Josse, Laura, Matthijs and Peter did not use past tense inflections, as was shown in § 5.2.7. In Table 5.15, this order of aspectual, modal and tense distinctions is summarised:

---

[32] In the literature on child language acquisition, various proposals have been made about restrictions that the Language Acquisition Device places on the mapping of a meaning to newly acquired forms. Pinker's (1984:177) *Unique Entry Principle* according to which "no complete set of grammatical feature values may be encoded by two or more distinct morphemes" is an example of such a learning principle. Clark's (1987:2) *Principle of Contrast* stating that "every two forms contrast in meaning" is another example.

[33] This specification is aspectual and differs from the modal [-realised] specification proposed by Hoekstra & Hyams (1998). See Chapter 3 for more discussion on the feature [-realised].

[34] I did not consider imperatives, even though they may be examples of modal utterances that reveal early modal distinctions. Although the children in my sample use imperatives early (as early as stage II), the early imperatives are restricted to only a few lexical items: *kijk* ('look!') is used by all children very often to draw attention, sometimes *kom* ('come!') or *pas op* ('be careful!') is used. This restricted use does not change during the four stages that I examined. The restricted use itself and the absence of any lexical increase of imperative forms indicate that imperative forms are not productive yet.

Table 5.15: Schema of sequence of development of specified aspectual, modal and tensed forms in early child Dutch

Stage	Area	Form	Meaning
II	Aspect	RP	[+completed]
		RI	[-completed]
III	Modality	SF	[+realis]
		PV	[-realis]
≥ IV	Tense	SF-ø	[-past]
		SF-suffix	[+past]

This sequence of development is compatible with the Aspect-before-Tense Hypothesis (cf. Bronckart & Sinclair, 1973; Antinucci & Miller, 1976). However, the aspectual distinction between completed and uncompleted events does not only appear before the tense distinction between past and present events, but also before the modal distinction between realis and irrealis events; the observations suggest the order Aspect >> Modality >> Tense. The observed order could either be universal or language-specific. When the order of specification of the features [± completed], [± realis] and [± past] has an inherent logic in the sense that the one has to be derived from the other or the order is determined by cognitive maturity, the observed sequence might be universal. If the order of specification is dependent on the kind of encoding and children pick up the one kind of encoding (say auxiliaries) earlier than the other kind of encoding (say inflection), then the order of specification would be language-specific.[35] As far as I know, there is no immediate support for the universal perspective, while there are indications that properties of the packaging or encoding of inflection affect the ease and hence, rapidity of acquisition (Slobin, 1977, 1982). Given this, it may be hypothesised that for Dutch children syntactic cues are more salient than morphological cues and that (i) sentence-final placement of participle and infinitive is responsible for the early acquisition

---

[35] Note that the opposition between an auxiliary versus inflection is presumably too simplistic, as position of the verb as well as richness of inflection play a role as well.

of the completed-uncompleted distinction, and that (ii) the modal distinction between sentence-final and sentence-initial placement of the main verb is more prominent than the tense distinction between the suffixes [*ø*] and [*-de*]/[*-te*]. This issue must be addressed in future research.

5.3.4   Re-interpreting the experimental results

Chapter 3 included an examination of not only corpus data, but also experimental results. What does my conclusion that the MRE in Dutch RIs is a developmental effect mean for the experimental results in Chapter 3? At first glance, the experimental data are consistent with the developmental hypothesis: the modal predominance of RIs in the experimental results could very well be an effect of age/linguistic proficiency as the children in the experiment are relatively old (2;10 on average). Apart from age, this conclusion is confirmed by a simple test. As described in Chapter 2, the third stage (i.e. the stage of the modal shift) is characterised by children's use of approximately as many FINs as RIs. Before this stage, the number of RIs exceeds the number of FINs. After this stage, the situation is turned around and FINs are predominant. Only 8 subjects in the experiment used less FINs than RIs, whereas 18 subjects used as much FINs as RIs or even more FINs than RIs. Therefore, I conclude that the experimental results are representative of a fairly late stage within the RI-period.

In order to place the experimental results in the developmental picture, I will compare corpus results and experimental results. The experimental results from the Dutch subjects are repeated in Table 5.16. Table 5.17 contains an overview of the corpus data analysed along the same lines as the experimental data (the semantic approach, see § 3.4.10). Thus, first the numbers of utterances expressing a modal event and expressing an ongoing event are calculated. On the basis of this, the proportions of verb forms in these two contexts (modal and ongoing) have subsequently been determined. Note that the PVs in Table 5.16 contain (precursors of) PICs (see Chapter 3, § 3.2.2):

Table 5.16: Distribution of forms in modal and ongoing condition, experimental results for Dutch subjects (n = 26), data from Chapter 3

Form	#	Modal condition (n = 298)		Ongoing condition (n = 562)	
		#	%	#	%
RI	149	101	34	48	9
SF	487	42	14	445	79
PV	216	155	52	61	11
PIC	8	0	0	8	1

Table 5.17: Distribution of forms used for modal and ongoing meanings, corpus results from all six children (collapsed)

Form	#	Modal condition (n = 1517)		Ongoing condition (n = 1643)	
		#	%	#	%
RI	**1248**	**912**	**60**	**336**	**20**
SF	1221	93	6	1128	69
PV	691	512	34	179	11

A comparison between the two tables shows that the children in the corpus study used relatively more RIs and relatively less PVs than the children in the experimental study. Also, the relative number of SFs in the experimental study is higher than the relative number of SFs in the corpus study. What does this show? Relating the experimental results to the longitudinal patterns observed in this section, it becomes clear that most of the children in the experiment have already undergone the modal shift. Most of the ongoing RIs have been pushed out by specified finite alternatives (predominantly SF_θ's). Although children rely in the modal condition more on RIs than in the ongoing condition, there is, nevertheless, a noticeable effect of modal PVs: more obvious than in the corpus data, the modal RIs are pushed out by modal PVs (34 % RIs vs. 52 % PVs in the experiments

against 60 % RIs vs. 34 % PVs in the corpus study). This confirms the earlier conclusion that the subjects in the experiment are linguistically proficient and represent a relatively old age group. In all, I conclude that the comparison between experimental results and corpus results in the light of development shows that the modal predominance in the experimental results is consistent with the modal predominance in the corpus data. Both can be interpreted as effects of a modal shift and hence, as an effect of the sequence of development in Dutch child language.

5.3.5 Recapitulation

This section started with the introduction of the Grammaticalisation of Finiteness Hypothesis (GoF), which predicts a particular sequence of the development of FINs in Dutch child language based on a mix of previously reported findings and ideas on how knowledge of finiteness is built up incrementally. The GoF is confirmed: FINs develop from a stage in which finiteness is absent to a lexical-finiteness-stage, followed by a grammatical-finiteness- stage. In the lexical-finiteness-stage, lexical finiteness markers come in and subsequently, these lexical finiteness markers are combined with PVs. The lexical finiteness markers are unanalysed forms that primarily denote tense and modal meanings. In the final, grammatical stage, verbs are analysed and inflection comes in. The acquisition of inflection induces a generalisation over the early lexical finiteness markers and infinitives. Children re-analyse the forms and an adult-like category of verbs is acquired. This generalisation results in the acquisition of verb movement. Strikingly, exactly this sequence of development was found in the longitudinal data of six Dutch-speaking children. Moreover, it was shown that the observed sequence of development is imprinted in changes of RIs over time. The increase of SF_θ's and the overuse of ongoing PVs cause a modal shift in RIs; the modal shift leads to the MRE as it was described in Chapter 3. Additionally, it was argued that the NCC, that is, the exclusive use of RIs for incomplete events, could follow from the same condition as the MRE: the Elsewhere Condition. An alternative explanation led to the hypothesis that the aspectual feature [±completed] is specified before the modal distinction [±realis] is made; both these distinctions surface before Dutch children formalise the tense distinction [±past].

5.3.6   Evaluation of hypotheses

Briefly evaluating the impact of the above observations for two influential hypotheses about children's grammatical representations, I conclude that the incremental growth of knowledge is not in line with the maturational version of the Reduced Competence Hypothesis (RCH), because this hypothesis predicts an instantaneous development. The observations are consistent with the Full Competence Hypothesis (FCH). The FCH, however, does not contribute new insights, because this hypothesis does not explain small changes over time in the child data. In order to account for such changes, an additional explanation is needed. The uniformity of the observations in different stages within a sample of six children strengthens this conclusion. The observations confirm that children access grammatical knowledge in a piecemeal fashion. In the earliest stages, frequencies and distributions in the input play a crucial role (the no-overlap stage). Over time, children become independent of the input, overcome the misleading information in the input and get their language system to work on the basis of the knowledge they have extracted from the input. Children derive morphological knowledge on the basis of an accumulation of lexical items, whereas the acquisition of morphology guides the children to the syntactic operation of verb movement.

In the following section, I link the acquisition of inflection to a second syntactic change that takes place. Given that (i) inflection encodes tense and agreement features and (ii) agreement connects finite verb and subject, it is expected that the acquisition of inflection effects subject use.

5.4   The development of subject use

Like the drop of finiteness (resulting in RIs), the asymmetry between child and adult language with regard to the dropping of subjects is one of the topics that FCH studies have paid much attention to. According to the FCH, like adults, children are equipped with the Extended Projection Principle (EPP), which is the principle that states that all sentences have a subject. Therefore, it is expected that children use subjects in a way that is

consistent with UG principles. However, spontaneous speech data show that children leave out overt subjects more frequently than adults do. The general solution given by studies that advocate the FCH is that even though quantitative data show that the children drop subjects more often than adults do, *the way in which* children drop the subject is fully in accordance with principles of the (adult) grammar: children drop the subject where empty categories are licensed. This is: (i) when the verb is an infinitive (Krämer, 1993; Wexler, 1994; Sano & Hyams, 1994; Schütze & Wexler 1996; Schütze, 1997) or (ii) when CP is absent (Rizzi, 1992, 1994; Haegeman, 1995). Scholars consider their proposals to be confirmed when they find a correlation between subject drop and RIs and a similar decrease of RIs and subject drop over time.

In this section, it will be shown that there are phases in which children drop subjects in a way that is incompatible with the FCH. The data indicate that rule-governed subject drop is learned: the pattern of subject drop in FINs as well as RIs changes over time. Only in the final developmental stage, i.e. stage IV, can the patterns of subject drop be considered adultlike and hence, in accordance with adult grammar. The changes suggest that from onset the children do not know that subjects are required: they leave out subjects in RIs as well as in FINs. It will be argued that the acquisition of inflection plays a crucial role in the development of subject drop.

5.4.1 Predictions

In Chapter 1, Underspecification and Truncation Hypotheses have been discussed as exponents of the FCH. According to both accounts, no development of subject use *within* the set of RIs is expected. As subject drop is either dependent on RIs (underspecification of I) or correlated with RIs (absence of CP), subject drop and RIs go hand in hand. The Agreement and Tense Omission Model or ATOM (Schütze & Wexler 1996; Schütze, 1997) is the only theory that provides the possibility to model developmental patterns of subject use within RIs. According to the ATOM, subject drop in RIs follows from the underspecification of Tense, as the underspecified Tense provides a licit environment for the empty category PRO. According to the ATOM, RIs are the effect of underspecification of Tense and/or Agreement. Some RIs match the feature matrix [+tense, -agreement]. In

these RIs, Tense is specified: PRO is not licensed and hence, overt subjects are required. Taking the developmental view, we can deduce that when Tense becomes specified *before* Agreement, the proportion of overt subjects in RIs is expected to increase. I will come back to the ATOM in section 5.5. In this section, I discuss the ATOM in the context of the observed combination of developments in RIs, i.e. modal shift and subject use. Turning to the predictions for FINs, the Truncation model predicts a transition from a stage in which subjects are dropped in FINs to a stage with no (or less) subject drop in FINs (Rizzi, 1992, 1994; Haegeman, 1995). This transition is the effect of maturation of the rule 'CP = root'; projection of a C domain implies absence of a licit environment for the null subject in FINs and RIs. According to the Underspecification view, no development of subject use in FINs whatsoever is expected or predicted (Krämer, 1993; Wexler, 1994; Sano & Hyams, 1994).

5.4.2   Method

In addition to the modal codes (see § 3.3.2), the files have been annotated with codes about subject use. Subjects in all positions are included.[36] Vocative utterances, mentioning the addressee explicitly (*Jakob, eten!* 'Jakob, eat!'), are not counted as utterances with subjects. In the transcripts, these

---

[36] Haegeman (1995) makes a distinction between different subjects: initial null subjects in root clauses (ia), non-initial null subjects in root clauses (ib) and null subjects in embedded clauses (ic):

(i)	a.	NULL SUBJECT heb 't zo koud	Hein 2;9
		have it so cold	
		'I am so cold'	
	b.	in de creche # heef NULL SUBJECT dat gezien	Hein 2;9
		in the creche have seen that	
		'I have seen that in the creche'	
	c.	als NULL SUBJECT niet # bang is	Hein 2;5
		if/when not afraid is	
		'when it is not afraid'	

As I am interested in the correlation between absence of finiteness and subject drop (and not in specific predictions from the truncation hypothesis) I did not make this distinction.

could be recognised by a comma between the addressee and the rest of the utterance. Imperatives containing a bare stem (*Kijk!* 'Look!'), are excluded from the analysis as these forms do (usually) not contain a subject in adult Dutch. It is unclear if they can allow for grammatical subjects. The sentences that are included are RIs, SFs and PVs.

5.4.3 Results

The graphs 5.3-5.8 on the next page give the frequencies of omitted subjects in RIs and FINs. Recall that FINs are all sentences that contain a verb that is finite (SFs and PVs), i.e. placed in first or second position and/or carries inflectional morphology. The raw numbers are given in the Tables 5.18 and 5.19:

Table 5.18: Subject drop (null subjects = NS) in RIs, numbers per stage, data from all six children

	Abel		Daan		Josse		Laura		Matthijs		Peter	
	RI	NS	RI	NS	RI	NS	RI	NS	RI	NS	RI	NS
I	n.a.	n.a.	5	5	n.a.	n.a.	18	18	40	40	26	26
II	33	31	54	50	99	94	56	36	127	113	114	69
III	84	73	66	37	76	56	276	201	112	85	57	22
IV	42	35	71	58	64	55	65	55	51	46	29	22
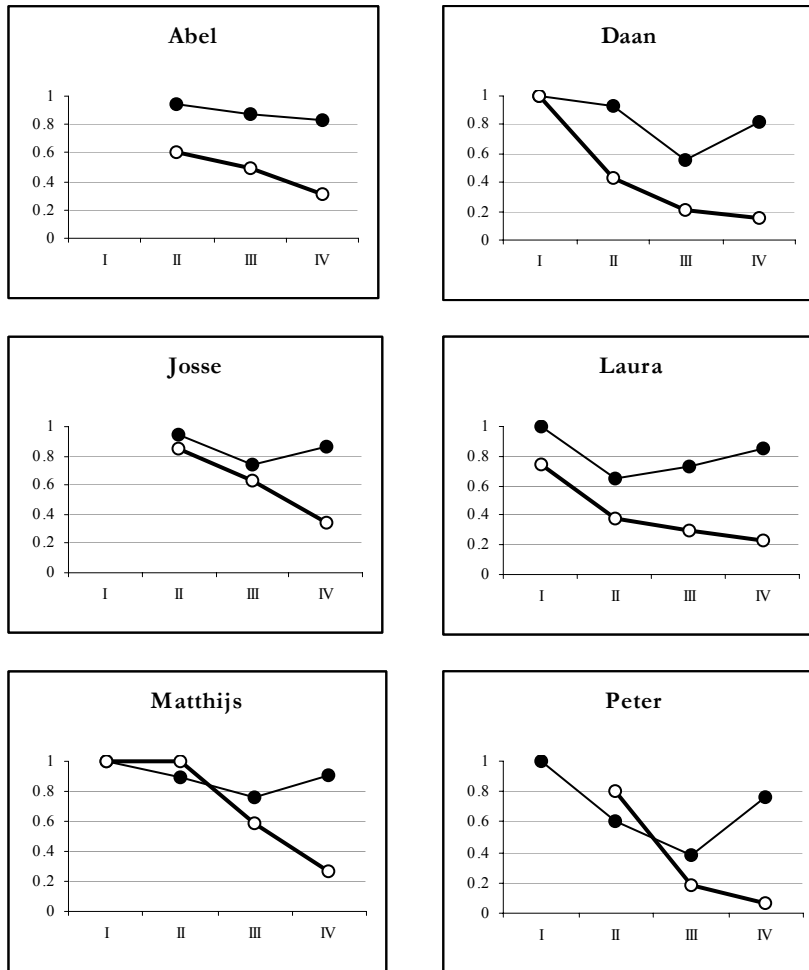
Figures 5.3 - 5.8: Percentages of subject drop in RIs and FINs four subsequent developmental stages, data from all six children
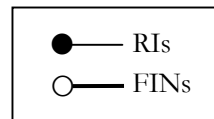
Table 5.19: Subject drop (null subjects = NS) in FINs, numbers per stage, data from all six children

	Abel		Daan		Josse		Laura		Matthijs		Peter	
	FIN	NS	FIN	NS	FIN	NS	FIN	NS	FIN	NS	FIN	NS
I	n.a.	n.a.	2	2	n.a.	n.a.	4	3	8	8	0	0
II	10	6	47	20	13	11	26	8	6	6	5	4
III	198	97	246	50	98	62	348	103	85	50	95	17
IV	275	85	621	92	437	152	372	87	524	139	627	41

I start by describing the most eye-catching developments, generalising across the six children. Later on, I point out some child-specific peculiarities. The development can be divided in two parts. In the *early* development, which continues until stage III, the proportion of subject drop in RIs and FINs decreases. In the *late* development, which starts after stage III, RIs and FINs show deviating patterns: the proportions of subject drop in RIs increase, while the proportions of subject drop in FINs continue their decrease. These deviating developments are most clearly visible in the final stage, i.e. stage IV: more than in any of the preceding stages, the children prefer to drop the subject in RIs and to use an overt subject in FINs. The graphs show that the development of subject drop in RIs can be modelled through a U-shaped curve. Laura's curve indicates that some child-specific variation occurs: Laura reaches the lowest point of the U, i.e. the moment at which overt subject use in RIs is at its peak, somewhat earlier than the other children. Given the uniform results of the other five children, it seems plausible that in Abel's case, the increase of subject drop in RIs still has to come and takes place after stage IV.

The question is whether or not the U-shaped development is statistically reliable. As before, this is tested with the permutation test. The null hypothesis states that all stages have equal average proportions of subject drop and there is no development. More precisely, the null hypothesis states that the proportions of subject drop observed over stages II, III and IV are drawn from a single distribution. Table 5.20 below contains the probabilities that the observed average differences between subject drop RIs in stage III

and stages II/IV are due to chance and the p-values when one of the children is left out of the analysis (Jack knife).

Table 5.20: Probabilities that the observed average difference between subject drop in stage III and subject drop in stages II/IV is due to chance. The probabilities are calculated over the entire group of six children and over 6 groups of five children, with one specific child omitted at a time

	p-value
All children	**0.001**
Omitted	
Abel	0.001
Daan	0.013
Josse	0.004
Laura	0.001
Matthijs	0.004
Peter	0.004

The conclusion is clear: the probability that the observed difference in proportion of subject drop between stage III and stages II and IV is due to chance is very small (p = 0.001). That is, the difference is highly significant. In addition, a single child does not disproportionately affect the probability. This means that we can speak of a U-shaped development of subject drop in RIs.

5.4.4   Interpretation of subject drop in RIs and FINs

The U-shaped development of subject drop in RIs indicates that between stage III and IV a reorganisation takes place. In this section, I will explain how this reorganisation falls in place (and is expected to occur) if we connect the observations regarding subject drop to earlier observations about the acquisition of inflection. In brief, the subject requirement in FINs

is in fact an antecedent-anaphor relation and follows from the anaphoric nature of agreement inflection in Dutch. By implication, the children need to have access to the agreement paradigm in order to find out that agreement in their target language is anaphoric and overt DP subjects (i.e. antecedents) are required (§ 5.4.5). Knowledge of the agreement paradigm will lead to identification of the infinitives as contrasting non-agreeing forms (§ 5.4.6).

5.4.5   Obligatory subjects in FINs

Starting with the patterns in adult Dutch, it can be observed that subject drop in finite main clauses is not allowed, apart from a number of constrained colloquial speech settings.[37] The requirement to use lexical subjects makes Dutch a so-called non *pro*-drop language. The generalisation that underlies that lexical subjects can be left unrealised in a language with rich agreement. When agreement is poor, lexical subjects are required (cf. Rizzi, 1982). For instance, the agreement paradigm of Italian expresses number as well as various person features. Italian exemplifies a rich paradigm and, therefore, allows for drop of pronominal subjects; see (20).

---

[37] Haegeman (1995) calls this *diary drop*. Van Kampen dubs the phenomenon *confession mode*. Under the analysis of Rizzi (1994) and Haegeman (1995) the examples of subject drop in Dutch exemplify topic drop. Haegeman (1995) gives the following examples to illustrate that the subject can only be dropped when placed in sentence-initial position. She argues that the subject is placed in the same extraposed position as the object in (ii), that is spec, CP:

(i)   NULL SUBJECT heb het al gezien
        have it already seen
        'I have seen it already'

(ii)   NULL OBJECT heb ik al gezien
        have I already seen
        'I have seen it already'

(iii)   * Dat boek ken NULL SUBJECT niet
        that book know not

Van Kampen (1997), however, stresses that subject drop in adult Dutch is restricted to first person subject drop, and hence, illustrates not just topic drop (because if it were topic drop such restrictions were not expected) but a highly marked register.

(20)   mangia una mela
        eat-*fin* an apple

Dutch is clearly more restricted in this respect, as shown in (21):

(21)   * eet een appel
        eat-*fin* an apple

Over the years, various proposals have been made to capture the relation
between drop of subject and richness of agreement. Recently, Koeneman
(2000) developed a theory that is empirically well-motivated and that
captures a wide range of languages. According to Koeneman, rich
agreement must (at least) encode the following three binary features:
[αspeaker], [αaddressee] and [αsingular].[38] If there is no evidence for
one of these three features in the agreement paradigm of a language,
agreement in this language is poor according to Koeneman's proposal.
Furthermore, poor agreement is non-argumental and cannot function as a
full DP subject. Rather, poor agreement is anaphoric and requires a full DP
subject as antecedent.[39] According to this definition, agreement in Dutch is
poor: there is not enough evidence to postulate the feature [αaddressee].
In Figure 5.9, Koeneman's hierarchical interpretation of the Dutch regular
present tense paradigm is given:

---

[38] The advantage of this system is that the entire paradigm is described with three binary
features. When the agreement paradigm is described in terms of person and number, there is
a feature that has three values (person) and a feature with two values (number). The features
[αspeaker] and [α addressee] are needed to distinguish first person and second person from
third person forms.

[39] Note that this is syntactic anaphoricity, relating to a variable that must be locally bound.

$$[\;\alpha\text{sg},\,\alpha\text{sp}\;]$$

$$[\;+\text{sg},\,\alpha\text{sp}\;] \qquad\qquad [\;-\text{sg},\,\alpha\text{sp}\;]$$

$$[\;+\text{sg},\,+\text{sp}\;] \qquad [\;+\text{sg},\,-\text{sp}\;]$$

$$[\;\text{-}\varnothing\;] \qquad\qquad [\;\text{-}t\;] \qquad\qquad [\;\text{-}en\;]$$
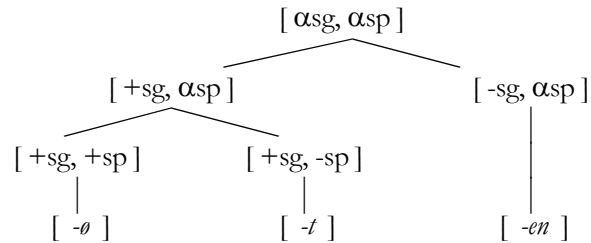
Figure 5.9: Agreement in the Dutch present tense paradigm, according to Koeneman (2000)

From this proposal, it follows directly that children must search the input and find evidence for the three features before they determine if an overt subject is required.[40] In section 5.2, it was concluded that the onset of the acquisition of inflection is between stages III and IV. Thus, before stage III, the children will not know that FINs need overt subjects. Hence, it is expected that the children 'overdrop' subjects in FINs prior to stage III (from an adult point of view). Given the curves of FINs in the graphs 5.3 – 5.8 (most clearly in the data of Josse, Laura, Matthijs and Peter), this expectation is borne out.

Proposals along similar lines, i.e. according to which the acquisition of overt subject use is related to the acquisition of inflection, can be found in Jaeggli & Hyams (1988) and in the work of Guilfoyle & Noonan (1992). In the literature, there is another kind of explanation for the overdrop of subjects by young children. According to this hypothesis, subject drop that seems - at first sight – to be banned by the grammar, is analysed as being topic drop. As Dutch is a language that allows topic drop, subject drop that is topic drop is grammatically legitimate (De Haan & Tuijnman, 1988). The

---

[40] Many more null subject/pro-drop proposals imply that it is required for the children to have knowledge of the inflectional paradigm in order to find out the properties of the target language. In this respect, Koeneman's proposal must be viewed in the line of a tradition that contains many proposals with some small differences. I have chosen Koeneman's proposal as it has a large empirical coverage and is clarifying. The proposal, moreover, exemplifies a structure building view and a very close relation is proposed between morphology and syntax. With regard to language development the influence of the acquisition of morphology surfaces directly in a growing phrasal marker.

implication of this analysis is that children overlicense topics. It is argued that this overlicensing is caused by deficits in pragmatic knowledge (Bromberg & Wexler, 1995). The study carried out by De Cat (2002), however, suggests that this alternative explanation, which analyses subject drop in early child language as topic drop, does not hold. According to De Cat, who examined topic drop in early child French, it is unlikely that children misanalyse certain subjects as topics. On the contrary, already in the earliest stages that De Cat looked at (which are the earliest stages in syntactic development), the French children seemed to know the properties of topics.

5.4.6   Optional subjects in RIs

The increase of overt subjects in RIs is an effect of the acquisition of inflection. (How) does this relate to the U-shaped development of subject drop in RIs? In order to answer this question, I will first turn to the final stage. What are children expected to learn with regard to subject drop in RIs? If we turn to RIs in adult Dutch, it can be concluded that here, subject use is optional. More specifically, there are cases in which the subject is truly optional, cases in which subjects must be omitted or cases in which subjects have to be present. In (22) and (23), examples are given in which the subject is optionally present (the bold-faced part is the RI). In (24), an example of obligatory subject use, and (25) exemplifies obligatory subject drop in RIs.[41] Following the common analysis of subjects in infinitival clauses, I notated the dropped subject as PRO or 'big pro' (Chomsky, 1986; Haegeman, 1992).

(22)   **Parel/PRO altijd bij ons blijven?** Dat zie ik al gebeuren!
       Pearl/PRO always with us stay-*inf*? That see I already happen!
       'Parel forever stay with us? I don't wanna see that happening!'
                                                                    (UTVM)[42]

---

[41] In the first optional examples in (22) the subject was present in the original sentence, while it was absent in the original sentence in (23).

[42] From the recommendable Dutch novel *Uit talloos veel miljoenen*, written by W.F. Hermans

(23)   Zat ik naar 'Het kleine huis op de prairie' te kijken…**Zij/PRO**
        brullen!
        sat I to 'The little house on the prairy' to watch … she/PRO weep-*inf*
        'I was watching the 'little house on the prairy', when she suddenly
        started to weep'   (Magazine)

(24)   Hij/*PRO achter ons aansjokken
        he/PRO behind us walk-*inf*
        he came slouching behind us   (UTVM)

(25)   Wacht, *ik/PRO de wekker proberen voor morgenochtend
        wait, I/PRO the alarmclock try-*inf* for tomorrow morning
        'Wait, I am going to try the alarmclock for tomorrow morning'
                                                                (MVH)[43]

I will no go into the rules – assuming that there are rules - that underlie the
patterns in (22)-(25) above; for now, the relevant observation is that RIs in
adult Dutch do not require subjects in the way FINs do. Given that the
infinitive does not have any agreement features, infinitival morphology is
not anaphoric and does not require a lexical subject as antecedent. As
knowledge of inflection, in particular of agreement, comes in between stages
III and IV, it is not expected that children are able to make a systematic
distinction between inflected forms (FINs) and infinitives (RIs) prior to
stage IV. When children indeed acquire inflection between stages III and
IV, as we concluded earlier on the basis of paradigmatic/lexical variation
and inflection errors, then it is expected that the children make this
distinction between stages III and IV. More specifically, it is predicted that
between stages III and IV, RIs and FINs start to show a deviating pattern
with regard to subject drop. This expectation is borne out. The conclusion is
that the observed systematic reorganisation of subject use is the effect of (i)
the children's specification of agreement inflection in Dutch as
[+anaphoric], and (ii) the ability to distinguish between agreeing finite
forms and the non-agreeing infinitive, that is, the identification of the
infinitive as a form that is not anaphoric and does not require a DP subject
as antecedent like FINs do.

---

[43] From the Dutch novel *Meneer Visser's hellevaart*, written by S. Vestdijk.

5.4.7   Early differences between FINs and RIs

A closer look at the graphs in 5.3 – 5.8 shows that the children seem to drop subjects more often in FINs than in RIs, even in the earliest stages. This could indicate that children make the grammatical distinction between RIs and FINs earlier than I claimed it to be made. There are, however, indications that the early distinction (before stage IV) has a different cause. I illustrated this in (26):

(26)   a.   dies **dieis** gogel   Abel 1;11.26
            that-is that-is bird
            'That is a bird'

       b.   **weejdik**ook   Daan 2;00.22
            want-I-also
            'I also want to have that

The sentences in (26) exemplify clitisation between finite verb and subject.[44] This kind of clitisation takes place in FINs but not in RIs. The difference between FINs and RIs follows from properties of the input. Due to Verb Second, finite verb and subject are adjacent in Dutch. By implication, subject and finite verbs are candidates for clitisation. For the children the adjacency (and clitisation) of finite verb and subject poses a segmentation problem: it is not clear where the subject ends and the verb begins. This results in forms like *kwil* (*ik wil* 'I want'), the inverted version *willik* (*wil ik* 'want I') or *tis* (*dit is* 'this is'), as illustrated in (26) above. In the case of infinitives, the segmentation problem does not occur: due to the SOV base order in Dutch, subject and infinitive are not adjacent. It is expected that this difference between subject-verb clitisation between finite verb and infinitive in Dutch will lead to relatively many overt subjects in FINs, especially in the early developmental stages when children are still figuring out the segments on a syntactic level, rather than on morphological level.

---

[44] Note that this kind of clitisation is difficult to study systematically with corpus data because there is a strong influence of choices made by the transcribers of the corpus data in this respect.

5.4.8   Conclusion

In this section, a second syntactic effect of the acquisition of inflection has been discussed: the appearance of rule-governed subject use, i.e. obligatory subject use in FINs and optional lexical subjects in RIs. The developmental patterns of subject drop in the data of Abel, Daan, Josse, Laura, Matthijs and Peter point to a rather radical change between stages III and IV. Before this moment, subject drop in RIs and FINs followed a similar type of development (and the small differences that appeared could be related to a phonological difference), but after stage III, subject drop in FINs continues its gradual decrease while subject drop in RIs suddenly starts to increase again. This change co-occurs with the acquisition of inflection, as was concluded in section 5.2. In this section, I proposed that this co-occurrence is not accidental but follows from the establishment of an anaphoric relation between agreeing verb and lexical subject. Consequently, the acquisition of inflection leads to the identification of infinitives as non-agreeing verbs and hence, as forms that do not grammatically require an antecedent. Thus, when inflection comes in, the patterns of subject drop in RIs and FINs start to diverge.

5.5   An alternative scenario for changes in RIs

RIs undergo two changes. I argued that the changes are correlated by a third factor, namely the rise of FINs. In other words, they both follow from new knowledge about finiteness that children collect in different stages. In this section, I will turn to a scenario that might have been an alternative if it had not made the wrong predictions for the course of development. This scenario is based on the Agreement and Tense Omission Model, or ATOM (Schütze & Wexler, 1996; Schütze, 1997) that was explained in Chapter 1.

What goes wrong in the ATOM? According to the ATOM, RIs follow from an underspecified grammatical system, more specifically from the underspecification of two functional categories that are represented in morpho-syntax by inflection: Agreement and/or Tense. Underspecification of one of the two categories suffices for the infinitive to appear. Thus, RIs

correspond to sentences that are either specified for [+agreement, -tense], [-tense, +agreement] or [-tense, -agreement]. When both Tense and Agreement are specified, a finite verb is selected. According to Schütze (1997: 232), [+tense] RIs are restricted to present tense. These RIs are bound by tense to speech time, which is the default time as long as the children cannot make a distinction between present and past. Unlike the tensed RIs, [-tense] or untensed RIs are not bound to a specific point in time, hence the untensed RIs must be modal. Thus, theoretically speaking, the ATOM can model the modal shift in RIs as the effect of an increase of [-tense] RIs, implying that RIs become underspecified. The second change over time concerns subject use. Theoretically speaking, the ATOM is also capable of modelling this development. According to the ATOM, the difference between RIs in which either Tense or Agreement is underspecified is noticeable in subject use in RIs. In RIs that are underspecified for Agreement, the overt subject receives default Case, as Agreement is taken to be responsible for Case. RIs that are [-tense] provide a licit environment for PRO, as Tense is claimed to be responsible for subject licensing. It follows that specification of Tense leads to the appearance of RIs with overt subjects (and default Case). By implication, the emergence of a clear preference for subject drop implies that Tense becomes underspecified.

In sum, the ATOM can explain changes in the use of RIs as well as changes in subject drop. However, according to the ATOM the two observed developments imply a development from [+tense] RIs to [-tense] RIs, which would be a reverse or backwards development: a specified stage is followed by an underspecified stage. This sequence of development contradicts the basic assumption that underlies the ATOM and other underspecification accounts, namely the idea that the underspecified stage is the initial stage. Therefore, I conclude that, the ATOM is inconsistent with the observed changes of RIs.

5.6 A note on methodology

I want to finish this chapter with a final remark on methodology. In this chapter, I have motivated claims about the development of children's

competence by looking at performance data. Methodologically, this is tricky. However, I only drew conclusions if observations could be generalised over a set of six children. This minimises the risk of drawing far-reaching conclusions based on performance noise considerably. Moreover, embedding my claims into a theoretical framework that relates phenomena via grammatical rules leads to predictions about co-occurring changes that result from grammatical development. When the changes indeed co-occur in the predicted way, it can be taken as evidence that the production data are reliable. The more such co-occurrences can be formulated beforehand and the more they are actually found in the data, the more reliable these data can be considered to be. In this respect, this chapter illustrates how language production data can be telling with regard to language competence.

5.7 Summary

This chapter has stressed the value of longitudinal research. A detailed investigation of changes over time reveals that children gradually build up knowledge of finiteness (the Grammaticalisation of Finiteness Hypothesis). In Dutch, the grammatical encoding of finiteness is not immediately accessible as the inflectional paradigm is rather poor and there is only marginal lexical overlap between finite verbs and infinitives in the input. This leads to a sequence of development in which finiteness is initially absent and finiteness first comes in in the form of lexical elements. Only later on, is finiteness re-analysed as a grammatical category realised by inflection. The acquisition of inflection introduces verb movement as an operation that generalises over sentences. It was argued that the modal shift in RIs that takes place in stage III is an effect of the stepwise acquisition of finiteness, more specifically, of the more frequent use of tensed forms than of modal alternatives for RIs in stage III. A factor that plays a role in this respect is the overuse of lexical finiteness markers in the stage that precedes the acquisition of inflection. The modal shift in RIs explained the observation that Dutch child RIs display a MRE, that is, that they are predominantly used to denote modal events. The syntactic effect of the acquisition of inflection is two-fold. Verb movement is induced by the tense features of inflection, whereas the agreement features of inflection lead to the acquisition of rule-governed subject use. This second effect shows up in

a comparison of the development of subject drop in FINs and RIs: the children that were investigated start to make a clear distinction between subject drop in FINs and RIs when they have knowledge of inflection, that is between stages III and IV. Relating the observations in this chapter to the structure building view on grammar, I concluded that children are not able to move the verb from start, nor can they project functional structure. Rather, they need to have knowledge of inflection first in order to be able to take this syntactic step. Children have knowledge about universal dominance relations between phrases, but they have to detect how the properties of language they hear can be brought back to this universal basic sentence structure. This is achieved by extracting structural information on the basis of growing lexical knowledge.

CHAPTER 6

Summary and Implications

This dissertation dealt with the acquisition of finite sentences, more specifically, with the acquisition of auxiliaries and verbal inflection. Normally developing Dutch-speaking children acquire the basis for the auxiliary and inflectional system of their target language when they are between ages of two and three years old. The first step in this development is the appearance of the so-called root infinitives, commonly abbreviated as RIs. These sentences contain a verb, i.e. a potential carrier of finiteness features, but initially this verb appears exclusively in non-finite form. Sentences such as *Bob op bank zitten* ('Bob on sofa sit-*inf*') or *Schoen aantrekken* ('Shoe on put-*inf*') are characteristic examples of RIs: in these sentences a finite auxiliary (*Bob moet op de bank zitten* 'Bob has to sit on the sofa') or finite inflection (*Bob zit op de bank* 'Bob sits on the sofa') is lacking. After the first non-finite stage, children pass a number of stages in which finite sentences and RIs are used alongside each other. A recurring question in this thesis is what the status of RIs in early child speech is. The answer to this question is of importance for determining the knowledge children have in the initial stage and for interpreting changes in the expression of finiteness that take place later.

The first chapter provided an overview of the literature on RIs. The second chapter focused on the statistical methods that were applied. In the remaining chapters, spontaneous speech data from six Dutch-speaking children were analysed. In Chapters 3 and 4, data from subsequent developmental stages were collapsed in order to test a number of hypotheses. Chapter 5 dealt with the small steps children take from one stage to another. In this thesis, patterns in Dutch child language have been compared to patterns in other languages; in particular to child English. In Chapter 3, results from an experiment with Dutch and English two- and three-year-olds were presented.

6.1   Summaries of the chapters 3, 4 and 5

6.1.1   Temporal, modal and aspectual denotation

RIs lack all overt expressions of tense, modality or aspect that adult sentences characterise. How do we assign this sentence a temporal, modal or aspectual interpretation? In the literature, three different answers to this question have been given. According to the No Tense Hypothesis, RIs are untensed and can receive all temporal interpretations: past, present and future (Behrens, 1993; Wijnen, 1997; Lasser, 1997). According to the Modal Hypothesis, RIs contain structural features that give rise to a modal interpretation. The modal features are either represented by a null auxiliary (Ferdinand, 1996; Ingram & Thompson, 1996) or by infinitival morphology (Hoekstra & Hyams, 1998). The third hypothesis states that RIs obey a Non-Completedness Constraint (NCC) and do not denote completed events (Lasser, 1997).

Naturalistic language data from six normally developing Dutch-speaking children showed that future, or modal RIs are predominant. The number of RIs that denote an event that took place prior to speech time is negligible. Hence, RIs in early child Dutch display what has been termed a Modal Reference Effect (MRE). Non-modal RIs are too frequent to confirm the Modal Hypothesis, however. The No Tense Hypothesis and the NCC are confirmed because there is no fixed temporal reference and past/completed RIs are very sparse. An empirical argument for one of the Modal Hypotheses, namely Hoekstra & Hyams' (1998) Infinitival Morphology Hypothesis (IMH), is based on a comparison between corpus data from Dutch and English-acquiring children. These suggest that English RIs do not display a MRE. According to the IMH, this is due the absence of infinitival morphology in English RIs. The outcome of an experiment with Dutch and English two and three-year-olds, presented in Chapter 3, confirmed the observation that Dutch RIs are more frequently modal than English RIs, but the strong predictions from the IMH are not borne out.

The remainder of Chapter 3 concentrated on the question how to understand the experimental results. First of all was the difference in interpretation between Dutch and English RIs considerably smaller than in previously reported corpus data. How could the difference between the two research

methods be explained? I pointed to an asymmetry between the Dutch and English corpus data: in the set of English RIs, utterances with first and second person subjects are excluded, whereas they are not in the set of Dutch RIs. Since first and second person subjects correlate with modality ('*I* want' and '*You* must'), many modal utterances are left out in the set of English RIs. Re-analysis of the Dutch data confirmed this hypothesis: exclusion of RIs with first and second person subjects from the set of Dutch RIs led to a significant decrease of the proportion of modal RIs in this set. The difference between Dutch and English corpus RIs became smaller and, therefore, more comparable with the observed difference in the experiments.

Secondly, how could the observed difference in meaning be accounted for, given my earlier conclusion that the Modal Hypotheses, and hence, the IMH, are untenable? As an alternative explanation, it was argued that the Heterogeneous Set Effect (HSE) occurs in English but not in Dutch. English RIs present a heterogeneous set because they contain 'real' non-finite RIs and finite sentences in which the inflectional morpheme of the finite verb is not realised (which can be the result of either drop or overgeneralisation of other paradigmatic forms). Dutch corpus data confirmed the hypothesis that children tend to overuse bare stems in finite contexts. In Dutch, as opposed to English, such sentences do not end up as RIs, because of SOV/Verb Second and infinitival morphology in Dutch. This explanation was supported by the following independent observations: English RIs contain relatively many stative predicates, allow for topicalisation and contain WH-words. It could not be concluded whether or not the HSE accounts sufficiently for semantic differences between Dutch and English RIs. Studies report deviating results with regard to the size of the difference and the size of the HSE in English is unknown.

By the end of this chapter, two questions were not yet answered: (i) why are Dutch RIs predominantly modal and display an MRE, and (ii) why are Dutch RIs not used to denote completed aspect and obey the NCC?

6.1.2 Types of verbs

De Haan (1987) discovered that Dutch children use different types of verbs in RIs and SFs. RIs contain predicates that denote activities, whereas SFs primarily contain auxiliary-like predicates such as modals, copula and aspectual

verbs. More recent interpretations of De Haan's observation describe this dichotomy between RIs and SFs as an eventive-stative distinction (Ferdinand, 1996; Wijnen, 1997; Hoekstra & Hyams, 1998) or as a [±telic] versus [αtelic] asymmetry (Gavruseva, 2001). The fourth chapter started with a test to determine which of the generalisations suits the Dutch data. I concluded that there is evidence for a weak version of the Eventivity Constraint (EC), that is, both SFs and RIs can be stative as well as eventive, but RIs are significantly more often eventive than SFs.

In an attempt to explain the EC, I disentangled three relevant factors: early obedience to semantic selection restrictions, cognitive immaturity and patterns in the input. An inventory of input patterns showed that state-denoting predicates appear predominantly as finite inflected forms in the input, whereas event-denoting predicates are most often sentence-final infinitives. Given that Dutch children are able to derive novel infinitival forms only about halfway the RI-period (empirical evidence for this claim has been provided in Chapter 5), the verb form-verb type correlations in children's production data must reflect verb form-verb type correlations in the input. Semantic selection restrictions play a role in RIs that are used by the children to give commands. Stative predicates in commanding utterances violate selection restrictions, since the former are specified for [+control], whereas the latter are [-control]. Finally, cognitive immaturity surfaces in the absence of RIs that express intended states. The interaction of state and intention triggers an epistemic modal reading. On the basis of results from various production studies and comprehension tasks, it was concluded that children in the RI-age do not have access to epistemic modality. The multiple-factor account for the EC presented in Chapter 4 leads to the prediction that stative RIs denote either states at speech time or desired states. This prediction is borne out.

6.1.3   Changes over time

In the chapters 3 and 4, data from the six children collected over a longer period of time (approximately one year) were collapsed. In the fifth chapter, longitudinal analyses have been performed on these data. It turned out that during the RI-period, finite sentences come in and become more frequent over time. Initially, all overt reflections of finiteness are absent: there are no inflected main verbs or auxiliaries. The first finite forms to appear are modal verbs and the copula *is* 'is'. These forms do not carry inflection and denote modality and

present tense. Absence of inflection and restriction of finiteness to a small set of verbs that have little lexical content, led me to the conclusion that finiteness starts out as a lexical feature in child Dutch. The lexical-finiteness-stage consists of two substages: the first substage is characterised by simplex lexical markers of finiteness, the second substage by periphrastic verbs in which lexical finiteness markers appear in combination with infinitives. It was argued that a lexical growth of finite verbs provides children with the prerequisites for verb-form-analysis. The emergence of inflection errors was taken as conclusive evidence for the acquisition of inflection. On the basis of verb movement theory, it was concluded that the morphological generalisation over verbs, introduced by inflection, triggers verb movement in the child grammar as a strategy to economise the syntactic representation and derive different sentences from one basic structure. In sum, the findings confirmed the Grammaticalisation of Finiteness Hypothesis.

Inflection is acquired fairly late by Dutch children, that is, after the first finite sentences appear. This was argued to be the effect of (i) homomorphemes and zero-affixes in the Dutch inflectional paradigm, (ii) of relatively little lexical variation within the inflected forms in the input, and (iii) of relatively little lexical overlap between inflected forms and non-inflected infinitives in the input. Examination of finite forms and infinitives, showed that lexical overlap emerges only in the later developmental stages. This observation added a developmental dimension to the results from Chapter 4.

The longitudinal perspective of Chapter 5 provided an explanation for the two unexplained observations from Chapter 3: the MRE and the NCC. Under the assumption that the Elsewhere Condition is working from the beginning in children's lexicon, the infinitive in early child Dutch is a vulnerable form: the infinitive will only be frequent as long as children have not yet learned specified tensed and modal alternatives for the unspecified infinitive. Driven by the Elsewhere Condition, the increasing number of finite forms causes the gradual pushing out of RIs from the children's repertoire. This process explained the Modal Shift in RIs, i.e. the observation that RIs are in later developmental stages more often modal than in early developmental stages. This hypothesis is supported by empirical findings: it was found that eventive verbs forms with a present tense denotation are earlier productive than modal eventive verbs forms. One reason for this effect is children's overuse of periphrastic verbs in present tense contexts during the lexical-finiteness-stage. The alternative

explanation for the MRE has two advantages (in comparison to the Modal Hypotheses from Chapter 3): non-modal RIs as well as changes over time in the temporal/modal denotation of RIs can be captured.

The NCC could also be explained as the result of obedience of the Elsewhere Condition: completed alternatives for RIs are acquired very early in the form of root participles (RPs). A second interpretation of the observations has been given, however. According to this view, RIs and RPs denote aspectual contrasts: the verb form in RIs is specified for [-completed], whereas RPs contain a [+completed] verb form. This second view is consistent with the hypothesis that children are very early sensitive to aspectual distinctions. On the basis of the data from the six children, the order of appearance of formal distinctions in early child Dutch would be: Aspect >> Modality >> Tense. During the RI-period, modal distinctions expand. In the earliest stages, only forms that denote necessary events were found. The acquisition of periphrastic verbs introduces forms for possible events.

The remainder of the Chapter 5 dealt with changes in subject use. Inflection in Dutch does not only contain tense features, but also agreement features: the finite verb agrees in person and number with the subject. On the basis of a recently developed theory, it was tested whether the acquisition of inflection has an effect on subject use. It was also tested whether the patterns of subject use confirm the moment of the acquisition of inflection as it was determined earlier in Chapter 5. It turned out that the developmental patterns of subject drop could be directly related to the acquisition of inflection. The conclusion was that two syntactic properties of Dutch are triggered by the acquisition of verbal inflections: verb movement and rule-governed subject use.

6.2 Evaluation of previous research

CHAPTER 3 The Modal Hypotheses (Ingram & Thompson, 1996; Ferdinand, 1996; Hoekstra & Hyams, 1998) predicted exclusive modal use of Dutch RIs. Corpus as well as experiment results showed that Dutch children use RIs for modal and ongoing/present tense events, but hardly for completed/past events. RIs that denote events ongoing at speech time (i.e. present tense) cover one quarter of all interpretable RIs. Hence, the Modal Hypotheses were not confirmed, unlike the No Tense Hypothesis (Behrens,

1993; Wijnen, 1997) and the NCC (Lasser, 1997). The predominance of modal RIs confirmed the claim the Dutch RIs display the MRE. Experimental findings confirmed that RIs in Dutch and English child language have a different meaning: English RIs do not show the MRE.

CHAPTER 4   I did not find any support for the absence of lexical overlap between the verb types in SFs and RIs during the RI-period (De Haan, 1987; Ferdinand, 1996) nor for the Telicity Hypothesis (Gavruseva, 2001). Provided that the EC is interpreted as a tendency and not as a true constraint, the results support the existence of an EC in early child Dutch (Ferdinand, 1996; Wijnen, 1997; Hoekstra & Hyams, 1998). Finally, the No Tense Approach (Wijnen, 1997) turned out to be a less adequate model for the EC than the Modal Approach (Ferdinand, 1996; Hoekstra & Hyams, 1998) or an input-driven approach (Schlichting, 1996; Pine, Lieven & Rowland, 1998).

CHAPTER 5   I found evidence for the existence of a no-overlap stage, followed by a stage in which finite verbs and infinitives show lexical overlap (Wijnen, 2000). The findings suggested furthermore that the acquisition of inflection correlates with the acquisition that Dutch is a non *pro*-drop language. This is in line with work from Jaeggli & Hyams (1988) and Guilfoyle & Noonan (1992). Results regarding the stepwise development of finiteness are inconsistent with the maturational version of the RCH (Lebeaux, 1988; Radford, 1988; 1990). The results are compatible with the FCH (cf. Poeppel & Wexler, 1993, amongst others). The scope of the FCH is limited with respect to developmental issues, however, and the hypothesis does not contribute to our insights into the process of acquiring language.

I want to close this section with a remark about the implications for the theoretical controversy with which I began the first chapter: De Haan (1987) versus Poeppel & Wexler (1993). De Haan argued that children have to acquire grammatical knowledge while Poeppel & Wexler defended the FCH and claim that children have access to grammatical knowledge from early on. Both positions were empirically motivated with data from respectively the Dutch boy, Tim, and the German boy, Andreas. Considering the developmental path that Dutch children take, the data from these two children represent two different developmental stages. The no-overlap pattern of Tim indicates that he is in stage II. However, Andreas must be well beyond this stage. In the file that is studied, Andreas uses 231 finite sentences and 51 RIs. Thus, finite

sentences are clearly predominant. There are more indications suggesting that Andreas is a quick learner. He does not only use many finite predicates but he also uses many *different* finite predicates. Considering the observations in this thesis, this points to fairly advanced knowledge. In addition, Andreas drops relatively few subjects. According to Poeppel & Wexler, in a set of 197 finite sentences, 180 sentences have overt subjects (ca. 91%). In a longitudinal study of subject drop in child Dutch, Haegeman (1995) shows that the curve of subject drop (in finite sentences and RIs) resembles the curve of RIs. They both show a similar decrease around the same time. If the developmental curve of subject drop in German child language is like Haegeman's curve for Dutch, Andreas' behaviour fits a rather late developmental stage. Even if all of his 51 RIs contain dropped subjects, he generally uses overt subjects (that is, in 73 % of the cases). Based on these comparisons, I think it is justified to conclude that Poeppel & Wexler's data do not contradict De Haan's data. On the contrary, the data reported in both studies are compatible with the developmental pattern described in this thesis. Tim's data represent the developmental stage that precedes the stage that is compatible with Andreas's data.

6.3 Implications for future research

The implications of the present study for future research cluster around two topics: the development of experimental methods for studies on the relation between verb forms and meanings and longitudinal studies of the development of inflection and auxiliaries.

6.3.1 Experimental methods

Chapter 3, § 3.4, stresses the difficulty of testing young children. The final experimental design that I used did not overcome all problems. With respect to the experimental method, I assumed the following. If the action takes place in the here-and-now, the child utterance denoting the action received an ongoing present tense interpretation. If the action does not take place in the here-and-now or did not take place in the past, a modal interpretation was assigned. Thus, the situation at speech time is always taken as the situation of reference. However, even though young children seem to speak very much about the

here-and-now, it is not excluded that they refer to a situation that is disjoint from the here-and-now. Consequently, a wrong interpretation is assigned to the child utterance.[1] Because of this problem, I did not make any strong claims on the basis of only the experiment. Since other studies that used different methods (and hence, suffered from other problems as, for instance, the comprehension task as described by Schönenberger et al., 1995), report similar observations, I concluded that my experimental results contribute to the robustness of the observation that Dutch and English RIs differ in meaning. It must be emphasised, however, that the experiment is exploratory and that much can be improved with regard to its design.

In Chapter 5, it was shown that age-effects influenced the experimental results. Thus, one of the implications for future research is that we must think about an experimental setting that does not have the age/proficiency bias and enables a comparison between children from different stages in the RI-period: a stage in which RIs do not have any specified equivalents yet (stage I/II) and a stage in which there are specified equivalents available (stage III/IV). For early child Dutch, it is expected that the two groups use RIs for different meanings: the first and younger group will use them for modal as well as ongoing activities, whereas the second and older group are expected to use other forms than RIs for ongoing activities and RIs still for modal activities. However, to develop an experiment that is also suitable for young children, we first have to figure out which problems arise with the younger age-group.

The setting itself may be problematic. Furthermore, the modality involved in the experiment may pose a problem. With regard to the situation, the presentation of movies on a laptop can be distracting or imposing. Both effects are undesirable.[2] Embedding modal and ongoing conditions in the children's daily pattern, i.e. in every-day life, may provide a situation that has less effect on whether or not young children can be tested. How can this be achieved? The children that have been tested went to daycare centers where they usually followed a strict every-day-routine. This routine creates the opportunity to find

---

[1] Note that the same assumption underlies the corpus method.

[2] To minimise the distracting effect, we covered the keyboard with a flat cardboard box so that the children could not press buttons or see lights flickering. This helped a lot, but was not always sufficient as there were children that wanted to play with the mouse or wanted to touch the screen.

slots during the day in which one activity is ongoing and another activity is modal, such as the 15 minutes before lunchtime or before the children go outside to play. Such an "experiment" - or rather a means to control naturalistic speech - is quite similar to the fairly uncontrolled design that I used. Advantages would be that the items as well as the situations are familiar and not artificial; less clarification and explanation would be needed. By implication, the situation is less demanding than the original experimental setting. As for the modality involved in the experiment, the discussion in Chapter 4 (§ 4.5) suggests that intentions and desires can be better understood by two- and three-year olds if they are their own (i.e. first person) intentions/desires than if they are someone else's intentions/desires; the concept of desires/intentions of two- and three-year olds seems to be limited, i.e. non-representational, and does not go beyond 'Me in the here-and-now'. Thus, a modal condition expressing intentions/desires has to utilise first person intentions/desires. Note that this restriction does not seem to apply to deontic modality. The children I studied use deontic RIs most frequently to give commands to others.

It would be desirable to elicit speech in a production task and, additionally, to also test comprehension. A comprehension task does not suffer from the critical assumption I mentioned earlier, since the child itself is the interpreter. Schönenberger et al. (1995) carried out a comprehension experiment, but they tested only a few subjects. Moreover, their results showed a strong ongoing bias as has been discussed in Chapter 3. On the basis of pilots that I carried out, I concluded that the children tended to focus on the activity that was expressed and that they ignored the variable modal/ongoing. This was evident in children's selection of pictures that display the activity, i.e. the ongoing picture. A factor that probably made the modal condition difficult for the children and which lead to the ignoring of the modality in the modal triggering sentences, was that modality was expressed in the experiment as third person intentions and desires. The experiment could be improved by using deontic modality, as this kind of modality does not seem to have the limitations that desires/intentions have. Furthermore, it may be worth trying to use an eye-tracker-experiment instead of a picture-pointing task. As the situation is less demanding, it may be a useful strategy, especially for the younger children.

6.3.2 Longitudinal research

Chapter 5 illustrated the impact of longitudinal research. The longitudinal analysis that is proposed gives rise to new hypotheses that provide a point of departure for future research. A first hypothesis concerns the acquisition of functional categories. It has been found that finiteness is acquired in a piecemeal fashion. Children's production data suggest a developmental path from 'absent' via 'lexical' to 'grammatical' finiteness. On the basis of this observation, it may be hypothesised that a generalisation can be made over functional categories. Such a hypothesis would state that the development of grammatical categories follows the path 'absent-lexical-grammatical'.

A second hypothesis concerns the order in which the semantic distinctions in the domains of tense, modality and aspect are formalised. It is hypothesised that this is dependent on the type of encoding in the target language. Verb placement, as well as richness of verbal morphology, seem to play a role. The earliest distinction is aspectual, namely that between completed and incomplete events: RPs versus RIs. In both cases, the verb that carries the aspectual specification is placed in sentence-final position. Modal and temporal distinctions require a broadened syntactic analysis as for these distinctions the child has to focus on sentence-initial and sentence-internal positions (first or second position). The modal distinction between realised (or factual) events and irrealis (or non-factual) events follows the early aspectual distinction: this distinction is formalised by verb first/second versus sentence-final placement of the main verb. The temporal distinction between past and present events is formalised by morphological differences on the verb in first/second position. This order suggests that Dutch children initially concentrate on the analysis of sentence-final words (which may be in line with Slobin's (1973) *Pay-attention-to-the-end-of-word-Principle*), followed by a general syntactic distinction between sentence-final and, roughly, sentence-initial verbs (i.e. placed in first/second position), which in turn is followed by the morphological fine-tuning of verbs in first/second position. The question whether or not this sequence of development can be generalised, remains open. It could be answered with the help of comparable cross-linguistic studies that compare Dutch results with results from target languages that differ from Dutch in their temporal, modal and aspectual encoding.

A third issue for further research is the hypothesis that the effect of Elsewhere Condition can be noticed in children's production data collected over a longer period of time. The Elsewhere Condition makes specific predictions for changes over time: when children have access to specified tensed or modal alternative for RIs, RIs will not be used anymore. More detailed data enlarge the chance of lexical overlap between the verb types in RIs, on the one hand, and SFs and PVs, on the other. This could show whether or not the replacement of RIs is consistent with the predicted effect of the Elsewhere Condition. Note that in order to be able to test the effect of the Elsewhere Condition, it is important to keep the temporal and modal meanings of the compared utterances constant. Therefore, modal PVs that express dynamic and deontic possibility must be left out of the set of modal utterances, as children do not (or hardly) express these types of modality in their RIs.

To sum up, this study provides a basis for cross-linguistic comparisons of longitudinal data. Ideally, the longitudinal data are denser than the data investigated in this thesis. I will point to one specific example from this thesis to exemplify the need for cross-linguistically comparable longitudinal data. In Chapter 5, I argued that the MRE in Dutch RIs is an effect of developmental patterns. Put simply, RIs are left for modal meanings, because the modal equivalents for RIs become productive later on than ongoing equivalents. Suppose that both the methodological artifact and the HSE do not sufficiently account for the semantic differences between Dutch and English RIs. In this case, developmental patterning may be a third factor that contributes to cross-linguistic differences in meaning. Given that the MRE in early child Dutch follows from the modal shift, the expectation is that there is no modal shift in early child English. One counteracting force of the modal shift is already provided by the HSE. Assume that English finite forms appear later than RIs (as in Dutch). Then, given that overuse of bare forms (i.e. inflection drop/incorrect inflection) is the effect of the emergence of finite forms, English RIs become a heterogeneous set over time. Another option would be that English children learn specified modal alternatives for RIs earlier than Dutch children. From work by Bloom, Tackeff & Lahey (1984) and Gerhardt (1991), we know that various modal catenatives like *gonna*, *wanna, hafta* and *needta* come in early in child English. It is unclear, however, if this early occurrence can be related to the absence of an MRE in English RIs.

A longitudinal comparison between early child Dutch and early child English has to go further. The experimental data presented in this thesis show that a comparison between Dutch and English is more complicated: besides RIs, English children use RPs, i.e. root participles containing a present participle (e.g. *Josie walking*). Data discussed by Hyams (2001) suggest that RIs and the present RPs overlap in meaning as both are used for ongoing events, the difference being that English RIs are less restricted and that these are also used for past and future denotations. However, when children try to map different forms onto different meanings from early on, it is unclear why children allow for overlap between RIs and RPs. Thus, to get a grip on the meaning of RIs in early child English, English RIs should not only be compared to SFs but also to RPs. A comparison of frequencies of these forms over time to meaning properties of RIs over time may deepen our insight into properties of RIs in early child English, and, finally, also into the strategies children apply while learning language. In sum, in order to make cross-linguistic comparisons to investigate universal patterns in children's use of early verb forms, generalisations based on data that come from different developmental stages do not suffice. The status of a form may change over time as a function of development, simply because a verb form stands in relation to the other verb forms that a child knows. As the whole range of verb forms that children know grows over time, it is expected that this one form will change as an effect of this growth. As the target set of verb forms differs from language to language (even between two Germanic languages like Dutch and English the differences are already considerable), it is expected that children acquiring different languages show their own developmental path; cross-linguistic comparisons can only be made with care.

I want to close this chapter with two final suggestions for extensions of the longitudinal research carried out in this thesis. Firstly, apart from a comparison between languages, it may be fruitful to compare the sequence of development of normally developing children, as has been described in this thesis, to the sequence of development of children with a specific language impairment (SLI). In the last couple of years, the discussion about RIs has been extended to SLI-children. Rice & Wexler (1996) proposed, for instance, that SLI-children have an *extended* optional infinitive stage. Turning to SLI-effects in child Dutch, we know from De Jong (1999) that the errors of Dutch SLI surface in inflection, in the sense that inflection is omitted or incorrectly used and SLI children tend to use RIs more often/longer than normally developing

children. A detailed comparison over time of the acquisition of inflection, on the basis of parameters such a lexical/paradigmatic variation, errors and subject use, may help us to determine where the SLI children deviate from the normally developing children with regard to their use of verbal inflection.

Secondly, the sequence of development of normally developing Dutch children that learn Dutch as their first language can also be compared to the development of children and adults that acquire Dutch as a second language (simultaneous or successive). For instance, Cornips (2000) has observed bilingual children of Moroccan and Turkish descent (between 6 and 13 years old) overuse the periphrastic construction *gaan* + *inf* ('go' + inf) and use it to denote ongoing events. This overuse is reminiscent of the overuse that I reported in the data of six young children that acquire Dutch as their first language. The question arises whether or not these two examples of the apparently same type of overuse also have the same cause. Furthermore, the Dutch-Moroccan/Turkish bilinguals that have the same age as the Dutch monolinguals examined in this thesis seem to make different errors with regard to agreement inflection than Dutch monolingual children (Cornips, p.c.). Closer examination of these errors, coupled with comparisons with the errors Dutch monolinguals and with the properties of the Moroccan and Turkish verbal inflectional systems would help us to understand more about bilingualism and the processes involved in second language learning.

APPENDIX 2.1:   SELECTED FILES

	I	II	III	IV
Abel	n.a.	1;10.30	2;01.02	2;05.17
		1;11.12	2;01.16	2;07.15
		1;11.26	2;02.19	2;07.29
			2;03.02	
Daan	1;08.21	2;00.22	2;04.14	2;08.13
	1;09.09	2;00.19	2;04.28	2;08.27
	1;10.16	2;01.21	2;05.11	2;09.10
Josse	n.a.	2;00.07	2;03.28	2;07.20
		2;00.21	2;04.11	2;08.04
				2;08.18
Laura	1;09.04 (01)	2;00.05 (08)	2;04.01 (17)	3;02.09 (34)
	1;09.18 (02)	2;00.19 (09)	2;04.15 (18)	3;03.00 (35)
		2;01.02 (10)	2;05.00 (19)	3;03.02 (36)
			2;05.17 (20)	3;04.06 (37)
			2;06.10 (21)	
Matthijs	1:09.30	1;11.10	2;04.24	2;10.22
	1;10.13	1;11.24	2;05.01	2;11.03
		2;00.09		2;11.19
Peter	1;07.18	1;09.20	2;00.28	2;03.07
		1;10.03	2;01.26	2;03.21

APPENDIX 2.2: FREQUENCIES OF RIS AND FINS, AND MLU IN STAGES I-IV

		RI	**FIN**	
	Stage	N	N	MLU (N_{MO}/N_{UT})[1]
Abel	I	n.a	n.a.	n.a.
	II	33	10	1.31 (1041/795)
	III	84	198	1.87 (3027/1622)
	IV	42	275	2.19 (2779/1268)

		RI	**FIN**	
	Stage	N	N	MLU (N_{MO}/N_{UT})
Daan	I	5	2	1.10 (423/382)
	II	54	49	1.41 (1880/1331)
	III	66	246	2.07 (2788/1348)
	IV	71	621	2.73 (4907/1796)

		RI	**FIN**	
	Stage	N	N	MLU (N_{MO}/N_{UT})
Josse	I	n.a	n.a.	n.a.
	II	99	11	1.32 (1257/955)
	III	76	98	1.98 (1831/926)
	IV	64	439	2.32 (3386/1459)

---

[1] The MLU is the ratio of the total number of morphemes from a speaker in a file and the total number of utterances in this file. The ratios give the MLU of all files within a stage. 'N_{MO}' is the number of morphemes in this stages and 'N_{UT}' is the number of utterances.

		RI	**FIN**	
	Stage	N	N	MLU (N_{MO}/N_{UT})
Laura	I	18	2	1.33 (420/315)
	II	56	26	1.42 (992/697)
	III	276	348	1.87 (4010/2139)
	IV	65	372	2.84 (3101/1090)

		RI	**FIN**	
	Stage	N	N	MLU (N_{MO}/N_{UT})
Matthijs	I	40	8	1.07 (1021/955)
	II	127	6	1.45 (1594/1098)
	III	112	85	1.83 (1841/1006)
	IV	51	524	2.55 (1998/1565)

		RI	**FIN**	
	Stage	N	N	MLU (N_{MO}/N_{UT})
Peter	I	26	0	1.00 (46/46)
	II	114	5	1.41 (1099/782)
	III	57	95	2.21 (965/436)
	IV	29	627	3.01 (3271/1085)

APPENDIX 3.1:   CODING SYSTEM

I   CODES: DENOTATION

In (1), (2) and (3) some examples of interpretations and codes are given. Each $ introduces a new field with different types of codes. In field 1, the modal value is given (modal = M, non-modal = N). If the utterance was modal, and the kind of modality could be determined, the M is followed by DY for dynamic modality, and by DE for deontic modality. If the modal value could not be determined, this field contained an O. Field 2 gives a specification of field 1. If the utterance was modal, the kind of modality was specified: 'nes' for necessity and 'pos' for possibility. After the colon, additional information was given as to whether the utterance was desiderative and expressed a wish ('des'), was regulating and expressed a command ('reg') or expressed a capacity ('cap'). If the utterance was non-modal, the completedness ('c'), ongoingness ('o') or prospectiveness ('p') of the event expressed in the utterance, was determined. If this value was unclear, the utterances was assigned an 'u' in this field. Field 3 gives information about the verb form (this was a while range of codes for RIs, simple finite verb, periphrastic verbs, modal verbs, copula, etc.).

(1)   *PET: Peter woef hebben
                Peter dog have
       %mod: $MDY $nes:des $RI
       %par: on the verge of crying
       *MOT: o
       *MOT: wil je even je +...
                want you *part* you
       *MOT: wil je even je woef hebben ?
                want you *part* your dog have
                'Do you want to have your dog?'
       %act: <aft> kisses PET .   Peter 2;00.28

(2)　*LAU:　die s pakke [= plakken, EB], hoor
　　　　　　　that glue *interj*
　　　%mod:　**$N $o $RI**
　　　*LAU:　die
　　　　　　　that
　　　*LAU:　pakke [= plakken]
　　　　　　　glue
　　　%mod:　**$N $o $RI**
　　　*MOT:　is dat plakken ["] ?
　　　　　　　is that glueing?
　　　*MOT:　ben je daarmee aan het plakken?
　　　　　　　are you therewith on the glue
　　　　　　　'Are you glueing with that?'　　　　Laura 2;05.17

(3)　*MAT:　mama !
　　　*MOT:　wat is er ?
　　　　　　　what is there
　　　　　　　'What's wrong?'
　　　*MAT:　Ieke spugen !
　　　　　　　Ieke throw up
　　　%mod:　**$N $c $RI**
　　　*MOT:　ging Ieke spugen ?
　　　　　　　went Ieke throw up
　　　　　　　'Did Ieke throw up?'
　　　*MAT:　ja .
　　　　　　　yes
　　　*MOT:　oh, getsie
　　　　　　　oh yuk　　　　　　　　　　　　Matthijs 1;11.24

II　CODES: SUBJECTS

Each RI is provided with a code for person of the subject (F(irst), S(econd) or T(hird)), number of the subject (S(ingular) or P(lural)) and overtness of the subject (O(vert) or N(ull)). In (4) and (5) some examples are given with a first person singular overt subject and a null subject respectively:

(4)   *ABE: ik een hand tekenen
               I a hand draw
               'I am drawing a hand'
               %mod: **$FSO** $N $o $RI   Abel 2;07.15

(5)   *LAU: oto niet neezett
               car not downput
               'don't put the car down'
               %mod: **$FSN** $N $o $RI   Laura 2;06.10

APPENDIX 3.2:   EXPERIMENTAL STORIES

Action: WASHING

Dirty dog (modal)

This is the story of the dirty dog who wants to go into the house to eat. He is so hungry and his food is in the house. But the dog is much too dirty to go into the house. Look how dirty it is! The doggy has to wash himself. Look! Next to the house is a bath tub, where he can go and wash himself.

Washing pig (non-modal)

This is the story about the dirty pig. The pig is waiting to get into the bath. See! Now he jumps into the bath tub. He is washing and washing and washing. He washes himself until he is completely clean. You see?

Action: RUNNING

Boy in danger (modal)

This is the story of a little boy who has to run away. Look at that little boy. Look how he is playing on the street. He throws the ball high up in the air. And then he catches it again. You see? But the boy is very silly. It is dangerous on the street. Oh oh! See what a happens! There comes a car! And now? The boy has to run away now.

Running girl (non-modal)

Here is a little girl. Her name is Lisa. Lisa is running. She is running home because it rains. Look how Lisa runs. She is running as fast as she can.

Action: DRINKING

Thirsty girl (modal)

This is the story about a little girl who is very thirsty. She is thirsty and she really wants to drink, but the glass in her hand is empty. Do you see the glass? Is it empty? Now the girl goes to her mother and holds up her empty glass.

Drinking boy (ongoing)

Here is Peter. Peter is sitting at the table. He has a big glass of lemonade. See how he picks up the glass and lifts it. Peter is drinking the lemonade. He drinks, drinks, drinks until his glass is empty

Action: CALLING

Girl wanting to call (modal)

This is the story of a little girl who wants to call her grandmother. Her granny has her birthday today. Look, there is the girl. She sits on the floor. There is the telephone on the cupboard. But the phone is too high and the girl is too small to reach the phone.

Calling boy calling (ongoing)

This boy is calling his friend. He talks and talks and talks for a long time.

APPENDIX 3.3:   OVERUSED BARE STEMS

Abel	ga ook niet (= dat gaat ook niet) (2x)	2;05.27
	jij bouw trein	2;05.27
	jij heb oorbel?	2;05.27
	ja, jij toren maak ?	2;07.15 (OV/RI)
	zit koekjes in	2;07.15
	hij heb even kots	2;07.15
	die heb rits	2;07.15
	je heb rits	2;07.15
	hij doe niet meer	2;07.29
	olifant ga niet door de	2;07.29
Daan	zoek Aart [?]	2;01.21
	pas ient (=past niet)	2;01.21
	Daan li (=ligt) in de wieg	2;04.14
	nee jij heb xxx	2;04.14
	(l)ig niet (in) de kist	2;04.28
	die heef snor	2;04.28
	deze ga in rijden	2;05.11
	deze ga auto is ie	2;05.11
	ha politie deze heb	2;05.11
	die heef band	2;09.27
	pak Rosa kuis [?]	2;08.27
	jij heef die heb ik ook al voor koffie	2;08.27
	doe zeer, daar mij hier kleine voet	2;08.27
	die heef geel	2;08.27
	val ze op hun neus, pof	2;09.10
	hij vies (=fietst?) om	2;09.10
Josse	trui heef Abel	2;00.07
	ape geef [= aap een kusje geven] 2;00.21 (OV/RI)	
	kom niet	2;00.21
	ga niet he?	2;03.28
	ga niet	2;03.28
	even zit Jos	2;04.11

	ga er af	2;04.11
	jij heb knoop	2;04.11
	toren bouw	2;04.11 (OV/RI)
	duur nog een tijdje (2x)	2;07.20
	Audrey slaap nog een tijdje	2;07.20
	kom &6 slager	2;07.20
	Ab heef ook buik	2;07.20
	Hanneke heef niet	2;07.20
	heef een snor	2;07.20
	Gerard heef geld	2;08.04
	au, doe mij zeer	2;08.04
	je kom eruit	2;08.18
	hier kom de baktor (=tractor)	2;08.18
	dat grote zon heef een gezicht	2;08.18
	jij teken een reiger	2;08.18
Laura	oh, bal pak	1;09.04 (OV/RI)
	boeke kijk	1;09.04 (OV/RI)
	ikke meer maak	2;04.01 (PV/RI)
	potte maak (2x)	2;04.01 (PV/RP)
	poppe saap	2;04.01
	paard rij	2;04.01 (OV/RI)
	mama e sie niet	2;04.01
	doe nou? (2x)	2;05
	die ook pak niet	2;05.17
	ekom papa aan!	3;02.09
	kom nou aan	3;02.09
	kom bij	3;03.00
	deese heb	3;03.00
	dese hoor niet daar	3;03.00
	dese hoor nou?	3;03.00
	nee jij vasthou mij	3;03.02
	kijk, Ernie huil	3;03.02
	pas so die (die past zo)	3;03.02
	dese blijf hier in deur	3;04.06
	dat doe	3;04.06 (OV/RI)
	sij heb net de auto	3;04.06
	ga sij met de auto	3;04.06

	ga naa de ziekenhuis	3;04.06
	sij maak n nat	3;04.06
	heb meneer nou op?	3;04.06
	dese heb wel	3;04.06
	ga Roos na ande sembad	3;04.06
Matthijs	pak boek	1;11.10
	mama help	1;11.10
	zit ook pitten in	2;10.22
	heef jij pleister?	2;10.22
	heef jij muggebulten	2;10.22
	Mirjam klim berg op	2;10.22
	die zat pitjes al uit	2;10.22
	Hannah heb ook n ijsbeer	2;11.03
Peter	binnen stap	1;09.20 (PV/RI)
	Peter stap	1.10.03
	Peter hier stap	1;10.03 (PV/RI)
	Peter maak open	2;00.28
	Peter los maak	2;00.28 (PV/RI)
	Peter stoel kijk	2;00.28 (OV/RI)
	Jiska maak los	2.01.27
	die ga even doorheen	2;03.07
	hij zeg toettoet	2;03.07
	hij zeg niet toettoet	2;03.07
	hij zeg weer piepo	2;03.07
	daar lig kikker	2;03.21
	nu heef kikker schaatsen	2;03.21
	die kikker heef ook om	2;03.21
	daar heef kikker sjaal om	2;03.21
	hij heef voetjes	2;03.21
	die kom uit bed	2;03.21
	heef eend schaatsen	2;03.21
	zeg niet klik	2;03.21
	hij ga omhoog	2;03.21
	Peter maak weer open	2;03.21

APPENDIX 4.1:   VERB TYPES IN RIS

Below all verbal predicates that appear in RIs are listed per child. Stative predicates that are non-type-shifting are bold-faced and in italics, while type-shifting predicates are only bold-faced.

Abel   **blijven 'stay'**, bouwen 'build', drinken 'drink', doen 'do', eten 'eat', fietsen 'cycle', gaan 'go', geven 'give', gooien 'throw', ***hebben 'have'***, helpen 'help', hikken 'hiccup', kijken 'look', kammen 'comb', kleien 'do clay modelling', klimmen 'climb', komen 'come', kopen 'buy', lachen 'laugh', laten zien 'show', lezen 'read', lopen 'walk', maken 'repair', (mee)nemen 'take', opeten 'finish food', pakken 'get', ***passen 'fit'***, poepen 'defecate', proeven 'taste', opschrijven 'write down', (s)choonmaken 'clean', (s)chrijven 'write', slapen 'sleep', **staan 'stand'**, stoppen 'stop', tekenen 'draw', tillen 'lift', trekken 'pull', uitkijken 'watch', vallen 'fall', vragen 'ask', wachten 'wait', zetten 'set', **zitten 'sit'**, zoeken 'search'.

Daan   aaien 'stroke', aankomen 'touch', bewaren 'keep', **blijven 'stay'**, boksen 'box', bouwen 'build', doen 'do', dichtgooien 'close', douchen 'take a shower', drinken 'drink', drijven 'float', dukke [= drukken, EB] 'press', eten 'eat', gooien 'throw', halen 'get', ***hebben 'have'***, **horen 'hear'**, indoen 'put in', kauwen 'chew', kijken 'look', kleien 'do clay modelling', k(l)imme 'climb', komen 'come', kneden 'knead', leggen 'put', lezen 'read', **liggen 'lie'**, maken 'repair', neerzetten 'put down', opeten 'finish food', opschiete 'hurry', owake [= openmaken?, EB], pakken 'get', (ka)pokmaken 'destroy', ophalen 'pick up', rijden 'ride', (s)chaatsen 'skate', schommelen 'swing', slapen 'sleep', slopen 'destroy', spelen 'play', **staan 'stand'**, stoppen 'stop', tekenen 'draw', timme [= timmeren, EB] 'hammer', uitd(r)ukke 'press out, vastmaken 'fasten', voetballen 'play soccer', wachten 'wait', zagen 'saw', **(in)zitten 'sit'**, zoeken 'search', zwemmen 'swim'.

Josse   afruilen 'exchange', bellen 'call', blazen 'blow', boeren 'belch', bouwen 'build', lazen [= blazen] 'blow', doen 'do', draaien 'turn', drinken 'drink', eten 'eat', gaan 'go', geven 'give', gooien 'throw',

*hebben **'have'***, helpen 'help', hoesten 'cough', huilen 'cry', kijken 'look', klimmen 'climb', kopen 'buy', koppetje duikelen 'turn head over heels', leegmaken 'empty', lezen 'read',   maken 'make', meenemen 'take', openmaken 'open', opzetten 'put on', pakken 'get', plassen 'urinate', prikken 'prick', roeren 'stir', schrijven 'write', slaan 'beat', slapen 'sleep', spelen 'play', springen 'jump', **staan 'stand'**, stampen 'stamp ones foot', sturen 'steer', tanken 'tank', tekenen 'draw', timmeren 'hammer', uithalen 'get out', uitstappen 'get out', vallen 'fall', vangen 'catch, vegen 'weep', verven 'paint', weggooien 'throw away', wachten 'wait', **zien 'see'**, zingen 'sing', **zitten 'sit'.**

Laura   aandoen 'put on', aankomen 'touch', balle(n) 'play with ball', bellen 'call', bouwen 'build', dansen 'dance', dichtdoen 'close', doen 'do', doese 'take a shower', d(r)inken 'drink', (op)eten 'eat', fuite [= fluiten?, EB] 'whistle', slije [= glijden] 'slide', sooie [= gooien] 'throw', *hebben **'have'***, huilen 'cry', kammen 'comb', kijken 'look', k(n)oeien 'spill', koken 'cook', komen 'come', kopen 'buy', lezen 'read', maken 'repair', nspatte [= natspatten, EB] 'wet', neerzetten 'put down', nemen 'take', omruilen 'exchange', ope(n)maken 'open', opeten 'finish food', opruimen 'clean', paatrije [= paard rijden] 'ride a horse', pakken 'get', plakken 'glue', poetsen 'polish', rij(d)en 'ride', rooje [= rollen] 'roll', s(ch)oonmaken 'clean', sille (?), s(l)apen 'sleep', s(l)apen doen 'sleep do', (s)pelen 'play', stappen 'step', stoppen 'stop', suise [= schuiven] 'shove', tekenen 'draw', uit(s)tappen 'get off', varen 'sail', toppe [= verstoppen] 'hide', voorlezen 'read', wassen 'wash', semme [= zwemmen] 'swim', zingen 'sing', **zitten 'sit'.**

Matthijs aandoen 'put on', aangeven 'hand', aanmaken 'put on', aanzetten 'put on', bakken 'bake', blaffen 'bark', binnenlopen 'walk in', bouwen 'build', brengen 'bring', doen 'do', dragen 'carry', draaien 'turn', drinken 'drink', duikelen 'turn somersaults', eten 'eat', fietsen 'cycle', gaan 'go', geven 'give', gooien 'throw', halen 'get', *hebben **'have'***, helpen 'help', hijsen 'lift', kijken 'look', kletsen 'chat', klimmen 'climb', komen 'come', kopen 'buy', knippen 'cut', kruipen 'crawl', lachen 'laugh', laten 'let', lezen 'read', liggen 'lie', maken 'repair', meekomen 'come with', naaien 'sew', neerzetten 'put

down', omdoen 'put on', opdoen 'put on', open maken 'open', ophalen 'pick up', opruimen 'clean', pakken 'get', poepen 'relieve nature', praten 'talk', rijden 'ride', roeren 'stir', schuiven 'shove', slapen 'sleep', (s)pelen 'play', spugen 'throw up', **staan 'stand'**, stappen 'step', strijken 'iron', sturen 'steer', taken [= takelen, EB] 'rig', tikken 'tick', trekken 'pull', uitdoen 'put off', vallen 'fall', vasthouden 'hold', vinden 'find', voeden 'feed', voorlezen 'read', wegbrengen 'take away', weghalen 'get away', **zien 'see'**, zingen 'sing', **zitten 'sit'**, zoeken 'search', zuigen 'suck', zwemmen 'swim'.

Peter   branden 'burn', dichtdoen 'close', doen 'do', draaien 'turn', dragen 'carry', drinken 'drink', duwen 'push', eten 'eat', geven 'give', gooien 'throw', ***hebben 'have'***, koken 'cook', kijken 'watch', **liggen 'lie'**, losmaken 'untie', nemen 'take', opbeuren 'lift', pakken 'get', rijden 'ride', roeren 'stir', schaatsen 'skate', slapen 'sleep', **staan stand'**, stappen 'step', stukmaken 'destroy', tikken 'tick', trekken 'pull', vallen 'fall', **voelen 'feel'**, vangen 'catch', zetten 'put', **zien 'see'**, **zitten 'sit'.**

APPENDIX 4.2:   VERB TYPES IN SFS

Below all verbal predicates that appear in SFs are listed per child. Stative predicates that are non-type-shifting are bold-faced and in italics, while type-shifting predicates are only bold-faced.

Abel   ***(be)doelen 'mean'***, doen 'do', gaan 'go', **hebben 'have'**, **horen 'hear'**, huilen 'cry', komen 'come', ***kunnen 'can'***, laten 'let', ***lukken 'work out'***, maken 'repair', ***moeten 'must'***, ***mogen 'be allowed'***, nemen 'take', **passen 'fit'**, **staan 'stand'**, vallen 'fall', vinden 'find', ***weten 'know'***, ***willen 'want'***, zetten 'put', **zien 'see'**, ***zijn ' be'***, **zitten 'sit'**, ***zullen 'will'***.

Daan   doen 'do', draaien 'turn', ***durven 'dare'***, gaan 'go', *hebben 'have'*, ***hoeven 'must'***, **horen 'hear'**, koesen (?), komen 'come', krijgen 'get', ***kunnen 'can'***, laten 'let', leggen 'put', **liggen 'lie'**, ***lukken 'work out'***, maken 'repair', ***moeten 'must'***, ***mogen 'be allowed'***, pakken 'get', **passen 'fit'**, regenen 'rain', slapen 'sleep', smeren 'spread', spelen 'play', stoppen 'stop', vallen 'fall', ***lekker vinden 'like'***, ***weten 'know'***, ***wil 'want'***, **zie 'see'**, zingt 'sings', **zit 'sit'**, ***zijn 'be'***, zoeken 'search', ***zullen 'will'***.

Josse   doen 'do', gaan 'go', **hangen 'hang'**, ***heten 'be called'***, klimmen 'climb', komen 'come', krijgen 'get', ***kunnen 'can'***, ***moeten 'must'***, ***mogen 'be allowed'***, slapen 'sleep', springen 'jump', **staan 'stand'**, vallen 'fall', vinden 'find', ***weten 'know'***, ***willen 'want'***, **zeggen 'say'**, ***zijn 'be'***, **zitten 'sit'**, ***zullen 'will'***.

Laura   blaffen 'bark', doen 'do', ***durven 'dare'***, gaan 'go', sooien [= gooien] 'throw', ***hebben 'have', hoeven 'must'***, **horen 'hear'**, knoeien 'spill', komen 'come', ***kunnen 'can'***, ***mogen 'be allowed'***, ***moeten 'must'***, **passen 'fit'**, vallen 'fall', vinden 'find', ***weten 'know'***, ***willen 'want'***, **zien 'see'**, ***zijn 'be'***, zingen 'sing', **zitten 'sit'**, zwemmen 'swim'.

Matthijs bijten 'bite', botsen 'bump', **denken 'think'**, doen 'do', gaan 'go', geven 'give', gooien 'throw', ***hebben 'have'***, ***hoeven 'must'***,

huilen 'cry', klimmen 'climb', komen 'come', ***kunnen 'can'***, **liggen 'lie'**, ***lukken 'work out'***, ***lusten 'like'***, maaien 'mow', maken 'repair', ***moeten 'must'***, ***mogen 'be allowed'***, **passen 'fit'**, regenen 'rain', rijden 'ride', slapen 'sleep', uittrekken 'take off', vinden 'find', wachten 'wait', ***weten 'know'***, ***willen 'want'***, worden 'become', zetten 'put', **zien 'see'**, ***zijn 'be'***, **zitten 'sit'**, zoeken 'search'.

Peter   branden 'burn', dansen 'dance', **denken 'think'**, doen 'do', draaien 'turn', dragen 'carry', gaan 'go', geven 'give', gooien 'throw', halen 'get', ***hebben 'have'***, **horen 'hear'**, eten 'eat', kijken 'look', komen 'come', ***kunnen 'can'***, **liggen 'lie'**, lopen 'walk', maken 'repair', ***moeten 'must'***, ***mogen 'be allowed'***, pakken 'get', **passen 'fit'**, rammelen 'rattle', rijden 'ride', **staan 'stand'**, vallen 'fall', ***willen 'want'***, ***zeggen 'say'***, zetten 'put', **zien 'see'**, ***zijn 'be'***, **zitten 'sit'.**

APPENDIX 5.1:   LEXICAL GROWTH OF INFINITIVES AND FINITE VERBS, GROWTH OF LEXICAL OVERLAP

Table A5.1: Abel

	Infinitives		Finite verbs[2]		Lexical overlap	
Stage	Types (new)	Acc.	Types (new)	Acc.	Types	Acc.
I/II	bouwen, doen, drinken, eten, gooien, hebben, kijken, lezen, (s)chrijven, teke(nen), trekken, vallen, zitten	13	kan, wil	2	-	0
III	blijven, fietsen, helpen, kammen, kleien, komen, kopen, laten, lopen, maken, opschrijven, pakken, passen, poepen, proeven, slapen, staan, (s)choonmaken, tillen, uitkijken, zoeken	34	doe, gaat, heb, is, kom, luk, vind, mag, moet, pas, zie, zit, zul	15	doe/doen heb/hebben zit/zitten	3

---

[2] I take here the stem of the finite forms. It various cases, the children use finite forms with endings on [-t]or [-en]. This list, however, focuses on the accumulation of lexical items that appear in finite and infinitival form and on the accumulation of lexical overlap. The paradigmatic variation of the different finite forms can be found in Appendix 5.2.

| IV | gaan, geven, hikken, klimmen, lachen, nemen, opeten, optillen, stoppen, vragen, wachten, zetten | 46 | (be)doel, hoor, huil, neem, maak, sta, val, weet, zet | 24 | gaat/gaan kom/komen maak/maken neem/nemen sta/staan val/vallen zet/zetten | 10 |

Table A5.2.: Daan

	Infinitives		Finite verbs		Lexical overlap	
Stage	Types (new)	Acc.	Types (new)	Acc.	Types	Acc.
I/II	aaie(n), dichtgooien, doen, drinken, d(r)ukke, eten, gooie(n), horen, indoen, kauwe(n), kijken, k(l)imme(n), lezen, liggen, opete(n), o(p)schiete(n), owake (openmaken), pakken, schommele(n), staan, timme(ren), zagen, zitten, zoeke(n)	24	gaat, is, kan, lig, moet, wil, zie, zit, zing	9	lig/liggen zit/zitten	2
III	bouwen, hebben, inzitten, (ka)pokmaken, kleien, kneden, maken, rijden, (s)chaatsen, stoppen, tekenen, uitdrukken	36	doe, durf, heb, hoef, luk, mag, pak, speel, stop, val, weet, zul	21	doe/doen heb/hebben pak/pakken stop/stoppen	6

| IV | aankomen, bewaren, blijven, boksen, douchen, drijven, komen, leggen, ophalen, neerzetten, slapen, slopen, spelen, vastmaken, voetballen, wachten, zwemmen | 53 | draai, hoor, koes (?), kom, krijg, laat, leg, maak, pas, regen, slaap, smeer, vind, zoek | 35 | hoor/horen kom/komen leg/leggen maak/maken slaap/slapen speel/spleen zoek/zoeken | 13 |

Table A5.3.: Josse

	Infinitives		Finite verbs		Lexical overlap	
Stage	Types(new)	Acc.	Types (new)	Acc.	Types	Acc.
I/II	bellen, (b)lazen, doen, drinken, eten, gooien, hebben, hoesten, kijken, kopen, maken, pakken, schrijven, spelen, staan, stampen, vallen, vegen, zingen, zitten	20	heet, kan, kom, mag, rink (spring)	5	-	0

| III | afruilen, boeren, bouwen, geven, helpen, huilen, lezen, meenemen, openmaken, opzetten, prikken, roeren, slapen, sturen, tekenen, timmeren, uithalen, vangen | 38 | doe, ga, krijg, is, moet, wil | 11 | doe/doen | 1 |
| IV | afgooien, draaien, gaan, klimmen, koppetje duik(el)en, leegmaken, plassen, slaan, springen, tanken, uitstappen, verven, wachten, weggooien, zien | 53 | hang, klim, slaap, sta, val, vind, weet, zeg, zit, zul | 21 | ga/gaan klim/klimmen slaap/slapen spring/springen sta/staan val/vallen zit/zitten | 8 |

Table A5.4.: Laura

	Infinitives		Finite verbs		Lexical overlap	
Stage	Types (new)	Acc.	Types (new)	Acc.	Types	Acc.
I/II	doen, ete(n), kammen, kijke(n), k(n)oe(i)en, koken, lezen, pakke(n), poetsen, rij(d)e(n), s(l)apen, (s)pele(n), tekenen, zitten	14	doe, hoor, hoe(f), kan, moe(t), val, wil, zit	8	zit/zitten doe/doen	2

| III | aankome(n), balle(n), bellen, bouwe(n), danse(n), doese (douchen), d(r)inke(n), sooie (gooien), hebben, huilen, komen, maken, neerzette(n), nemen, ope(n)maken, opeten, omruile(n), paa(r)drij(d)e(n), p(l)akke(n), rooje (rollen), saapen doen (slapen doen), s(ch)oonmake(n), suise(schuiven), stoppe(n), uit(s)tappe(n), varen, toppe (verstoppen), wassen, semme (zwemmen), singe (zingen) | 45 | ga, sooi (gooi), heb, kom, is, pas, mas (mag), zie, sing (zing) | 17 | kom/komen gooi/gooien heb/hebben zing/zingen | 6 |
| IV | aandoen, slije (glijden) dichtdoen, kope(n), indoen, nspatte (natspatten), opruime(n), sille (?), stappe (n), voorlese (voorlezen) | 55 | blaf, durf, k(n)oei, swem (zwem), vind, weet | 23 | knoei/knoeien zwem/zwemmen | 8 |

Table A5.5: Matthijs

	Infinitives		Finite verbs		Lexical overlap	
Stage	Types (new)	Acc.	Types (new)	Acc.	Types	Acc.
I/II	aanmaken, brengen, doen, draaien, dragen, drinken, eten, gooien, halen, helpen, kijken, knippen, komen, kruipen, lachen, lezen, liggen, naaien, pakken, poepen, rij(d)en, roeren, schuiven, slapen, (s)pelen, spugen, staan, strijken, sturen, tikken, vinden, zingen, zitten, zoeken, zuigen, zwemmen	36	mag, pas	2	-	0
III	aandoen, aanzetten, bakken, binnenlopen, bouwen, duikelen, fietsen, gaan, geven, hebben, hijsen, kletsen, klimmen, kopen, laten, maken, neerzetten, omdoen, opdoen, opruimen, praten, takelen, uitdoen, vallen, vasthouden, voeden, zien	63	doe, heb, ga, gooi, kan, klim, is, moet, regen, uittrek, zie, zit	14	doe/doen ga/gaan gooi/gooien heb/hebben klim/klimmen zie/zien zit/zitten	7

| IV | aangeven, blaffen, openmaken, ophalen, meekomen, stappen, trekken, voorlezen, wegbrengen, weghalen | 73 | bijt, bots, denk, geef, hoef, huil, kom, lig, luk, lust, maak, maai, rijd, slaap, vind, wacht, weet, wil, word, zet, zoek | 35 | geef/geven kom/komen lig/liggen maak/maken rijd/rijden slaap/slapen vind/vinden zet/zetten zoek/zoeken | 16 |

Table A5.6: Peter

	Infinitives		Finite verbs		Lexical overlap	
Stage	Types (new)	Acc.	Types (new)	Acc.	Types	Acc.
I/II	doen, draaien, drinken, duwen, geven, gooien, hebben, koken, liggen, nemen, pakken, rij(d)en, roeren, slapen, staan, stappen, trekken, vallen, vangen, zetten, zitten	21	eet, mag, zit	3	zit/zitten	1

| III | branden, dichtdoen, dragen, eten, kijken, losmaken, (op)beuren, stukmaken, voelen, zien | 31 | doe, draai, ga, gooi, haal, heb, hoor, is, kan, kijk, kom, maak, moet, pak, wil | 18 | doe/doen draai/draaien eet/eten gooi/gooien heb/hebben kijk/kijken maak/maken pak/pakken | 9 |
| IV | tikken, schaatsen | 33 | brand, dans, denk, draag, geef, lig, loop, pas, rammel, rijd, sta, val, zeg, zet, zie | 33 | brand/branden draag/dragen geef/geven lig/liggen rijd/rijden sta/staan val/vallen zet/zetten zie/zien | 18 |

APPENDIX 5.2:   PARADIGMATIC VARIATION

	Stage III	Stage IV
Abel	is/was 'is/was'	is/zijn/was 'is/are-*pl*/was'
	doe/doet 'do-*sg*/does'	heb/heeft/had
	vin/vindt 'find-*sg*/finds'	'have-*sg*/has/had-*sg*'
	past/passen 'fit-*sg*/fit-*pl*'	ga/gaat/gaan 'go-*sg*/goes/go-*pl*'
		doe/doet 'do-*sg*/does'
Daan	is/zijn 'is/are-*pl*'	heb/heeft 'have-*sg*/has'
	heb/heeft 'have-*sg*/has'	doe/doet/deed
	doe/doet 'do-*sg*/does'	'do-*sg*/does/ did *sg*'
	ga/gaat/gaan	hoort/horen 'hears/hear-*pl*'
	'go-*sg*/goes/go-*pl*'	val/vallen 'fall-*sg*/fall-*pl*'
Josse	ga/gaat/gaan	is/bent 'is/are-*sg*'
	'go-*sg*/goes/go-*pl*'	heb/heeft 'have-*sg*/has'
	doe/doet 'do-*sg*/does'	ga/gaat 'go-*sg*/goes'
		doe/doet 'do-*sg*/does'
		staat/stond 'stands/stood-*sg*'
Laura	is/zijn 'is/are'	is/ben 'is/am'
	doet(e)/doen 'does/do-*pl*'	heb/heeft 'have-*sg*/has'
	va(l)/valt 'fall-*sg*/falls'	ga/gaat 'go-*sg*/goes'
		doet(e) 'does'
Matthijs	is/was 'is/was'	is/zijn/was 'is/are-*pl*/was'
		heb/heeft/hebben
		'have-*sg*/has/ have-*pl*'
		doe/doet 'do-*sg*/does'
		gaat/gaan 'goes/go-*pl*'
		zit/zitten/zat 'sit-*sg*/sit-*pl*/sat-*sg*'
		vin/vindt 'find-*sg*/finds'

APPENDIX 5.3:   AGREEMENT ERRORS

Overgeneralisations second and third person singular

(1)   ik ook ziet ook niet   Abel 2;03.02
         I-*1sg* also sees-*2/3sg* also not
         'I do not see it either'

(2)   jij heeft kloopjes [= knoopjes, EB]   Abel 2;03.02
         you have-*2/3sg* buttons-*dim*
         'You have buttons'

(3)   ik doet weer   Josse 2;03.28
         I-*1sg* does-*2/3sg* again
         'I do it again'

(4)   doet het fordijn [=gordijn, EB]) even dicht   Josse 2;07.20
         (I)-*1sg* does-*2/3sg* the curtain just close
         'I am closing the curtain'

(5)   ik gaat me niet   Laura 3;03
         I-*1sg* goes-*2/3sg* me not
         'I am not going'

(6)   heef(t) ik   Matthijs 2;10.22
         has-*2/3sg* I-*1sg*
         'I have it'

(7)   ja, die heef(t) ik ook nodig   Matthijs 2;11.03
         yes, that has-*2/3sg* I-*1sg* also need
         'Yes, I need that as well'

(8)   ik heeft die vuilnisauto   Matthijs 2;11.19
         I-*1sg* has-*2/3sg* that garbage truck
         'I have that garbage truck'

(9)   is ik weer   Matthijs 2;10.22
         is-*2/3sg* I-*1sg* again
         'I am again'

(10)   koe ziet ik ook in boot   Matthijs 2;10.22
         cow sees-*2/3sg* I-*1sg* also in boat
         'I also see the cow in the boat'

(11)   ik heeft hem   Peter 2;03.21
         I-*1sg* has-*2/3sg* him
         'I have him'

Wrong Number

(12)   wij moet ook   Abel 2;07.29
         we-*pl* must-*sg* also
         'We have to as well'

(13)   xxx woont allemaal kindertjes   Abel 2;07.29
         xxx lives-*sg* all little children-*pl*
         'All little children live there'

(14)   eerst deze gaan deze auto   Daan 2;09.10
         first this go-*pl* this car-*sg*
         'First this car goes'

(15)   deze horen ook deze auto   Daan 2;09.10
         this belong-*p*l also this car-*sg*
         'This car also belongs ...'

(16)   eh er is hier mense drin   Laura 3;06.04
         eh there is-*sg* here people-*pl* here in
         'Eh there are people in here'

(17)   dat gaan deze eerst de vinger   Peter 2;03.07
         that go-*pl* this first the finger-*sg*
         'First this finger goes'

ABBREVIATIONS

Abbreviation	Full term
AUX-drop	*Auxiliary drop* The explanation for RIs according to which RIs contain a phonetically non-realised auxiliary.
ATOM	*Agreement and Tense Omission Model* Theory that states that Agreement and Tense are underspecified functional categories in child grammars.
BS	*Bare Stem* Finite or non-finite verb that appears as bare stem because inflection has been dropped (as an effect of processing limitations) or null inflection is overused. ex: Harry drink koffie ex. Harry drink coffee
EC	*Eventivity Constraint* RIs are restricted to event-denoting predicates.
FCH	*Full Competence Hypothesis* Functional categories are present from early on.
FIN	*Finite sentence*
FP	Finite sentence with participial main verb ex: Harry is drinking coffee
GoF	*Grammaticalisation of Finiteness (GoF) Hypothesis (final)* (i) In early child Dutch, the grammatical marking of finiteness by means of inflection and verb movement is preceded by a lexical-finiteness-stage. (ii) In the lexical-finiteness-stage, simple lexical finiteness markers precede complex lexical finiteness markers.

HSE	*Heterogeneous Set Effect*
	RIs comprise finite and non-finite sentences as an effect of inflection drop (restricted to languages with bare stem RIs).
MA	*Modal Approach*
	RIs obey the EC because they are modal.
MRE	*Modal Reference Effect*
	The effect that RIs are more frequently used for modal meanings than for non-modal meanings.
NCC	*Non-Completedness Constraint*
	The predicate of an RI cannot refer to a completed event.
NS	*Null subject*
	Phonetically non-realised subject
	ex: 0 heb gisteren te veel koffie gedronken
	0 have yesterday too much coffee drink-*part-past*
NTA	*No Tense Approach*
	RIs obey the EC because they lack tense.
OI-stage	*Optional Infinitive Stage*
	Stage in language acquisition in which finiteness is optional; characterised by the use of both RIs and SFs.
PIC	*Prepositional Infinitival Complement* (Dutch progressive)
	ex: Harry is liters koffie aan het drinken
	Harry is liters coffie on the drink-*inf*
PV	Finite sentence with *periphrastic verb*
	ex: Harry mag geen koffie meer drinken
	Harry may no coffee anymore drink-*inf*
PV_{lex}	Finite sentences with *periphrastic verbs* used in the *lexical-finiteness-stage* (stage III characteristic).

RI	*Root Infinitive*
	Nonfinite clause with infinitival main verb:
	ex: Harry koffie drinken
	Harry coffee drink-*inf*
RP	*Root participle*
	Nonfinite clause with participial main verb:
	ex: Harry koffie gedronken
	Harry coffee drink-*part-past*
	ex: Harry drinking
SF	*Finite sentence with simple verb*
	ex: Harry drinkt koffie
	Harry coffee drink-*fin*
SF_{gram}	Finite sentence with a simple verb that is grammatically marked for finiteness by inflection and movement (not before stage IV)
SF_{lex}	*Finite sentence with simple verb that is lexically marked for finiteness; contain auxiliary-like predicates, precursors of inflection and verb movement (stage II characteristic)*
SF_{θ}	*Finite sentence with simple thematic main verb*

TABLES AND FIGURES

TABLES

---

[1] All children behave very similarly, and differences between the subjects are marginal. Therefore, the data are collapsed. The same holds for the results in Table 4.6.

FIGURES

BIBLIOGRAPHY

Anderson, S. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Antinucci, F. and R. Miller (1976). "How children talk about what happened." *Journal of child language* 3: 167 - 189.

Armon-Lotem, S. (1995). "What Hebrew early verbs teach us about root infinitives." In: Wijnen, F. and C. Koster (Eds) *Proceedings of the Groningen Assembly on Language Acquisition (GALA 1995)*. Groningen: Center for Language and Cognition.

Astington, J. and A. Gopnik (1991). "Understanding desire and intention." In: A. Whiten (Ed.) *Natural theories of mind: The evolution, development and simulation of second-order representations*. Oxford: Basil Blackwell.

Avrutin, S. (1999). *Development of the syntax-discourse interface*. Dordrecht: Kluwer Academic Publishers.

Barbiers, S. (1995). *The syntax of interpretation*. HIL Dissertation. Leiden: Holland Academic Graphics.

Bar-Shalom, E. and W. Snyder (1998). "Optional infinitives in child Russian: a comparison with Italian and Polish." *Proceedings of the Generative Approaches to Language Acquisition (GALA 1997)*, Edinburgh, Scotland.

Bar-Shalom, E. and W. Snyder (1999). "On the relationship between root infinitives and imperatives in early child Russian." In: *Proceedings BUCLD 23*. Medford MA: Cascadilla Press.

Becker, M. (1999). *The use of modal verbs and modality in child German*. Unpublished manuscript, UCLA.

Behrens, H. (1993). *Temporal interpretation in German child language*. Dissertation UvA, Amsterdam.

Besten, H. den. (1983). "On the interaction of root transformations and lexical deletive rules." In: W. Abraham (Ed.) *On the formal syntax of the Westgermania: papers from the third Groningen Grammar Talks*. Amsterdam/Philadelphia: John Benjamins.

Bickerton, D. (1981). *Roots of language*. Ann Arbor: Karoma Press.

Blom, E. (2002). "On the use and interpretation of root infinitives in early child Dutch." In: S. Barbiers, F. Beukema and W. van der Wurff (Eds.) *Modality and its interaction with the verbal system*." Amsterdam/Philadelphia: John Benjamins.

Blom, E. and E. Krikhaar (2002). "Where to use a nonfinite clause? An experiment on the use of verb forms in early Dutch and English child language." In: J. Costa and J.M. Freitas (Eds.) *Proceedings of the Generative Approaches to Language Acquisition (GALA 2001)*, Palmela, Portugal.

Blom, E., E. Krikhaar and F. Wijnen (2001). "Root infinitives in Dutch and English child language: an experimental approach." *Proceedings BUCLD 25*. Medford MA: Cascadilla Press.

Blom, E. and F. Wijnen (2000). "How Dutch children's root infinitives become modal." *Proceedings BUCLD 24*. Medford MA: Cascadilla Press.

Blom, E., F. Wijnen and S. Gillis (1998). "What they hear is what you get?: Infinitives and modality in child language and child-directed speech." In: M. Perkins and S. Howard (Eds.) *New directions in language development and disorders*. New York: Kluwer/Plenum.

Bloom, L., K. Lifter and J. Hafitz (1980). "Semantics of verbs and the development of verb inflection in child language." *Language 56(2)*: 386 - 412.

Bloom, L., J. Tackeff and M. Lahey (1984). "Learning 'to' in complement clauses." *Journal of child language* (11): 391 - 406.

Bloom, P. (1990). "Subjectless sentences in child language." *Linguistic Inquiry* 21(4): 491 - 504.

Bode, S. de (1999). *The RI phenomenon in Russian*. Unpublished manuscript, UCLA.

Boezewinkel, S. (1995). *Van mij doen tot haar zusje had twee vriendinnetjes te spelen; een onderzoek naar de verwerving van pronomina in de Nederlandse kindertaal.* MA-thesis, Utrecht University.

Bol, G. and F. Kuiken (1986). "Het gebruik van pronomina bij kinderen van een tot vier." *Psycholinguistiek en stoornissen: toegepaste taalwetenschap in artikelen* 24: 47 - 58.

Boser, K., B. Lust, L. Santelmann and J. Whitman (1992). "The syntax of CP and V2 in early child German (ECG). The strong continuity hypothesis." *Proceedings of NELS 22*.

Bromberg, H. and K. Wexler (1995). "Null subjects in child wh-questions". In: *MIT Working Papers in Linguistics 26*: 221 – 247.

Bronckart, J. P. and H. Sinclair (1973). "Time, tense and aspect." *Cognition* 2(1): 107 - 130.

Brun, D., S. Avrutin and M. Babyonyshev (1999). "Aspect and its temporal interpretation during the optional infinitive stage in Russian." *Proceedings of BUCLD 23*, Medford MA: Cascadilla Press.

Bybee, J. and S. Fleischman (1995). Modality in grammar and discourse: an introductory essay. In: J. Bybee and S. Fleischman (Eds.) *Modality in grammar and discourse.* Amsterdam: Benjamins.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge MA: MIT Press.

Carlson, G.N (1977). *Reference to kinds in English.* Dissertation, University of Massachusetts, Amherst. Published 1980 by Garland Press: New York.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language: its nature, origin and use.* New York: Praeger.

Chomsky, N. (1993). "A minimalist program for linguistic theory." In: K. Hale and S. Keyser (Eds.) *The View from Building 20.* Cambridge MA: MIT Press.

Chung, S. and A. Timberlake (1985). "Tense, aspect and mood." In: T. Shopen (Ed.) *Language typology and syntactic description. Volume III: Grammatical categories and the lexicon.* Cambridge: Cambridge University Press.

Clahsen, H. (1982). *Spracherwerb in der Kindheit.* Dissertation, University of Tübingen.

Clahsen, H. (1990). "Constraints on parameter setting: a grammatical analysis of some acquisitional stages in German child language." *Language acquisition* (1): 361 - 391.

Clahsen, H., S. Eissenbeiss and M. Penke (1994). "Lexical learning in early syntactic development." In: T. Hoekstra and B. Schwartz (Eds.) *Language acquisition studies in generative grammar.* Amsterdam/Philadelphia: John Benjamins.

Clahsen, H. and M. Penke (1992). "The acquisition of agreement morphology and its syntactic consequences: new evidence on German child language from the Simone-corpus." In: J. Meisel (Ed.) *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition.* Dordrecht: Kluwer Academic Publishers.

Clahsen, H. and K.-D. Smolka (1986). "Psycholinguistic evidence and the description of V2 phenomena in German." In: H. Haider and M. Prinzhorn (Eds.)*Verb second phenomena in Germanic languages.* Dordrecht: Kluwer Academic Publishers.

Clark, E. (1987). "The principle of contrast: a constraint on language acquisition." In: B. MacWhinney (Ed.) *Mechanisms of language acquisition.* New York: Springer-Verlag.

Comrie, B. (1976). *Aspect: an introduction to the study of verbal aspect and related problems.* Cambridge MA: Cambridge University Press.

Comrie, B. (1985). *Tense.* Cambridge MA: Cambridge University Press.

Cornips, L. (2000). "The use of gaan + infinitive in the narratives of older bilingual children of Moroccan and Turkish descent." In: T. van der Wouden and H. de Hoop (Eds.) *Linguistics in the Netherlands.* Amsterdam: John Benjamins.

Crisma, P. (1992). "On the acquisition of Wh-questions in French." *Geneva Generative Papers* 0(1), University of Geneva.

De Cat, C. (2002). *French dislocation.* Dissertation, University of York.

Deen, K. U. (1997). *The interpretation of root infinitives in English: is eventivity a factor?.* Unpublished manuscript, UCLA.

Deen, K. U. and N. Hyams (1999). "The forms and interpretation of finite and nonfinite verbs in Swahili." *Proceedings BUCLD 26.* Medford MA: Cascadilla Press.

Dowty, D. (1979). *Word meaning and montague grammar.* Dordrecht: Reidel.

Enç, M. (1987). "Anchoring conditions for tense." *Linguistic Inquiry* 18(4): 633 - 657.

Erb, M. C. (2001). *Finite auxiliaries in German.* Dissertation Tilburg University

Evers, A. (1981). "Two functional principles for 'Move V'." In: W. Abraham (Ed.) *GAGL* 19: 96 - 110.

Evers, A. and J. van Kampen (1995). "Do-insertion and LF in child language." In: J. Don, B. Schouten and W. Zonneveld (Eds.) *OTS Yearbook 1994.* Utrecht University.

Evers-Vermeul, J. (1997). *Pronomina en casus.* MA-thesis, Utrecht University.

Felser, C. (2000). "Aspectual complement clauses and the (un-)availability of verb raising.: In: H. Janssen (Ed.) *Verbal projections.* Tübingen: Niemeyer.

Ferdinand, A. (1996). *The development of functional categories. The acquisition of the subject in French.* HIL Dissertation, Leiden University.

Flavell, J. H., E. R. Flavell, F. L. Green and L. J. Moses (1990). "Young children's understanding of fact beliefs versus value beliefs." *Child development* 61: 915 - 928.

Fries, N. (1999). *Wie finit ist der Imperativ und wie infinit darf er sein? Oder auch: Wie infinit ist der Imperativ und wie finit muss er sein?* Paper presented at the Tagung der Deutsche Gesellschaft für Sprachwissenschaft 2000, Marburg, Germany.

Gavruseva, E. (2001). *Aspect, eventivity and finiteness in early child grammars.* Unpublished manuscript, University of Iowa.

Gavruseva, E. (2002). "Aspect, eventivity, and finiteness in early child grammar." In: J. Costa and M.J. Freitas (Eds.) *Proceedings of the Generative Approaches to Language Acquisition (GALA 2001)*, Palmela, Portugal.

Gerhardt, J. (1991). "The meaning and use of the modals HAFTA, NEEDTA and WANNA in children's speech." *Journal of Pragmatics* 16: 531 - 590.

Ginneken, J. van (1917). *De roman van een kleuter*. Nijmegen: Malmberg.

Giorgi, A. and F. Pianesi (1996). *Tense and Aspect: from semantics to morphosyntax*. New York [etc.], Oxford University Press.

Gonsalves, J. (1998). *Relations between Conceptual and Semantic Development: Preschoolers' Understanding of Modality across Linguistic and Nonlinguistic domains*. Dissertation, Clark University.

Good, P. (1999). *Resampling methods: a practical guide to data analysis*. Boston: Birkhauser.

Gopnik, A. (1993). "How we know our minds: The illusion of first-person knowledge of intentionality." In: A. Goldmand (Ed.) *Readings in philosophy and cognitive science*. Cambridge MA: MIT Press.

Gordon, P. (1986). "Level-ordering in lexical development." *Cognition* 21: 73 - 93.

Grimshaw, J. (1994). "Minimal projections and clause structure." In: M. Suner, B. Lust and J. Whitman (Eds.) *Syntactic theory and first language acquisition: cross-linguistic perspectives. Volume I: Heads, projections, and learnability*. Hillsdale N.J: Lawrence Erlbaum Associates.

Grodzinsky, J. and T. Reinhart (1993). "The innateness of binding and coreference." *Linguistic Inquiry* 24: 69 - 102.

Guasti, M. T. (1994). "Verb syntax in Italian child language: finite and nonfinite verbs." *Language acquisition* 3: 1 - 40.

Gueron, J. and T. Hoekstra (1995). "The temporal interpretation of predication." In: A. Cardinaletti and M. T. Guasti (Eds.) *Syntax and Semantics 28*. San Diego: Academic Press.

Guilfoyle, E. and M. Noonan (1992). "Functional categories and language acquisition." *Canadian journal of linguistics* 37: 241 - 272.

Haan, A. de (1996). *De verwerving van morfologische finietheid in het Nederlands*. MA-thesis, Groningen University.

Haan, G. de (1987). "A theory-bound approach to the acquisition of verb placement in Dutch." In: G. de Haan and W. Zonneveld (Eds.) *Formal parameters of generative grammar: OTS Yearbook 1987*. Dordrecht: ICG.

Haan, G. de and K. Tuijnman (1988). "Missing subjects and objects in child grammar." In: P. Jordens and J. Lallemans (Eds.) *Language Development.* Dordrecht: Foris.

Haegeman, L. (1994). "Root infinitives, tense and truncated structures." *Language acquisition* 4(3): 205-255.

Haegeman, L. (1995). "Root null subjects and root infinitives in early Dutch." In: F. Wijnen and C. Koster (Eds.) *Groningen Assembly on Language Acquisition (GALA 1995).* Groningen: Center for Language and Cognition Groningen.

Halle, M. and A. Marantz (1993). "Distributed morphology and the pieces of inflection." In: K. Hale and S. Keyser (Eds.) *The view from building 20.* Cambridge MA: Cambridge University Press.

Halliday, M. A. K. (1975). *Learning how to mean. Explorations in the Development of Language.* London: Edward Arnold.

Hamann, C. and K. Plunkett (1998). "Subjectless sentences in child Danish." *Cognition* 69: 35 - 72.

Harris, T. and K. Wexler (1996). "The optional-infinitive stage in child English." In: H. Clahsen (Ed.) *Generative perspectives on language acquisition.* Amsterdam/Philadelphia: John Benjamins.

Hirst, W. and J. Weil (1982). "Acquisition of epistemic and deontic meanings of modals." *Journal of child language* 9: 659 - 666.

Hoekstra, T. and N. Hyams (1995). "Missing heads in child language." In: F. Wijnen and C. Koster (Eds.) *Groningen Assembly on Language Acquisition (GALA 1995).* Groningen: Center for Language and Cognition Groningen.

Hoekstra, T. and N. Hyams (1998). "Aspects of root infinitives." *Lingua* 106: 81 - 112.

Hoekstra, T. and P. Jordens (1994). "From Adjunct to Head." In: T. Hoekstra and B. Schwartz (Eds.) *Language acquisition in generative grammar.* Amsterdam/Philadelphia: John Benjamins.

Hollebrandse, B. and T. Roeper (1995). "The concept of Do-insertion and the theory of INFL in acquisition." In: F. Wijnen and C. Koster (Eds.) *Groningen Assembly on Language Acquisition (GALA 1995).* Groningen: Center for Language and Cognition Groningen.

Holmberg, A. and C. Platzack (1991). "On the role of inflection in Scandinavian syntax." In: W. Abraham, W. Kosmeijer and E. Reuland (Eds.) *Issues in Germanic syntax.* Berlin/New York: Mouton de Gruyter.

Hyams, N. (1996). "The underspecification of functional categories in early grammar." In: H. Clahsen (Ed) *Generative approaches to first and second language acquisition.* Amsterdam/Philadelphia: John Benjamins.

Hyams, N. (1999). "Underspecification and modularity in early syntax." In: M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan and K. Wheatley (Eds.) *Functionalism and formalism in linguistics.* Amsterdam/Philadelphia: John Benjamins.

Hyams, N. (2001a). "Finiteness, aspect and mood in early grammar: a cross-linguistic perspective." *Proceedings of CLS 36.*

Hyams, N. (2001b). "Now you hear it now you don't: the nature of optionality in child language." *Proceedings of BUCLD 25.* Medford MA: Cascadilla Press.

Hyams, N. and K. Wexler (1993). "On the grammatical basis of null subjects in child language." *Linguistic Inquiry* 24: 421-459.

Ingram, D. and W. Thompson (1996). "Early syntactic acquisition in German: evidence for the modal hypothesis." *Language* 72(1): 97 - 120.

Jackendoff, R. (1983). *Semantics and cognition.* Cambridge MA: MIT Press.

Jaeggli, O. and N. Hyams (1988). "Morphological uniformity and the setting of the null subject parameter." *Proceedings of NELS 18.*

Jong, J. de (1999). *Specific Language Impairment in Dutch: inflectional morphology and argument structure.* Groningen Dissertations in Linguistics 28, Groningen University.

Jordens, P. (1990). "The acquisition of verb placement." *Linguistics* 28: 1407 - 1448.

Kampen, J. van (1997). *First steps in WH movement.* Dissertation Utrecht Institute of Linguistics OTS.

Kampen, J. van (2001). "Bootstraps at two for lexicon and discourse." Unpublished manuscript, Utrecht Institute of Linguistics OTS.

Kaper, W. (1976). "Pronominal case-errors." *Journal of child language* 3: 439 - 441.

Kayne, R. (1994). *The antisymmetry of syntax.* Cambridge MA: MIT Press.

Kiparsky, P. (1973). ""Elsewhere" in phonology." In: S. Anderson and P. Kiparsky (Eds.) *A Festschrift for Morris Halle.* New York: Holt, Rinehart and Winston.

Klein, R., (1974). "Word order: Dutch children and their mothers." *Publications of the Institute of General Linguistics*, Department of Linguistics, University of Amsterdam.

Klein, W. (1994). *Time in language.* London/New York: Routledge.

Koeneman, O. (2000). *The flexible nature of verb movement*. Dissertation Utrecht Institute of Linguistics OTS, Utrecht University

Koster, J. (1975). "Dutch as an SOV language." *Linguistic Analysis* 1: 111 - 136.

Krämer, I. (1993). "The licensing of subjects in early child language." In: C. Phillips (Ed.) *Papers on case and agreement II: MIT working papers in linguistics 19*.

Kratzer, A. (1989). "Stage and Individual level predicates." In: G. N. Carlson and F. J. Pelletier (Eds.) *The generic book*. Chicago/London: The University of Chicago Press.

Kratzer, A. (1991). "Modality." In: A. von Stechow and D. Wunderlich (Eds.) *Semantics: an international handbook of contemporary research*. Berlin: De Gruyter.

Kuczaj, S. A. and M. Maratsos (1975). "What a child 'can' say before he 'will'." *Merril-Palmer Quarterly* 21: 89 - 111.

Kursawe, C. (1994). *Fragesätze in der deutsche Kindersprache*, University of Düsseldorf.

Lakoff, G. (1966). "Stative adjectives and verbs in English." In: A. G. Oettinger (Ed.) *Mathematical linguistics and automatic translation. Report NSF-17*. Cambridge MA, The Computation Laboratory, Harvard University.

Lasser, I. (1997). *Finiteness in adult and child German*. Dissertation, Max Planck Institute,. Nijmegen.

Lebeaux, D. (1988). *Language acquisition and the form of the grammar*. Dissertation, Amherst MA.

Levin, B. and M. Rappaport (Hovav) (1995). *Unaccusativity: at the syntax-lexical interface*. Cambridge MA: MIT Press.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

MacWhinney, B. (1995). *The CHILDES project: tools for analyzing talk*. Hillsdale NJ: Lawrence Erlbaum Associates.

Madsen, S. and J. Gilkerson (1999). *A pattern of differences between bare forms and root infinitives*. Unpublished manuscript, UCLA.

McDowell, J. (1987). *Assertion and modality*. Dissertation, University of Southern California.

Palmer, F. R. (1986). *Mood and modality*. Cambridge: Cambridge University Press.

Partee, B. (1984). "Nominal and temporal anaphora." *Linguistics and philosophy* 7: 243 - 286.

Peters, A. (1982). *The units of language acquisition.* Cambridge UK: Cambridge University Press.

Petter, M. (1998). *Getting PRO under control.* HIL Dissertations, Leiden University.

Phillips, C. (1995). "Syntax at age 2: crosslinguistic differences." *MIT Working Papers in Linguistics* 26: 325-382.

Pierce, A. (1989). *On the emergence of syntax: a cross-linguistic study.* Dissertation MIT.

Pierce, A. (1992). *Language acquisition and syntactic theory: a comparative analysis of French and English child grammars.* Dordrecht: Kluwer Academic Publishers.

Pine, J., E. Lieven and C. Rowland. (1998). *Tense optionality and children's use of English verb morphology: testing the Optional Infinitive Hypothesis.* Paper presented at the Child Language Seminar, Sheffield.

Pinker, S. (1984). *Language learnability and language development.* Cambridge MA: Harvard University Press.

Poeppel, D. and K. Wexler (1993). "The full competence hypothesis of clause structure in early German." *Language* 69(1).

Portner, P. (1999). "The semantics of mood." *GLOT International* 4(1): 3 - 9.

Powers, S. (1994). *The acquisition of case in Dutch.* Nijmegen: Max Plank Institute.

Powers, S. (1996). *The growth of the phrase marker: evidence from subjects.* Dissertation, University of Maryland.

Powers, S. (1998). "Binarity and singularity in child grammar." In: B. Hollebrandse (Ed.) *New perspectives on language acquisition.* UMOP 22, University of Massachusetts, Amherst.

Pustojevski, J. (1991). "The syntax of event structure." In: B. Levin and S. Pinker (Eds.) *Lexical and conceptual semantics.* Cambridge MA: MIT Press.

Quirk, R. (1985). A comprehensive grammar of the English language. New York.

Radford, A. (1988). "Small children's small clauses." *Transactions of the philological society* 86: 1 - 46.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: the nature of early child grammars of English.* Oxford UK: Basil Blackwell.

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

Reichenbach, H. (1947). *Elements of symbolic logic.* New York: Macmillan.

Rizzi, L. (1982). *Issues in Italian Syntax.* Dordrecht: Foris.

Rizzi, L. (1992). "Early null subjects and root null subjects." *Geneva Generative Papers P* 0(1): 102 - 115.

Rizzi, L. (1994). "Some notes on linguistic theory and language development: the case of root infinitives." *Language acquisition* 3(4): 371 - 393.

Rizzi, L. (1997). "The fine structure of the left periphery." In: L. Haegeman (Ed.) *Elements of grammar. Handbook of generative syntax.* Dordrecht: Kluwer Academic Publishers.

Roeper, T. (1996). "The role of Merger Theory and formal features in acquisition." In: H. Clahsen (Ed.) *Generative perspectives on language acquisition. Empirical findings, theoretical considerations, crosslinguistic comparisons.* Amsterdam/Philadelphia: John Benjamins.

Roeper, T. and B. Rohrbacher (1994). *Null subjects in early child English and the theory of Economy of Projection,* University of Massachusetts/University of Pennsylvania.

Rousseau, A. (1998). "Control and raising in and out of subjunctive complements." In: B. Joseph, A. Ralli and M. L. Rivero (Eds.) *Comparative syntax of Balkan languages.* Oxford: Oxford University Press.

Sano, T. and N. Hyams (1994). "Agreement, finiteness, and the development of null arguments." *Proceedings of NELS 24.*

Santelmann, L. (1994). *The acquisition of verb second in child Swedish.* Dissertation, Cornell University.

Schaeffer, J. (1990). *The syntax of the subject in child language.* MA-thesis, Utrecht University.

Schaerlakens, A. M. and S. Gillis (1987). *De taalverwerving van het kind: een hernieuwde orientatie in het Nederlandstalige onderzoek.* Groningen: Wolters - Noordhoff.

Schlichting, L. (1996). *Discovering of syntax. An empirical study in Dutch language acquisition.* Dissertation, Nijmegen University.

Schönenberger, M., A. Pierce, K. Wexler and F. Wijnen (1995). "Accounts of root infinitives and the interpretation of root infinitives." *Geneva Generative Papers* 3(2) 47 - 71.

Schütze, C. and K. Wexler (1996). Subject case licensing and English root infinitives. *Proceedings of BUCLD 20.* Somerville MA: Cascadilla Press.

Schütze, C. and K. Wexler (1999). An elicitation study of young English children's knowledge of tense: semantic and syntactic properties of optional infinitives. *Proceedings BUCLD 24,* Medford MA: Cascadilla Press.

Schüzte, C. (1997). *INFL in child and adult language: agreement, case and licensing.* Dissertation MIT.

Searle, J. R. (1983). *Intentionality: an essay in the philosophy of mind.* Cambridge: Cambridge University Press.

Shepherd, S. C. (1981). *Modals in Antiguan Creole, child language acquisition, and history.* Dissertation, Stanford University.

Slobin, D. I. (1973). "Cognitive prerequisites for the development of grammar." In: C. E. Ferguson and D. I. Slobin (Eds.) *Studies in child language development.* New York: Holt, Rinehart & Winston.

Slobin, D. I. (1977). "Language change in childhood and history." In: J. Macnamara (Ed.) *Language learning and thought.* New York: Academic Press.

Slobin, D. I. (1979). *The role of language in language acquisition.* Paper presented at the fiftieth Annual Meeting of the Eastern Psychological Association, Philadelphia.

Slobin, D. I. (1982). "Universal and particular in the acquisition of language." In: E. Wanner and L. Gleitman (Eds.) *Language acquisition: the state of the art.* Cambridge: Cambridge University Press.

Speas, M. (1994). "Null arguments in a theory of economy of projection." *University of Massachusetts Occasional Papers in Linguistics* 17: 179 - 208.

Steedman, M. (1977). "Verbs, time and modality." *Cognitive science* 1: 216 - 234.

Stephany, U. (1986). "Modality." In: P. Fletcher and M. Garman (Eds.) *Language acquisition.* Cambridge: Cambridge University Press.

Tóth, I. (2000). *Inflected infinitives in Hungarian.* Dissertation Tilburg University.

Travis, L. (1994). "Event phrase and a theory of functional categories." *Canadian Linguistic Association Meeting,* University of Calgary.

Vendler, Z. (1967). *Linguistics in philosophy.* Ithaca NY: Cornell University Press.

Verhulst - Schlichting, L. (1985). "De ontwikkeling van het werkwoord: plaats, vorm en type." *Interdisciplinair Tijdschrift voor Taal- en Tekstwetenschap* 5(3): 285 - 298.

Verkuyl, H. (1972). *On the compositional nature of the aspects.* Dordrecht: Reidel.

Wanner, A. (1999). *Verbklassifizierung und aspektuelle Alternationen im Englischen.* Tübingen: Niemeyer.

Weissenborn, J. (1994). "Constraining the child's grammar: Local Wellformedness in the development of verb movement." In: B. Lust, M. Suner and J. Whitman (Eds.) *Syntactic theory and first language acquisition.* Hillsdale N.J: Lawrence Erlbaum Associates.

Wellman, H. M. (1990). *The child's theory of mind.* Cambridge MA: Bradford/MIT Press.

Wellman, H. M. and J. D. Woolley (1990). "From simple desires to ordinary beliefs. The early development of everyday psychology." *Cognition* 35: 245 - 275.

Wexler, K. (1992). "Optional infinitives, head movement, and the economy of derivations in child grammar." *Occasional paper, MIT dept. Brain & Cognitive Science 45.*

Wexler, K. (1994). "Finiteness and head movement in early child language." In: N. Hornstein and D. Lightfoot (Eds.) *Verb movement.* Cambridge MA: Cambridge University Press.

Wijnen, F. (1997). "Temporal reference and eventivity in root infinitives." *MIT Occasional papers in linguistics* 12: 1 - 25.

Wijnen, F. (1999). *Verb placement in Dutch child language: a longitudinal analysis.* Unpublished manuscript, Utrecht Institute of Linguistics OTS.

Wijnen, F. (2000). "Input, intake and sequence of syntactic development." In: M. Beers, B. van den Bogaerde, G. Bol, J. de Jong and C. Rooijmans (Eds.) *From sound to sentence.* Center for Language and Cognition, Groningen.

Wijnen, F. and G. Bol (1993). "The escape from the optional infinitive stage." In: A. de Boer, J. de Jong and R. Landeweerd (Eds.) *Language and Cognition 3: Yearbook of the research group for theoretical and experimental linguistics of the University of Groningen .*

Wijnen, F. and L. Elbers (1998). *Sequence in syntactic development: an output-as-input approach*, Unpublished manuscript, Utrecht Institute of Linguistics OTS/Psychological Laboratory Utrecht University.

Wijnen, F., M. Kempen and S. Gillis (2001). "Bare infinitives in Dutch early child language: an effect of input?" *Journal of child language* 28(3).

Wimmer, H. and J. Perner (1983). "Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception." *Cognition* 13: 103 - 128.

Zuckerman, S. (2001). *The acquisition of "optional" movement.* Groningen Dissertations in Linguistics 34. Groningen University.

Zwart, J.-W. (1993). *Dutch syntax: a minimalist approach.* Groningen Dissertations in Linguistics 10. Groningen University.

Zwart, J.-W. (1997). *Morphosyntax of verb movement. A minimalist approach to the syntax of Dutch.* Dordrecht: Kluwer Academic Publishers.

SAMENVATTING IN HET NEDERLANDS

Deze dissertatie gaat over de verwerving van finiete zinnen, meer specifiek over de verwerving van werkwoordelijke inflecties en hulpwerkwoorden. Normaal ontwikkelende Nederlandstalige kinderen leren de basis voor het inflectie- en hulpwerkwoordsysteem van hun doeltaal wanneer ze ruwweg tussen de twee en drie jaar oud zijn. De eerste stap in deze ontwikkeling wordt gemarkeerd door het verschijnen van zogenaamde 'root infinitives' vaak afgekort tot RIs. Deze zinnen bevatten een werkwoord, i.e. de potentiële drager van finietheidskenmerken, maar het werkwoord verschijnt aanvankelijk louter in niet-finiete vorm. Zinnen als *Bob op bank zitten* of *Schoen aantrekken* zijn karakteristieke voorbeelden van dergelijke RIs: een hulpwerkwoord (vgl.: *Bob wil op de bank zitten* of *Bob moet op de bank zitten*) of inflectie (vgl.: *Bob zit op de bank* of *Ik trek een schoen aan*) ontbreekt. Na dit eerste stadium doorlopen kinderen een aantal volgende stadia die gekenmerkt worden door het gebruik van RIs en verschillende soorten finiete zinnen naast elkaar. Een terugkerende vraag in deze dissertatie is wat de status is van RIs in de taal van jonge kinderen.

In de hoofdstukken 3, 4 en 5 wordt geconcludeerd dat RIs uitingen zijn zonder een specifieke modale of temporele referentie, maar met een aspectuele betekenis. De kinderen gebruiken RIs zolang ze nog geen werkwoordvormen kennen die een gespecificeerde modale of temporele betekenis hebben. De gespecificeerde alternatieven voor RIs. i.e. uitspellingen van de functionele categorieën Modaliteit en Tempus, worden door de kinderen op een stapsgewijze manier geleerd. Vastgesteld wordt dat de lexicale encoderingswijze van Tempus door middel van een hulpwerkwoord door Nederlandse kinderen eerder geleerd wordt dan de grammaticale encoderingswijze door middel van inflectie. Het gevolg is dat (i) overgeneralisatie plaatsvindt van hulpwerkwoord-gebruik en niet alleen Modaliteit maar ook Tempus lexicaal uitgespeld wordt, en (ii) dat modale onderscheidingen geformaliseerd worden voor temporele onderscheidingen. Omdat inflectie in het Nederlands niet alleen de propositie voorziet van een tijdsindex, maar ook subject en werkwoord aan elkaar relateert door middel van congruentie, wordt aandacht besteed aan veranderingen in subjectgebruik die het gevolg kunnen zijn van de verwerving van inflectionele morfologie.

In de hoofdstukken 3, 4 en 5 worden uitingen van zes Nederlandstalige kinderen tussen de 2 en 3, 5 (i.e. de RI-periode) geanalyseerd. In de hoofdstukken 3 en 4 wordt geabstraheerd van veranderingen die plaatsvinden naarmate de kinderen ouder worden: observaties van verschillende meet-momenten worden samengenomen. In hoofdstuk 5 staan RIs en finiete zinnen in verschillende ontwikkelingsstadia centraal. Verder wordt Nederlandse kindertaal vergeleken met andere kindertalen. In het bijzonder wordt een vergelijking getrokken tussen Nederlands- en Engelstalige kinderen op grond van uitingen ontlokt in een experimentele situatie. De hoofdstukken 3, 4 en 5 worden vooraf gegaan door een eerste hoofdstuk met een literatuuroverzicht en een tweede hoofdstuk met uitleg over de gehanteerde statistische methode en geanalyseerde taaldata. Het zesde hoofdstuk vat de belangrijkste conclusies en bevindingen samen, evalueert eerder voorgestelde hypotheses en geeft de implicaties voor toekomstig onderzoek.

HOOFDSTUK 1 PERSPECTIEVEN OP ROOT INFINITIVES   In dit hoofd-stuk wordt een uitgebreid overzicht gegeven van de verschillende verklaringen voor het gebruik van RIs door jonge kinderen en voor de kenmerken van RIs. De relevante literatuur is onderverdeeld in drie typen benaderingen. Volgens een groep van wetenschappers zijn RIs het resultaat van beperkingen in de kindergrammatica. Deze beperking uit zich in het systematisch ontbreken van functionele categorieën. Volgens een tweede groep van wetenschappers is er geen fundamenteel verschil tussen de grammaticale kennis van kinderen en volwassenen. Er wordt betoogd dat RIs het gevolg zijn van het ontbreken van niet-grammaticale kennis. Kinderen weten bijvoorbeeld niet dat (of waar) finietheid verplicht uitgedrukt moet worden of ze kennen de lexicale items niet die finietheid uitdrukken, waardoor de functionele categorie [αfiniet] (IP of Tense) ondergespecificeerd blijft in vroege kindertaal. Een derde tak van onderzoek relateert RIs en eigenschappen van RIs aan patronen in het taalaanbod.

HOOFDSTUK 2 STATISCHE METHODES   In veel gevallen zijn de kindertaaldata die besproken worden in deze dissertatie kwantitatief te marginaal om standaard statistische significantietoetsen te gebruiken. Daarom wordt gebruikt gemaakt van de permutatie-toets om de significantie van de gedane observaties te bepalen. Het voordeel is dat voor gebruik van deze toets geen extra aannames gedaan hoeven worden over de verdeling van de geobserveerde data. Volgens deze methode wordt een geobserveerd verschil

afgezet tegen een verschil berekend op basis van toevalsverdelingen. Deze toevalsverdelingen komen tot stand door de gedane observaties willekeurig te verdelen over posities. Op grond van een herhaling van deze toevalsprocedure (bijv. 1000 of 5000 keer) wordt de waarschijnlijkheid bepaald dat het geobserveerde verschil afwijkt van het toevalsverschil (i.e. de p-waarde). Is de p-waarde hoog, i.e. hoger dan de significantiedrempel van 0.05, dan wordt de H0 bevestigd: het is waarschijnlijk dat er geen significant verschil is tussen de geobserveerde verdeling en de toevalsverdeling. Is de p-waarde laag, i.e. lager dan de significantiedrempel van 0.05, dan wordt de H0 niet bevestigd: het is waarschijnlijk dat er verschil is tussen de twee verdelingen. Deze wijze van toetsen wordt toegepast om te bepalen of er (i) een betekenisverschil is tussen de RIs van Nederlands- en Engelstalige kinderen en (ii) om te bepalen of er een verschil is tussen ontwikkelingsstadia. Om vast te stellen of generalisatie over de zes Nederlandstalige kinderen die in deze dissertatie bestudeerd worden, mogelijk is, wordt een zogenaamde 'Jack-knife' techniek toegepast: er wordt berekend of weglating van een enkel kind - steeds een ander kind uit de groep van kinderen - leidt tot een vertekening van het groepsgemiddelde. Als dit niet het geval is, kan gegeneraliseerd worden over de groep van kinderen.

HOOFDSTUK 3 TEMPORELE, MODALE EN ASPECTUELE DENOTATIE
Volgens de Geen Tense Hypothese (*No Tense Hypothesis*) hebben RIs geen gefixeerde temporele verankering. De voorspelling is dat RIs gebruikt worden om handelingen te beschrijven die voor, tijdens en na het spreekmoment plaatsgevonden hebben, plaatsvinden of zullen plaatsvinden. Volgens de Modaal Hypothese (*Modal Hypothesis*) zijn RIs modaal. Er zijn twee varianten van de Modaal Hypothese: RIs bevatten een niet-uitgesproken modaal hulpwerkwoord (*Modal Drop Hypothesis*) of het infinitiefsuffix [*-en*] in RIs heeft een modale betekenis (*Infinitival Morphology Hypothesis* of IMH).. Volgens de Niet-Compleetheidsbeperking (*Non-Completedness Constraint*) verwijzen RIs zonder uitzondering naar niet afgeronde handelingen. Spontane taaldata van de zes Nederlandstalige kinderen laten zien RIs met een modale betekenis in de meerderheid zijn ten opzichte van RIs met een tegenwoordige tijd verwijzing. Het aantal RIs dat verwijst naar een handeling die afgerond is ten tijde van het spreekmoment, is verwaarloosbaar. Hoewel modale RIs in de meerderheid zijn, zijn er teveel tegenwoordige-tijd-RIs om de Modaal Hypothese te bevestigen. Er is wel sprake van een Modaal Referentie Effect (*Modal Reference Effect* of MRE). De Geen Tense Hypothese wordt bevestigd omdat er geen

gefixeerde temporele betekenis is. Nederlandse RIs vertonen een Niet-Compleetheidsbeperking

Een van de argumenten die gegeven is voor de IMH is dat Engelse RIs geen MRE vertonen: dit wordt verklaard als het gevolg van het ontbreken van infinitiefmorfologie in het Engels. In het derde hoofdstuk is het verschil in betekenis tussen Nederlandse en Engelse RIs nader onderzocht door middel van een experiment. De experimentele methode voorkomt een aantal methodologische bezwaren van corpusonderzoek en maakt een valide taalvergelijking mogelijk. De resultaten van het experiment bevestigen een betekenisverschil tussen Nederlandse en Engelse RIs: de Nederlandse RIs zijn vaker modaal dan de Engelse. Preciezer uitgedrukt: de kans dat Nederlandse kinderen een RI voor modale betekenissen gebruiken is groter dan de kans dat Engelse kinderen een RI voor modale betekenissen aanwenden.

Het geobserveerde verschil tussen Engels en Nederlands is echter aanzienlijk kleiner dan het verschil zoals dat gesuggereerd wordt door de resultaten van corpusonderzoek. Wat verklaart het verschil tussen de twee onderzoeksmethodes? Engelse corpus RIs zijn beperkt tot uitingen met het onderwerp in derde persoon enkelvoud, omdat alleen deze vorm de mogelijkheid biedt tot een onderscheid tussen finiete uitingen en RIs. Nederlandse corpus RIs, daarentegen, bevatten alle mogelijke onderwerpen. Het experiment vertoont deze asymmetrie tussen de twee talen niet, omdat alle uitingen in het experiment een onderwerp in derde persoon enkelvoud hebben. Omdat juist uitingen met een onderwerp in eerste en tweede persoon enkelvoud een correlatie vertonen met modaliteit ('ik wil' en 'jij moet'), leidt de asymmetrie tussen de twee talen in de geselecteerde corpusdata tot een aanzienlijk betekenisverschil tussen RIs in de twee talen. Her-analyse van de Nederlandse corpus RIs laat zien dat weglating van RIs met onderwerpen in eerste en tweede persoon enkelvoud leidt tot een significante daling van het aantal modale RIs. Na her-analyse is de grootte van het betekenisverschil tussen RIs in Engelse en Nederlandse kindertaal in corpusdata meer vergelijkbaar met het verschil gevonden in het experiment.

Eerder in dit hoofdstuk is op grond van Nederlandse corpusdata de conclusie getrokken dat infinitiefmorfologie niet verantwoordelijk kan worden gehouden voor de betekenis van Nederlandse RIs. Deze conclusie heeft tot gevolg dat de eerder gegeven verklaring voor betekenisverschillen tussen Engelse en Neder-

landse RIs, niet meer geldt. Ter verklaring van het betekenisverschil tussen Engelse en Nederlands RIs is gewezen op het Heterogene Set Effect (*Heterogeneous Set Effect* of HSE) dat wel een rol speelt in Engels maar niet in Nederlands. Het gevolg van het HSE is dat de set van Engelse RIs niet alleen niet-finiete uitingen bevat, maar ook finiete uitingen waarvan inflectie is weggelaten. Dat jonge kinderen de neiging hebben om inflectie weg te laten in finiete contexten is aangetoond via de Nederlandse corpusdata. Echter, doordat Nederlands een SOV/V2 taal is met aparte infinitiefmorfologie en het Engels een SVO taal is zonder infinitiefmorfologie, zijn dergelijke uitingen in het Nederlands geen RIs maar in het Engels wel. Deze verklaring voor het betekenisverschil tussen Nederlandse en Engelse RIs wordt ondersteund door een aantal onafhankelijke observaties (betrekking hebbend op werkwoords-type, vraagwoordgebruik en topicalisatie) die erop wijzen dat Engelse RIs meer verwant zijn aan simpele finiete zinnen dan Nederlandse RIs. Het is onduidelijk of het HSE daadwerkelijk een *voldoende* verklaring biedt voor het betekenisverschil tussen Engelse en Nederlandse RIs.

HOOFDSTUK 4 TYPES VAN WERKWOORDEN   Eerder onderzoek heeft uitgewezen dat kinderen in RIs systematisch andere werkwoordelijke predi-caten gebruiken dan in hun zinnen met een verbogen finiet hoofdwerkwoord, i.e. simple finiete zinnen. Er zijn verschillende generalisaties voorgesteld om dit verschil te beschrijven. De Geen Overlap Hypothese (*No Overlap Hypothesis*) stelt dat er geen enkele overlap is tussen de werkwoordstypes in RIs en finiete zinnen met een finiet hoofwerkwoord. De corpusdata van Abel, Daan, Josse, Laura, Matthijs en Peter laten zien dat er wel overlap is. Naast de Geen Overlap Hypothese zijn er hypotheses die een asymmetrie voor stellen tussen werkwoorden gebruikt als finiete hoofdwerkwoorden en infinitieven. Deze hypotheses vallen uiteen in een Teliciteit Hypothese (*Telicity Hypothesis*) en een Dynamiciteit Hypothese (*Dynamicity Hypothesis*). De resultaten wijzen in richting van de Dynamiciteit Hypothese, omdat RIs vrijwel altijd een predicaat bevatten dat een gebeurtenis (*event*) denoteert en slechts zelden een predicaat bevatten dat een toestand (*state*) denoteert. Omdat toestandsbeschrijvende RIs niet uitgesloten zijn, wordt de conclusie getrokken dat er sprake is van een zwakke restrictie.

In de literatuur worden twee verschillende verklaringen gegeven voor de beperking van RIs tot eventieve predicaten: de een is gebaseerd op de Geen Tense Hypothese, de andere op de Modaal Hypothese (beide besproken in het

voorgaande hoofdstuk). De Geen Tense Benadering van de event-beperking kan echter niet verklaren waarom RIs dezelfde event-beperking vertonen als finiete uitingen met een perifrastisch werkwoord (i.e. een complexe werkwoordsvorm bestaande uit hulpwerkwoord en infinitief), terwijl de verklaring die de event-beperking in RIs relateert aan het overwegend modale gebruik van RIs deze overeenkomst wel kan verklaren. Verder kan de Geen Tense Benadering niet de overeenkomsten verklaren tussen volwassen en kindertaal met betrekking tot de event-beperking, terwijl de Modale Benadering ook deze overeenkomst kan verklaren. Het laatste voordeel dat de Modale Benadering heeft is dat ze de zwakte van de event-restrictie kan verklaren, en wel op twee manieren. Zoals geobserveerd in het voorgaande hoofdstuk, zijn niet alle RIs modaal. Een aantal niet-eventieve RIs is inderdaad niet-modaal. Verder is de Modale Benadering van werkwoordsselectie in RIs niet perse unificerend omdat er verschillende modaliteiten gebruikt worden in RIs. Bepaalde modale betekenissen - deontische modaliteit/commando's - sluiten statieve predicaten uit. Andere modale betekenissen - intentionaliteit - leiden in combinatie met een statief predicaat tot een soort van modaliteit die kinderen tussen de twee en drie door cognitieve beperkingen niet gebruiken, i.e epistemische modaliteit. De voorspelling dat statieve RIs beperkt zijn tot RIs met een tegenwoordige tijdsreferentie of tot RIs die een wens uitdrukken, komt uit.

Nederlandse RIs zijn verder zo vaak eventief als een gevolg van patronen in het taalaanbod. Nader onderzoek van het taalaanbod dat de zes Nederlandstalige kinderen laat zien dat deze kinderen statieve werkwoorden voornamelijk horen als geïnflecteerde hoofdwerkwoorden, terwijl ze eventieve werkwoorden voornamelijk horen als infinitieven. Zolang de kinderen nog niet zelf de mogelijkheid hebben om nieuwe werkwoordsvormen samen te stellen, zijn de vormen die ze gebruiken directe weerspiegelingen van de werkwoordsvormen in het taalaanbod. Op grond van het taalaanbod is de geobserveerde asymmetrie in werkwoordstype tussen finiete zinsinitiële werkwoorden en zinsfinale infinitieven verwacht.

HOOFDSTUK 5 VERANDERINGEN IN DE TIJD   In de vorige twee hoofdstukken zijn data verzameld over een langere periode (i.e. circa een jaar) samengenomen. In dit vijfde hoofdstuk wordt deze periode opgedeeld in een viertal ontwikkelingsstadia. Er wordt getoond dat finiete zinnen een stapsgewijze opkomst en groei vertonen. Aanvankelijk ontbreken alle verschijningsvormen van finietheid: er is noch inflectie noch zijn er hulpwerk-

woorden. De eerste finiete vormen die verschijnen zijn modale werkwoorden en koppelwerkwoorden. Deze werkwoorden dragen geen uiterlijke kenmerken van inflectie. Verder hebben zij geen andere lexicale betekenis dan het verschaffen van een modale- of tijdsbetekenis. Op grond hiervan wordt de conclusie getrokken dat Finietheid (Tempus) en Modaliteit in dit stadium lexicaal zijn en dat Finietheid nog niet gegrammaticaliseerd is tot inflectie. Deze ontwikkeling is voorspeld door de Grammaticalisatie van Finietheid Hypothese (*Grammaticalisation of Finiteness Hypothesis* of GoF). Ook een tweede ontwikkelingsvolgorde is voorspeld door de GoF. Het lexicale stadium bestaat uit twee sub-stadia. In sub-stadium 1 zijn de lexicale finietheidsmarkeerders simplex, in sub-stadium 2 worden ze gecombineerd met infinitieven tot perifrastische vormen. Na het lexicale stadium volgt het grammaticale stadium: inflectie wordt verworven. Inflectie introduceert werkwoordverplaatsing en geeft het kind de mogelijkheid economisch om te gaan met syntactische representaties. Inflectie (en daarmee de grammaticale categorie Tempus) wordt door Nederlandse kinderen pas laat verworven. Dit is het gevolg van (i) de weinig saillante kenmerken van inflectie in het Nederlands, (ii) relatief weinig lexicale variatie in de set van geïnflecteerde vormen in het taalaanbod, en (iii) weinig lexicale overlap tussen geïnflecteerde vormen en niet-geïnflecteerde infinitieven. Deze drie factoren vertragen het proces van segmentatie en daarmee de verwerving van inflectie. Een inventarisatie van de lexicale overlap in de vier ontwikkelingsstadia in de data van de zes kinderen toont aan dat de overlap pas in een later stadium ontstaat. Deze observatie plaatst de observaties van Hoofdstuk 4 in een ontwikkelingsperspectief.

Door de factor ontwikkeling als uitgangspunt te nemen voor de analyse van kindertaal, kunnen twee observaties van Hoofdstuk 3 verklaard worden: de voorkeur voor modale RIs en het ontbreken van RIs die verwijzen naar afgeronde handelingen. Wanneer aangenomen wordt dat de Elsewhere Conditie, waarvan betoogd is dat deze een rol speelt in het lexicon van een volwassen spreker, ook de selectie van woorden in het lexicon van het jonge kind organiseert, kan verklaard worden waarom Nederlandse RIs zo vaak modaal gebruikt worden. Infinitieven, vormen zonder tijds- en modaliteits-kenmerken, zijn de eerste werkwoordsvormen die Nederlandse kinderen leren. Wanneer vormen geleerd worden die semantisch gespecificeerd zijn, dan zorgt de Elsewhere Conditie ervoor dat de meer gespecificeerde vormen geselec-teerd worden en niet de ongespecificeerde infinitief. Op deze wijze worden in het taalleerproces RIs verdrongen door finiete zinnen. Doordat gespecificeerde

alternatieven voor RIs met een tegenwoordige tijdsbetekenis eerder productief zijn dan gespecificeerde alternatieven met een modale betekenis, worden de RIs van Nederlandse kinderen relatief vaker modaal gebruikt naarmate de kinderen ouder worden: RIs vertonen verschuiving naar modaal gebruik (*Modal Shift*). Het overwegend modale gebruik van RIs in de latere verwervingsstadia is groot genoeg om te leiden tot een dominantie van modale RIs wanneer alle data van alle stadia samengenomen worden. Een van de oorzaken voor de modale verschuiving is het overgebruik van perifrastische werkwoorden voor handelingen die in de tegenwoordige tijd plaatsvinden. De Elsewhere of ontwikkelingshypothese zoals voorgesteld in deze dissertatie heeft als voordeel (boven de andere verklaringen voor de modaliteit van RIs) dat ze verklaart waarom niet alle RIs modaal zijn en verklaart waarom RIs modaler worden naarmate kinderen ouder worden.

In Hoofdstuk 3 is verder gevonden dat RIs van Nederlandse kinderen niet verwijzen naar handelingen die afgerond zijn. Ook in dit geval is mogelijk de werking van de Elsewhere Conditie zichtbaar: aspectueel gespecificeerde alternatieven voor RIs worden vroeg geleerd in de vorm van zogenaamde root participles (RPs). Er is een tweede interpretatie van de observaties mogelijk en dat is dat infinitieven ongespecificeerd zijn voor Modaliteit en Tempus, maar niet voor aspectualiteit. RIs en RPs zijn tegenhangers: de eerste bevatten een vorm gespecificeerd als [-compleet], de tweede een vorm gespecificeerd als [+compleet]. Deze interpretatie betekent dat het aspectuele contrast [±compleet] gemarkeerd wordt voor het modale contrast [±realis] en het temporele contrast [±verleden tijd], respectievelijk. Het modale contrast tussen mogelijkheid en noodzakelijk lijkt ook een ontwikkeling te vertonen: de verwerving van perifrastische werkwoordsvormen breidt het scala aan expressies om modale handelingen te beschrijven uit. In de eerste stadia worden slechts noodzakelijke handelingen beschreven, later ook mogelijke handelingen.

Na een bespreking van overwegend semantische aspecten van de verwerving van finietheid, wordt ingegaan op een syntactisch effect van de verwerving van inflectie. Geïnflecteerde werkwoordsvormen vertonen congruentie met het zinssubject. Op grond van een recent ontwikkelde theorie wordt de hypothese getoetst of de verwerving van inflectie effect heeft op het gebruik van zinssubjecten door Nederlandse kinderen. Andersom gesteld: de patronen van het gebruik van zinssubjecten vormen een indicatie voor de verwerving van

inflectie. Het blijkt dat de ontwikkeling van subjectgebruik in de data van Abel, Daan, Josse, Laura, Matthijs en Peter compatibel is met de eerder getrokken conclusie over het moment waarop inflectie verworven wordt en de kinderen een onderscheid kunnen maken tussen geïnflecteerde finiete vormen en niet-geïnflecteerde infinitieven: vanaf het moment dat inflectie verworven wordt, beginnen RIs en finiete zinnen een tegenovergesteld patroon van subject-gebruik te vertonen.

HOOFDSTUK 6 SAMENVATTING EN IMPLICATIES De bevindingen in de hoofdstukken 3, 4 en 5 leiden tot een aantal hypotheses die getoetst kunnen worden in toekomstig onderzoek. Een eerste hypothese is dat het incrementele ontwikkelingstraject geschetst voor de functionele categorie Tempus (van afwezig, via lexicaal naar grammaticaal) kenmerkend is voor het ontwikkelings-traject van functionele categorieën in het algemeen. Een andere hypothese is dat lexicale encoderingen (via een hulpwerkwoord) eerder geleerd worden dan grammaticale encoderingen (via inflectie). Dit betekent dat de volgorde Aspectualiteit >> Modaliteit >> Tempus gevonden voor Nederlandse kindertaal niet universeel is, maar afhankelijk is van taalspecifieke en-coderingen. Voorbeelden van specifieke vragen die de voorgestelde analyses oproepen zijn: Gaan kinderen die de zogenaamde *pro*-drop talen leren, zoals het Italiaans, door een fase waarin ze subjecten in finiete zinnen overgebruiken? Hoe leren kinderen in een taal als het Afrikaans werkwoords-verplaatsing, in aanmerking genomen dat het Afrikaans geen enkel fonetisch-gerealiseerd inflectioneel suffix kent?

De uitgebreide aandacht die besteed is aan de methodes van experimenteel onderzoek van het gebruik van werkwoordsvormen bij zeer jonge kinderen, kan dienen als aanknopingspunt voor nieuw op te zetten experiment. Temeer daar de corpusmethode problematisch is met betrekking tot de verwerving van de betekenis van werkwoordsvormen, is verder experimenteel onderzoek (mogelijkerwijs met kinderen uit verschillende ontwikkelingsstadia) gewenst. Het is verder gebleken dat analyse waarin taalontwikkeling, meer specifiek de ontwikkeling van werkwoordsvormen, bestudeerd wordt, vruchtbaar is. De ontwikkelingsanalyse heeft verscheidene eerder gedane observaties in een nieuw licht geplaatst. Verder longitudinaal onderzoek is zeker gewenst, niet alleen in het Nederlands maar juist ook in talen die verschillen van het Nederlands om te onderzoeken in hoeverre kinderen die verschillende talen leren een vergelijkbaar pad doorlopen. Dit draagt bij aan een verdere

ontrafeling van de universele taalleermechanismen. Het is bekend dat Neder-
landstalige kinderen met een specifieke taalstoornis problemen hebben met het
gebruik van finiete werkwoorden. Ook kinderen en volwassen die het
Nederlands als tweede taal leren maken 'fouten' in het gebruik van finiete
werkwoorden. Gedetailleerde ontwikkelingsanalyses zoals uitgewerkt in deze
dissertatie kunnen bijdragen tot het begrip van de aard van de verschillen
tussen normaal ontwikkelende kinderen die Nederlands als hun eerste taal
leren, en (i) Nederlandstalige kinderen met een taalstoornis en (ii) kinderen en
volwassenen die Nederlands als tweede taal leren.

CURRICULUM VITAE

Elma Blom was born in Warnsveld on the 4th of December 1972. She attended the Baudartius College in Zutphen, where she received her Athenaeum diploma in 1990. She studied Dutch Language & Literature at Utrecht University and specialised in syntax-semantics and sociolinguistics. After obtaining a cum laude Master's degree in 1996, she started with psycholinguistic research at the Utrecht Institute of Linguistics OTS. The research carried out at the UIL-OTS resulted in this dissertation. Parts of the corpus study reported in this dissertation have been used for a joint Dutch-Flemish project on machine learning on the basis of naturalistic language input. The experimental research on early child English is carried out during an eight-month visit at the UCLA Linguistics Department. While preparing this dissertation, Elma Blom was also a student at the Royal Academy of Fine Arts in The Hague.