SPEECH ABOUT SPEECH

A developmental study on form and function
of direct and indirect speech

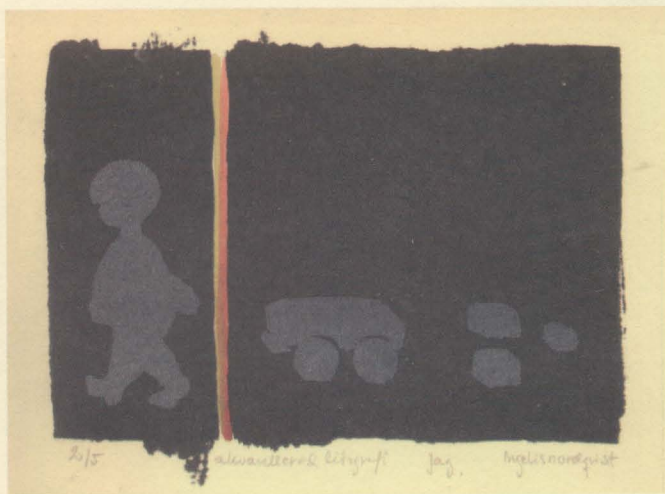ÅSA NORDQVIST

SPEECH ABOUT SPEECH

A developmental study on form and function of direct and indirect speech

ÅSA NORDQVIST

Doctoral Dissertation

Publicly defended in Lilla Hörsalen,
Humanisten, Göteborg University,
on March 3, 2001, at 10.00
for the degree of Doctor of Philosophy

SPEECH ABOUT SPEECH

A developmental study on form and function of direct and indirect speech

ABSTRACT

The present study explores the emergence and development of forms of direct and indirect speech, and functions connected to them, from a long-term perspective. More specifically, the research questions examined concern at what point in time children start to use direct and indirect speech, what the course of development looks like, and how direct and indirect speech are used by children and adults in different activities and in speech in comparison to writing. Four types of direct and indirect speech are distinguished according to their grammatical-structural and deictic properties: *indirect, free indirect, framed direct*, and *free direct speech*. These forms are argued to be used either to pass on information contained in earlier actual utterances (*speech reporting*) or to project speech onto fictional characters (*speech projection*). Different types of information can be packaged simultaneously into a form of direct or indirect speech (e.g., gestures, intonation, voice quality) to express several functions (e.g., to convey speaker perspective, express evaluation and plot advancement, and describe the referent).

The empirical data examined include longitudinal and cross-sectional data, and the use of direct and indirect speech by Swedish-speaking monolingual children who have just started to produce their first words, pre-school children telling a picture-elicited narrative and playing with a doll house, and school children and adults narrating a story in speech as well as in writing, are investigated. Quantitative as well as qualitative types of analysis are carried out, and the results show that direct speech appear before (around 26 months) indirect speech (around 35 months) in the children's production, and that the first occurrences are typically prompted by the parents. The emergence and development of use of the forms are found to be intimately connected with the development of understanding the minds and perspectives of others. The results of the 3-year-olds' language use, indicate a pragmatic awareness in that different types of forms are used in different types of activities, information is packaged simultaneously into the forms to express multifunctionality, and direct and indirect speech are integrated in narrative frames. The majority of the school-age children (i.e., 9-, 12- and 15-year-olds) includes few speech projections in their spoken and written narratives, and the narratives have a detached feel. Shifts of speaker perspective are more successfully employed in their spoken narrations than their written narratives, suggesting that they master *speaking within speaking* better than *speaking within writing*. The adult group of narrators is the only one clearly differentiating between the two modes of production, in that they include significantly more speech projections, and a greater number of descriptive speech act verbs and forms of free indirect speech in their written narratives than in their spoken ones.

KEY WORDS: direct and indirect speech, speech projections and speech reporting, long-term language development, linguistic information structuring, speech and writing, narratives

ACKNOWLEDGEMENTS

This dissertation has been slightly revised since the thesis defense. In addition to the people mentioned above, I would like to thank *Judy Reilly* who acted as my faculty discussant at the thesis defense and who gave me many new insights. I am also grateful to *Joakim Nivre* for encouraging me at an early stage to carry out analyses on free indirect speech, and I especially appreciate the valuable comments he provided on the previous edition of the thesis. Some of these comments have been taken into consideration in this final version. The remaining shortcomings are entirely my own.

The lithograph on the cover is entitled "jag" ('me') and was created by *Ingelis Nordqvist*. I am indebted to her for help with the layout of the cover.

Göteborg in June 2001

Åsa Nordqvist

<u>Errata</u>

Due to a formatting error, some page number references in the Table of Contents are misleading. From section *4.2 Three-year-olds* and onwards, the page numbers shown are too high by two. Hence, Section 4.2 starts at page 180 (not 182), Section 4.3 at page 201 (not 203), and so forth. This formatting error also affects all the cross-references made in the text that refer to any page between 180 and 298.

CONTENTS

SUMMARY OF THE THESIS

REFERENCES

APPENDIX

1   Introduction

Language can be used to refer to language and we can talk about talk. There are different ways of reporting speech that has been uttered at another time and/or in another place. One possibility is to use the form of direct speech, *Joanna said, "the sun is shining today"*, and another way is to use an indirect form, *Joanna said that the sun was shining that day*. A fundamental difference between these two forms lies in the speaker perspective or point of view of the reporter. In the case of direct speech the point of view is adapted to the original speech situation and speaker (Joanna), and the speech is claimed to convey the words exactly as they were originally uttered. In indirect speech the reporter relates the speech event from his or her own point of view, and the deictic elements are adjusted accordingly. However, these forms of direct and indirect speech may be used even when there is no actual speech event to report and to take a stance on. For example, direct and indirect speech may be used in cases of fiction, e.g., in the relating of a fictional narrative or when a child is playing make-believe. In these cases the characters and their speech, are, to a great extent made up and if there are earlier speech events that are referred to at all, these are only *seemingly* authentic speech events. Thus, the two forms may be used for a wide range of purposes and different conditions of use are related to them. For a child acquiring a langauge and learning to use forms of direct and indirect speech, the task is complex. Besides acquiring the mere syntactic structure and getting the syntactic aspects right (like word-order, deictic adjustments and co- or sub-ordination of clauses), the child needs a metalinguistic capacity to monitor and express her relation to her ongoing speech. Yet another aspect is how the forms are used for different purposes and to create certain effects in a variety of contexts. In order to report speech successfully the child also needs to make clear to the listener whose point of view is being taken. Due to the complexity of the task there is reason to believe that the ability to report speech develops over a long period of time.

1. Introduction

1.1 Aim and scope of the study

1.1.1 Aim

The purpose of the present study is to examine the emergence and development of forms of direct and indirect speech, and functions connected to them, from a long-term perspective. The empirical data to be investigated include Swedish-speaking monolingual children who have just started to produce their first words, pre-school children telling a narrative and playing with a doll house, and school children and adults narrating a story in speech as well as in writing (see below).

There are several motivations for this type of study. As was indicated above, the use of direct and indirect speech to refer to speech from another speech event is actually *speech-about-speech* and *speech-within-speech*. The reporting of speech is thus, in a sense a metalinguistic act and studying children's use of these forms provides us with information about the development of one aspect of metalinguistic awareness. The data is also comprised of written narratives, so the reporting of speech and dialogues in written discourse adds a further metalinguistic dimension in that the system of writing is a symbolic representation of spoken language. In this case the development of metalinguistic skills concerns *writing-about-speech* and *speech-within-writing*. Moreover, a reporting speaker needs to estimate the listener's knowledge of the world and provide her with enough information so that it is clear that speech is reported, whose speech is reported, and what the speech or message consists of. In the case of fictional narratives and use of direct and indirect speech, it is of particular importance to make clear shifts in perspective between the narrator and story characters. Thus, pragmatic considerations are required in order to succeed in reporting the speech. Investigating children reporting speech at different ages gives us an idea about the development of pragmatic competence, narrative competence, and the children's understanding of other people's minds (perspective taking).

In the introduction of this chapter, one example of direct speech and indirect speech respectively, were provided. The case of direct speech, *Joanna said, "the sun is shining today"*, linguistically consists of two separate main clauses, where the first clause contains a speaker identity (*Joanna*) and a verb of saying (*say*). In this clause, (that may either precede or follow the clause with the quote), the perspective is that of the reporter and the verb is thus inflected for past tense. In the second clause - the quote - the perspective is that of the reported speaker, and the verb is in the present tense and the adverb of time is *today*. The example of indirect speech, *Joanna said that the sun was shining that day*, in contrast, consists of a main clause and a subordinated clause and these clauses cannot typically be switched. The perspective is that of the reporting speaker and the deictic elements are adjusted accordingly (*said, was, that day*). Hence, in addition to pragmatic competence, use of these forms requires certain linguistic competence (e.g. to co- and sub-ordinate clauses). This is also an aspect worth examining in greater detail from a developmental perspective.

Direct speech has been shown to be produced by English-speaking children as early as at two years of age and indirect speech slightly later (Ely & McCabe 1993). However, although it is the case that the forms may be acquired by very young children, I will argue that *how* these forms are used – what functions they can have, how the forms may be employed in different activities and for different purposes, and how the linguistic information may be structured – is a process of long-term development. Therefore, it is not enough to establish when the first forms emerge, but it is also necessary to examine how the use of the forms develops over a longer period of time and within different types of contexts.

Although the issue of direct and indirect speech has been extensively examined in various research areas such as linguistics, philosophy, literary science and psychology, (see Nordqvist 2000a), relatively few studies have been concerned with its use by children. Hickmann contends that "[a]lthough an increasing number of studies have focused on children's discourse cohesion

..., surprisingly few have examined children's narrative strategies when reporting speech" (1993: 67-68). Further, Goodell and Sachs claim that "[v]ery little is known about English-speaking children's acquisition of the skills needed to report the speech of others" (1992: 395). To my knowledge, analyses of *Swedish* children's acquisition of these skills is even more scarce. In a later section (Section 2.4.4) several gaps in research will be pinpointed, and this thesis aims to fill in some of these gaps.

From this overview it is clear that there are several theoretical motives for carrying out this type of study. The results are intended to lead to greater knowledge about the development of linguistic, metalinguistic, pragmatic and narrative competence in general, and the development of forms and functions of direct and indirect speech in speaking and writing in Swedish in particular. An important additional motive for the present investigation is the need to study language development from a long-term perspective. Most studies dealing with language development have concentrated on a limited age range. In addition, when children's use of direct and indirect speech is studied, most often one type of use is focused on (e.g. experimentally elicited forms in a particular age group). In contrast, the data to be presented briefly in the section below cover a wide age range and include several types of data. Accordingly, one main aim of the thesis is also to create a theoretical and methodological framework that can handle these broad and heterogeneous data.

1.1.2 Types of data

The combined corpora underlying the present study have a size of close to 275,000 transcribed words, and include a total of 144 Swedish speaking subjects. The corpora were not collected for the purpose of analyzing direct and indirect speech, but they were found suitable for the present purpose of investigation. Moreover, the data are production data rather than data on language comprehension. Three sub-corpora can be distinguished (a detailed description of all data is presented in Section 3.2). The first one is a

longitudinal case study corpus comprised of audio and video recordings of a boy and a girl between 19 months and 4 years of age. Twenty-eight recordings per child were made and done so in the children's home-settings when the children were engaged in typical everyday activities with their caretakers (play, reading, eating routines). These recordings were then transcribed in computer format in order to facilitate analysis. The second corpus consists of fourteen dyads of 3-year-olds and their mothers playing make-believe with a doll house for about half an hour. These sessions were audio and video recorded and transcribed by hand. The third and final corpus is comprised of picture-elicited narratives by 3-, 4-, 5-, 9-, 12-, 15-year-olds and adults. The older subjects (the 9-, 12-, 15-year-olds and the adults) produced one spoken and one written version of the story. This design makes it possible not only to investigate the use of direct and indirect speech within narratives from a developmental perspective, but, in addition, to compare their use in speech and writing.

1.1.3 Research questions

The data selected allow for different types of analysis and aim to answer several research questions. A first question to explore is what the development of forms looks like.

Development of forms
When do the first forms of direct and indirect speech emerge, and what types are the first to appear? What is the frequency of the forms at different ages, and in different activities? What does the development of forms look like in the data?

This type of analysis is valuable and possible to carry out in all the types of data included in the present investigation.

The next question concerns how the forms are employed. This is of interest from a developmental perspective since one of the hypotheses of this thesis is that the forms may appear early but that the functional use changes during the course of development. One aspect concerns linguistic information

structuring. When quoting Joanna we may aim to depict her voice pitch along with the words she uttered. In this way information about the speaker and her speech is provided simultaneously by prosodic and verbal means. To what extent do children make use of options like these? Moreover, it is of particular interest to find out how the type of activity carried out influences the use of direct and indirect speech. Can we expect that the forms are used differently by the 3-year-olds playing with a doll house than the 3-year-olds telling a picture-elicited narrative?

The use of the forms
How are the direct and indirect speech forms used at different ages and in different types of activities? What does the linguistic information structuring of the forms look like? Do the subjects manage to convey shifts in speaker perspective?

The child acquires her language in interaction with her environment, and the parents are a crucial part of this environment. As will be shown later, adults adapt their speech to a certain extent when talking with a child, and this plays a major role in engaging the child in conversation. In the Longitudinal case study material, the data with the three-year-olds playing with a doll house and the three- and four-year-olds narrating the picture story, we can explore the language input provided by the adults. These data provide a window on the interplay between input and production in an early phase of acquisition and on what types of usage the adults encourage in the children. They do not, however, allow us to make inferences about input in later stages of development.

Input characteristics
What types of direct and indirect speech forms do the adults interacting with the children use, and how do they use the forms? Do they prompt their children to use direct and indirect speech?

The fourth and final research question concerns later language development and the patterns of use of these forms by school children,

adolescents and adults. These subjects all have more training in language use than the younger children and they have the experience of language instruction in school. More importantly, these children have learned how to read and write. As briefly discussed above, the use of direct and indirect speech in writing places even greater demands of metalinguistic and pragmatic competence on the reporter than their use in speech.

Later development in speech and writing
In what ways does school children's use of direct and indirect speech differ from younger children's use? How does the development of writing look in comparison with speaking? What about adults' use of direct and indirect speech? How do the forms develop in a long-term perspective?

The research questions are further elaborated in Section 2.5, and they are answered and discussed in Chapter 5.

1.2   Outline of the thesis

This thesis is divided into six chapters. The present chapter introduces the thesis by presenting the topic, briefly describes the data, and states the research questions. The next chapter, *Theory*, lays the theoretical framework of the thesis and presents previous research on language development, how linguistic information is structured and conveyed, what "narrative" refers to and some general aspects of narrative development. The chapter also includes a detailed and critical examination of previous research on "direct and indirect speech", and develops a theoretical model which the subsequent empirical analyses rest on. The chapter concludes with a reconsideration of the research questions stated in the introduction (above).

In Chapter 3, *Methodology*, I present the data design and describe the types of analysis that are carried out (including measurements, operationalizations, etc.). The outcomes of the analyses are presented and discussed in the subsequent chapter, *Results*. In Chapter 5, *General discussion*,

1. Introduction

I summarize and relate these findings to the research questions, and discuss some questions in regard to the methods used. In the final chapter (*Conclusion*), I formulate some broad assumptions deduced from the findings and make suggestions for further research.

2 Theory

This chapter presents the theoretical framework of the thesis and includes examinations of relevant areas of linguistic research. Since the main issue of this thesis is to investigate how direct and indirect forms of speech develop from a long-term perspective, we need to clarify what language development consists of and what the conditions are for acquisition of language. This is dealt with in Section 2.1. The following section, 2.2, is concerned with the structuring of linguistic information, what is meant by form and function, and what the conditions of use of spoken and written language consist of. Utterances of direct and indirect speech are typically embedded in narrative structures, and previous research related to narrative structure and types, narrative development etc., is then presented in Section 2.3. Section 2.4 is devoted to a detailed presentation and critical examination of direct and indirect speech, which results in a theoretical model (Section 2.4.2) that will serve as a base for the analyses of the data. The chapter concludes with Section 2.5 where the research questions stated in the introduction will be discussed in more detail and related to previous research.

2.1 Language development

This section consists of three sub-sections. The first section discusses language development as an open-ended and non-linear process; the second section emphasizes the fact that the language-acquiring child is a part of a social context; and the third and final section reviews research on the child's cognitive development and theorizing of mind.

2.1.1 An open-ended and non-linear process

When people interact by means of language, cognitive, communicative, and linguistic factors interact. Cognition is necessary for the planning and monitoring of speech, and for the perception and interpretation of speech.

Linguistic communication implies a sender and a receiver, and a successful interchange involves, among other things, a willingness to cooperate, rationality, and adaptation to the situational context (Grice 1975, Allwood 1976). As a sender you have to take into account the interlocutor's state of mind, and assess how much of the array of information to be talked about is already shared by the interlocutor and what the interlocutor needs to know to understand new information you might wish to convey. Furthermore, a set of conventional linguistic forms are used in this interchange. Berman & Slobin (1994) contend that these three main aspects also interact in a complex fashion in development. They characterize development in the following way:

> Younger children take fewer expressive options because: (a) cognitively, they cannot conceive of the full range of encodable perspectives; (b) communicatively, they cannot fully assess the listener's viewpoint; and (c) linguistically, they do not command the full range of formal devices. (Berman & Slobin 1994: 15)

Hence, development consists of, among other things, becoming conscious of perspectives that are encodable, adjusting to the listener (and her perspective), and finding appropriate and effective linguistic forms to convey intents and information. However, development should not necessarily be seen as a constant and linear growth of, for instance, lexical forms. Slobin's generalization that "new forms first express old functions, and new functions are first expressed by old forms" (1973: 184) implies that, although the language use may start out for the little child as a one-to-one mapping between form and function, development is more complex than acquisition of new forms with a built in constant function. "Across time, use of any given form is extended and hence reconstructed in a variety of interrelated ways --- while knowledge of linguistic forms is evident from a very young age, this knowledge is only partial since, with age, these forms take on different functions" (Berman 1996: 345-346). Strömqvist (1998) presents a similar view on language development:

... important parts of language development can be modeled in terms of reorganizations of the mental representations in the learner or, more precisely, reorganizations of the relation between linguistic forms and functions/content. These reorganizational processes are driven both by principles of human information processing (such as principles of clarity, economy, etc.) and by external factors (the input and the learner's adaptation to language usage in new sociocultural contexts), and they characterize the language user not only in childhood but also in adolescence and adulthood. (Strömqvist 1998: 99)

In other words, development involves reorganization of processes, and as Karmiloff-Smith states, "it is not obvious that the same outcome in behaviour is the result of the same underlying processes" (1981: 123).

These observations indicate that language development is more complex than only a linear growth and acquisition of new forms. Rather, langauge development can be said to consist of reorganizations and the acquisition of more functions for the same form. "Most generally, it will be the case that individual forms take on more functions with development. That is, as new functions emerge, they will recruit existing forms, and also stimulate the acquisition of relevant new forms" (Berman & Slobin 1994: 33). Thus, although the forms of direct and indirect speech may be acquired early by a child, the acquisition process of the possible functions tied to these forms is a long-term matter.

Language development is an open-ended process. Mental reorganizations continue to take place throughout life, and although some aspects of language, such as pronunciation skills or basic grammar are, to a large extent, acquired already at a relatively early stage of development, other aspects, like lexical knowledge and rhetorical abilities continue to grow; "there is neither any perfectly steady state nor any state of completeness in the linguistic career of a normal human being" (Strömqvist 1998: 98).

2.1.2 Importance of input and interaction

Certainly, the child does not grow up in a social vacuum and the language acquisition process cannot be explained solely in terms of mental processes

within the individual child. Rather, the child acquires her language in interaction with her environment. In studies of mainly Western middle-class families, it has been shown that child-directed speech (CDS) has certain systematic characteristics. It is generally higher in pitch, more exaggerated in intonation and slower in tempo than speech between adults (Fernald *et al* 1989; van der Weijer 1999). The utterances tend to be shorter, well-formed, and contain fewer complex sentences and subordinate clauses. Furthermore, false starts and hesitations are rare, whereas repetitions, expansions, and recasts of the child utterances frequently occur. In addition, the topics tend to be closely tied to the immediate context in order to engage the young child in conversation. Child utterances are embedded in larger and more complex informational and interactional structures by the parents, who also adapt to the child's perspective. In this way, the caregivers ensure joint attentional focus. Although the expanded version of the child's utterance also gives the child important syntactic and semantic information, this parental style of interaction has the major function of engaging the child in conversation. Pine (1994) states that there is now a general consensus that speech adjustments to young language learners are motivated by a desire to communicate rather than to teach language.

Describing the characteristics of this particular speech register (CDS) dominated the research field of input in the beginning of the 1970's (see the contributions in Snow & Ferguson 1977), but interest has since then shifted to the functions of these speech modifications, and to the kinds of procedure which children apply to the analysis of their input. Importantly, and as Lieven (1994) critically points out, most of the studies of input to children are based on children growing up in middle-class Western families living in urban areas in mainly English-speaking countries. There are cultures where (prelinguistic) infants are not spoken to, and no particular speech adjustments are made.[1] Yet, all normally developing children learn to speak. Hence, the functions of speech

---

[1] See Lieven (1994) for an overview.

modifications, and the relationship between the linguistic input and the child's production, are not straightforward.

Direct relationships between certain features of CDS and the acquisition of certain linguistic aspects in English certainly exist. Several training studies have demonstrated that enriched input facilitates the development of precisely those aspects of the langauge system to which the training is directed (e.g., Nelson *et al* 1973; Shatz, Hoff-Ginsberg & MacIver 1989). Density of maternal speech (or frequency of exposure to words) has been shown to have implications for vocabulary acquisition (Huttenlocher *et al* 1991), and Farrar (1990) suggests that recasts of the child's utterance can play a facilitating role in the acquisition of grammatical morphemes. However, Pine (1994) contends that the relationship between the linguistic input and the child's grammatical development is a very complex question, and research results have sometimes turned out to be quite contradictory. This can be explained by several problematic aspects. Firstly, as Pine (1994) points out, there are methodological shortcomings. The size of the corpora have often been small, and the datapoints few. Moreover, different statistical methods, variables and theoretical departures have been applied (for a discussion of this problematic question, see Richards (1994)). Secondly, correlations might not be straightforward or easy to interpret. For instance, Sokolov and Snow point out that "several maternal mechanisms may have the same effect, so absence or low frequency of these mechanisms may not reveal itself in outcome data" (1994: 42). Finally, and perhaps most importantly, children make use of the language they hear in different ways at different points in development. An implicit assumption has often been that the relationship between CDS and language acquisition should be the same regardless of the child's current level of linguistic ability and cognitive maturity. Yet on the contrary, it is likely that the child has different cognitive thresholds to achieve at different stages in development and is only apt to pick up the kind of linguistic information that is currently processable by her (cf. Berman & Slobin's characterization of 'development' in Section 2.1.1 above). Hence, a certain feature of, or

13

structure in the CDS may have great impact at one stage in development, but less relevance at another.

Despite the fact that the structure of the input to, and the attitudes towards children are different in different cultures, most children learn to talk and solve the problem of segmenting speech and relating content to form. Lieven (1994) reports that in many cultures where speech is not adjusted when speaking to children (cf. CDS), children listen to and overhear adults' speech to each other, imitate and echo what they hear, and are also asked by the adults to imitate. Due to a sensitivity of the frequency and routinization of utterances and events, and a quite remarkable memory of the relationship between an utterance not fully understood and the pragmatics of the situation in which it was initially uttered, the child registers distributions and meanings within the language. This indicates that children are bringing a fairly impressive range of interactive, processing, and generative skills to the task of learning to talk. However, children will only learn to talk in an environment of which they can make some sense and which has a structure of which the child is a part. Routines thus seem to facilitate the child's langauge acquisition process. Snow *et al* (1987) and Sokolov & Snow (1994) argue that there are two main ways in which parents can build instructive discourse frames for their children. One is to be highly responsive to the child's actions, gestures and vocalizations, and another is to provide the child with predictable texts and routines so that the child comes to recognize what the structures of those are. Children who have stories read to them on a regular basis are likely to become familiar with the genre, and for instance, how dialogues are reported. The caretaker thus provides the child with a valuable activity and culturally dependent discourse models. Indeed, one important function of CDS is to socialize the child into beliefs, feelings, and behaviors that are appropriate to her role in her own culture (Ely & Gleason 1995). This includes knowledge about activity-related language. Moreover, as will be discussed later in this thesis, the quality of early interaction has important consequences for narrative development.

In conclusion, on the one hand, we have a flow of speech directed to the child, and on the other, we have the child who is processing this flow and formulating hypotheses about the structure of the input langauge. This section has illustrated that CDS has certain (but only to some extent universal) features, but not yet enough is known about the particular effects of CDS, and about the language acquiring child as a processor and hypothesis maker. Nevertheless, there is reason to believe that the input, and the models of language that environment provides, are important to the langauge acquiring and developing child, and that the caretaker's willingness to interact with the child contributes to the child's linguistic, communicative, and pragmatic development.

2.1.3 Understanding of minds

In order to communicate effectively, the young child must not only adapt to physical constraints in the communication situation, but also discover that other people have minds. Piaget (e.g., 1926) is of the opinion that the pre-school child is cognitively egocentric in that she does not understand that other people's views and needs are different from her own. Piaget emphasizes the lack of social adaptation characteristic of egocentric speech, and he regards the process of decentration – from egocentrism to an ability to incorporate other people's perspectives in her own thinking – as completed by about age 6 or 7. The assumption of such late achievement of social adaptation in the child has been criticized, since detailed studies of the nature and quality of early interaction between infants and their caretakers have shown that the ability to take the perspective of others develops already from an early age. Vygotsky (1962) argues that cognitive development is determined by language (i.e., the linguistic tools of thought) and by the sociocultural experience of the child, and Mead (1934; see also Cooley 1902) states that the realization of self develops within the child through interaction processes with others. In the symbolic interactionism theory (Mead), three levels of the development of self are

15

distinguished. The first step is represented by the *I* conception, where the child is aware of her own existence and needs. In the second step, the child reflects herself in her interaction partners, and by applying their perspectives to her own, a concept of a social *me* emerges. In the final stage, the child is able to think in more abstract terms and to communicate with a *generalized other*.

Beal (1988) points out that the "awareness that states of knowledge may differ between individuals is critical to the communication process, because a message will be effective only if it provides the information that the other person will need in order to understand" (1988: 315). As was discussed in the previous section, the young child is situated in a context with nurturing and caring interlocutors from whom the child receives considerable conversational support (Ninio & Snow 1996). However, as the child grows older, more and more topics deal with other aspects than the I-here-and-now, and other demands are made upon the child; the child's conception of the speech situation becomes more central. Included in this development is emergent theorizing of mind, and in recent years *Theory of mind* has attracted quite a bit of attention within developmental psychology research (for an early account of theory of mind in these contexts, see Wimmer & Perner 1983).

Understanding intentionality, and that other people also have intentions, is important for the child's pragmatic development. Poulin-Dubois & Shultz (1988) contend that the ability of attributing intentions to others seems to begin at the age of 2;6 to 3 years of age, as an extension of the child's conception of agency. The ability to *speak* about other people's intentions develops during the third year of life. Three-year-olds have an explicit understanding of the distinction between real objects and events and decoupled representations such as dreams, thoughts, and images (Wellman 1988). Four-year-olds make further progress, in that they become aware that real objects are also mentally represented (Forguson & Gopnik 1988). At this age, children understand false belief, representational change, and the distinction between appearance and reality. Flavell (1988) distinguishes two "Levels", in which two- to three-year-olds at Level 1 understand that another person may or may not see something,

whereas at Level 2 (4-5 years of age) children understand that something seen may present different appearances or engender different visual experiences if the observer views it from different positions in space.

Leslie (1987, 1988) shows that the fundamental forms of make-believe (i.e., *object substitution* (like "Mummy is daddy" or "the banana is a telephone"), *pretend attribution of properties* (an object, event, or situation is imagined to have properties it does not have), and *imaginary object* (an object is said to exist where there is none)), emerge in the child between about 18 and 24 months. Leslie states that the mastery of pretending is an important step in understanding mental states; more exactly: "early pretend play is actually a primitive manifestation of the child's theory of mind" (1988: 24). Kavanaugh et al (1997) report that by age 2, children's pretend play often involves both familiar play partners and replica toys, and the child may, for instance, be found "feeding" the doll. The next step involves ascription of passive agency to replica objects. Thus, a 2-year-old may pretend that a doll placed on a bed is sleeping. After mastering this, children start pretending that replica toys can carry out their own make-believe actions, for instance, a toy doll may put another toy doll to bed. As the children grow older, their make-believe play becomes increasingly complex, and by age 4, they ascribe make-believe emotions and cognitions to inanimate objects. In addition to this, Montgomery & Montgomery (1999) show that already at the age of 3, children will accurately attribute intentions to moving dots on a computer screen, suggesting that children have little difficulty attributing mental states to inanimate objects.

Thus, children engage in make-believe play from an early age and use objects that are quickly ascribed agency and intentions. Although little points to the fact that children assign intentions to dolls before they attribute them to other people, there is reason to believe, (as Leslie does above), that playing make-believe with replica toys is an important step towards a more complex understanding that other people have independent minds and other perspectives.

2. Theory

The pre-school child is fairly well equipped with pragmatic skills and qualifications for being an effective communicator and interlocutor. However, although these young children learn relatively easily how to communicate their intended meanings, their understanding of the requirements for effective communication continues to develop into the early elementary school years. For example, Beal (1988) discusses the importance of the communicative *quality* of messages, and supported by research results, establishes that an understanding of the role of message quality in determining communication success or failure is gradually acquired in the early school years.

Entering school has important consequences for real conscious *metapragmatic* behavior according to Gombert (1990), and the development then proceeds until adolescence:

> It is again at the age of approximately 6-7 that we see the emergence of behaviour which ambiguously reveals reflection on or monitoring of language in terms of the relationship between it and its context of emission. This applies equally to both the awareness of referential ambiguities and the young speaker's consideration of the addressee's characteristics. However, it seems that this ability to process contextual indices continues to grow until adolescence as such indices become more complex and, probably, more familiar to the child. (Gombert 1990: 120)

Gombert argues that the increasingly sophisticated linguistic organizations sculpted by the adolescent depend to a great extent on the growth in the processing capacity of working memory. Perner (1988) points out that nonliteral use of language which does require complex analysis of communicative intention (for instance, irony and sarcasm) develops rather late. This is due to, according to Perner, the fact that younger children "do not have to develop their own communicative conventions but are socialized into an existing system" and that "there is no need for them to engage in sophisticated intentional analysis from the beginning" (1988: 289). As will be discussed further below (see Section 2.2.3), learning to read and write in school also contributes to increased metalinguistic and pragmatic awareness.

2.2 Linguistic information structuring

This section is introduced by a discussion about what constitutes a *form*, and what *function* consists of. In the subsequent section a range of functional dimensions are presented, and functions of specific relevance to this particular study will be distinguished. The final section deals with functional similarities and differences between two modes of language production – speaking versus writing.

2.2.1 Form vs function

As was hinted at in the section on language development, form and function interact in a complex fashion in language use and development. A relevant question is then what 'form' consists of, and what is meant by 'function'.

Jespersen (1924), taking a grammatical stand, distinguished the difference between *form*, *function*, and *notion*. Form was described as the concrete expression, or a sound (of a word or some other part of a linguistic expression), whereas function referred to the grammatical meaning, or rather, a syntactic class. The notion is then the content side.[1] In their book on narrative development, Berman & Slobin (1994: 4), like Jespersen, regard *forms* as linguistic devices that include grammatical morphemes, bound inflections, interclausal connectives and syntactic constructions, along with lexical items encoding notions of temporality, manner and causation. This is to say that form is, by and large, equivalent to *grammatical structures* of different kinds. These "grammatical and lexical forms of a language provide the speakers with an array of different 'expressive options' or 'rhetorical choices' for verbalizing particular concepts and relations in ongoing discourse" (Berman 1996: 344). This is also the sense in which I will hereafter regard 'form', that is, as a unit

---

[1] The function connects the expression with the content, and as a speaker (or writer) you start out having a notion and move through syntax (function) to the formal expression. In the case of listening (or reading), the process goes in the opposite direction. "Syntactic categories thus, Janus-like, face both ways, towards form, and towards notion. They stand midway and form the connecting link between the world of sounds and the world of ideas" (1924: 56-57).

19

with certain grammatical-structural properties. The forms of direct and indirect speech will be further defined in Section 2.4.2.1.

The question of what constitutes *function* is more tricky, and it has indeed been applied to different aspects in different studies (cf., for instance, Jespersen above relating function to syntactic class). In this thesis, the function of a linguistic form that is expressed at a certain point in time, refers to *how the form is used*. More specifically, what *context* it is used in and what *effects* the usage brings about, are focused on.

Berman & Slobin (1994) choose to regard *discourse function* as "purposes served by these forms [i.e., linguistic devices that include grammatical morphemes, bound inflections, interclausal connectives and syntactic constructions, along with lexical items encoding notions of temporality, manner and causation] in narrative discourse - purposes of constructing a text that is cohesive and coherent at all levels: within the clause, between adjacent clauses, and hierarchically relating larger text segments to one another" (1994: 4). To illustrate, the form of the English progressive aspect can be used for backgrounding information (example from Berman & Slobin 1994: 5):

And then he fell over with the dog - into the pond. He **was** just **sitting** - on the - edge before - with his dog, and pow - into the water.

Berman & Slobin thus regard 'purpose' as the defining characteristic of function. This needs to be critically considered. Most importantly, it is not clear what is meant by purpose. If it is meant to refer to intentional behavior, whose intentions are we to consider? It is likely that it is the narrator that has a certain purpose by using a certain option, but, in that case, to what extent is the receiver and the investigator able to decide the level of intentionality? My interpretation is that in the above definition, it is not only the purpose of the narrator that determines the function, but also the result achieved through the behavior. More specifically, it is actually *the effect* of the behavior (i.e., the perlocution (Austin 1962)), that the authors are including in their definition.

Indeed, as Allwood (1980) points out, investigating the achieved effect of a certain behavior is a way to approximate a speaker's original intention and this is a common and fruitful research method. The notion of purpose should nevertheless be carefully used in relation to a definition of function, since intentionality is not available to an outside observer.

Thus, what is possible to say anything about – for me as an observer and investigator – is what effects a certain use of direct or indirect speech brings about and in what context it is used. Later I will distinguish between two major functions in relation to use of direct and indirect speech. The first type, *speech reporting*, refers primarily to cases where information contained in earlier actual utterances is passed on. The second type, *speech projection*, is used in cases of fiction, typically when speech is ascribed to a doll or to a story protagonist. In order to grasp what constitutes these functions, and what distinguishes them, the child needs to experience contexts in which the forms (i.e., direct and indirect speech) occur, how they are conventionally used, and what (intentional) use of the forms can have as a result. As was discussed in Section 2.1.1, this is indeed a long-term form of development.

So, besides the major functions of speech reporting and speech projection, what other functions may be expressed through the use of direct and indirect speech? Examples of such functional dimensions will be discussed in further detail in the next section.

2.2.2 Filtering, packaging, and perspectivizing

When experiences are linguistically encoded, they are also *filtered* (Berman & Slobin 1994, Strömqvist 1996). That is, some aspects of the information array talked about filter through and others are filtered out in the linguistic encoding process. This means that the speaker chooses a perspective - e.g., what is topic and focus and what is to be backgrounded and foregrounded - and this has to be done through the set of options/forms provided by the langauge at hand. Certain verbs, word-order, and/or topic markers available in the language may

be employed to express certain perspectives, and language specific constraints then have an effect on how information is encoded. [2] Furthermore, linguistic information can be *packaged* in different ways. In Berman & Slobin's (1994) original use of this term, *packaging* refers to how narrative events are packaged into hierarchical (syntactic) constructions. However, in a more general sense (and as is used by Strömqvist 1996; Strömqvist, Ahlsén & Wengelin 1998), information cannot only be packaged in discourse by means of lexicon, morphology and syntax. Information can also be distributed over larger fragments of discourse, and in speech, phonetic and phonological aspects can convey additional information to what is strictly expressed by verbal means.[3] In the same simultaneous information structuring manner, one modality can be used multifunctionally; for instance, expressing a proposition with the help of words and at the same time by voice quality expressing a certain perspective is an example of this. Following Strömqvist, structuring an array of information from a certain point of view, can be said to have the function of *perspectivizing*.

Temporality (the expression of the location of events on the time line and temporal relations between events) is certainly a function highly relevant to the study of narratives, as is *connectivity* (that is to make a narrative coherent, by means of, e.g., syntactic conjunction, subordination and anaphoric devices). These functions have been examined carefully from a developmental point of view by several researchers and results from that type of studies will be presented in Section 2.3.2. However, in the context of direct and indirect speech, there are certain other functions that I wish to stress. Besides expressing a proposition and perspectivizing (mentioned above), direct and

---

[2] Strömqvist compares the English construction *land on the moon*, and the Russian equivalent *prilunit'sja* (1996: 4). In the English example, the word *land*, having no internal morphological structure, encodes the activity of landing, and the geographical goal has to be encoded by means of an entire prepositional phrase (*on the moon*). The Russian example (*prilunit'sja*), in contrast, consists wholly of a verb with an internal morphological structure which incorporates the goal. Thus, in the former case the information is distributed over a syntactic phrase and in the latter case a morphologically complex verb.

indirect speech forms can, for instance, be used to *vivify* a narration,[4] *creating a distance* between the narrator and the reported/projected speakers, and *advancing the plot* in a narrative. Typically these dimensions interact in a multifunctional fashion (see further Section 2.4.3.4 below). An interesting question to then investigate, is how information is packaged in these particular forms (i.e., direct and indirect speech) to construct perspective and express acts of speech, in different situations and at different ages.

In a later section of this thesis, Clark & Gerrig's (1990) analysis of quotations as demonstrations will be discussed. They point out that when Alice quotes George, she may depict the sentence he uttered, but at the same time she may choose to depict his emotional state, his accent and his voice. This is a further example of multifunctionality, and also of packaging, in this case within the spoken modality. In the section below, packaging will be discussed from a cross-modal perspective (speech versus writing).

2.2.3   Speech and writing

The prevailing view in the 20th century has been (and still is) to see spoken langauge as primary in relation to written language. Bloomfield's famous dictum that writing "is not langauge, but merely a way of recording language by means of visible marks" (1933: 21), implies that written language is only an (imperfect) rendition of speech. Speaking has been with us from early on in the history of mankind, whereas writing has existed only for a briefer period. It is also the case that speaking is acquired naturally and early on in every normal child, whereas writing and reading are usually acquired through deliberate instruction (which in turn depends on already acquired speaking abilities). However, although it can be established that writing has evolved from and builds upon speech, there are systematic differences between the two modes of

---

[3] Berman & Slobin (1994) also, however, mention a function of packaging where rhythm and tempo along with syntactic constructions allow for, e.g., the creation of suspense and surprise. They contend that this category is an elusive one.

expression and production. In contrasting the typical speech situation of face-to-face interaction, with the writing situation (where the act of writing typically occurs at a different point in time and place than the act of receiving and reading the same piece of text) certain modality-specific properties will become evident.

In the spoken face-to-face situation communication is multimodal although the focus is on speech. The whole body (including gestures, facial expressions, voice, etc.) can be used and expressive information can be packaged in a simultaneous manner (i.e., along the speech proper, it is possible to add paralinguistic features). The duration of the speech signal is very short, the speech is perceived and produced on-line, and the communication is highly interactive and mutually adaptive. Since the speed of speech processing is high and there is typically no room for extensive planning in advance, several hesitation sounds, repetitions and self corrections are made. Words like *eh* and *uhm* can also be used in order to hold the turn and/or to maintain the listener's attention and the speaker might choose to avoid long pauses and fill them out (Sacks *et al* 1974,; Allwood *et al* 1990, 1992). In the typical writing situation (as when producing the text in this thesis), the writer has more time to spend on planning and editing, and the receiver will only read the final edited version. There is no on-line feedback or mutual adaptation possible, and the writer does not risk her turn while writing. Moreover, written communication is typically monomodal, in that it relies on the visual modality only. Written language is linear to a great extent; expressive features are not distributed simultaneously in the same manner as in speaking.[5] (For a fuller account on these matters, see Strömqvist, Ahlsén & Wengelin 1998.)

---

[4] Vivifying can also be seen as a special case of perspectivizing, in that it helps foreground the vivified object.

[5] The extensive use of computer communication today, however, has contributed to the development of a written form of language that tries to capture a part of the multimodal dimension of spoken langauge. In order to convey information on emotional states along the written text, so called *smileys* can be used. A "happy face" :-) placed at the end of a sentence, for instance, is in some sense meant to resemble a real speech event where the speech is delivered at the same time as the speaker is smiling. However, despite its similarities with the real multimodal speech event, it is still linearily distributed in the text.

The differences between speech and writing in communication conditions and processing constraints also result in different forms of language. It has already been mentioned that normal speech contains several hesitation sounds and the like (a natural result of the fact that speech is produced on-line), while these markers are rare or non-existent in writing. The fact that spoken communication is interactive also makes feedback-signals and elicitors frequent in speech but not in writing (since it is typically non-interactive). Chafe (1982, 1985) argues for a notion of 'involvement' as characteristic of speech. Chafe regards markers of involvement to be, e.g., first person references, markers of a speaker's mental processes and monitoring of information flow (like hesitation sounds). According to a frequency list based on a Swedish spoken language corpus containing 1,2 million words, *jag* ('I') and *vi* ('we') are among the ten most common words and the speech management (Allwood *et al* 1990) marker *eh* ('uhm') also ranks high in the list (Allwood 1999). In the writing context, in contrast, the sender is *detached* from the addressee, and this leaves its mark on the written form, according to Chafe. Among other constructions, passives were shown to be more extensively used in writing (Chafe 1982).

Chafe (1982: 37) comments upon the fact that the speed of processing and production can be relaxed in writing: "As we write down one idea, our thoughts have plenty of time to move ahead to others. The result is that we have time to integrate a succession of ideas into a single linguistic whole in a way that is not available in speaking". He introduces the term 'integration', (i.e. "packing of more information into an idea unit than the rapid pace of spoken langauge would normally allow" (1982: 39)) as characteristic to writing. Comparing spoken langauge with written, Chafe found that integrated constructions like nominalizations, participles, and embedded and subordinated clauses were more common in writing. The notion of integration has clear affinities to that of packaging in that different types of information may be integrated, or packaged, into, for instance a participle or a direct speech construction.

However, it is not a simply stated fact that spoken communication is always interactive and face-to-face. The news presenter on TV is not interacting with his audience although the audience can see and hear him. Similarly, in a telephone conversation, body language cannot be used informatively (although changes in tone of voice certainly can). Likewise, it is perfectly possible to have written on-line communication by means of text telephone and computer talk systems. None of these situations qualify as "typical" situations of speech and writing respectively as they were described above, and it can be assumed that the different conditions of production will indeed result in different forms of language and language use. Tannen (1985) points out that what is thought of as spoken discourse is often spontaneous face-to-face conversation, and what is thought of as written discourse is expository prose. This is not by coincidence, she states, since there "is something typically written about message-focused communication, for it is the innovation of print that made it common to communicate on a large scale with others who are not in one's immediate context" and she agrees that "there is something typically oral about interpersonal involvement" (1985: 129). However, she points out that involvement (in Chafe's sense) is not necessarily tied to oral, spontaneous face-to-face conversation, but is also an important feature of written literary works, narratives, and creative writing. Although the medium is the written one, the style is characterized by interaction.

Why is it that literary language builds on and perfects features of mundane conversation? I believe it is because literary language, like ordinary conversation, is dependent for its effect on interpersonal involvement. It fosters and builds on involvement between speaker and hearer rather than focusing on information or message. It also depends for its impact on the emotional involvement of the hearer. In contrast, expository prose, associated with literate tradition in the way we have seen, depends for its impact on impressing the audience with the strength and completeness of its argument, that is, with aspects of the lexicalized message.

(Tannen 1985: 139-140)

As a consequence of this, Tannen prefers to talk about "relative focus of involvement" as typical of spoken and non-typical of written language. This insight is of importance for this thesis since the type of written data included in this study (i.e. narratives) is closer to what Tannen calls literary discourse than expository texts.

Learning to read and write involves an important step in the language development of the child. The fact that the child has to "break the code" and conceptualize that the written system is a symbolic representation of spoken language, has important implications for the child's linguistic awareness. Olson suggests that writing by its very nature is a metalinguistic activity since "writing ... is a representation of language" (1991: 261), and that metalinguistic knowledge (i.e., that language can become the object of thought and discussion) "is [not] a precondition of literacy but rather ... a product of literacy" (1991: 259). However, although the emergence of writing most likely contributes to a greater consciousness of language, there is still a considerable distance to cover for the language developing school child. As has been established above, speaking and writing are complex matters, and it takes time to learn what is specific to the different types of modality, production conditions, and genres (for example, not only is the child to learn what is characteristic of written expository texts, but also how to use literary language in order to write a lively story). Control over the construction of spoken and written discourse respectively, continues to develop over time.

Due to their close relationship, the spoken language influences the acquisition of written language and the increasing consciousness of characteristics of written language might also have an impact on spoken language. Michaels & Collins (1984) compared fourth-grade children's speech styles with their writing styles by having them watch the same film and then both tell and write a narrative account of it. The results showed, that the children who relied on paralinguistic channels in speaking were more likely to write a text that was ambiguous (i.e., it was not clear which character that was referred to). These children neglected to compensate for the loss of the

27

paralinguistic channel in writing by lexicalizing connections that were signaled paralinguistically in speaking. Hildyard & Hidi (1985) found that until twelve years of age, there were no differences between elementary school children's oral and written texts. After that point in time, however, the written texts were found to manifest a higher degree of structural complexity, evidencing a higher consciousness about modality specific properties in the older children. Nordqvist & Strömqvist (1995) found that connectors like *and then*, that are characteristic of oral telling, were as frequently used in written as in spoken narratives by 9-year-olds. Twelve- and fifteen-year-olds, on the other hand, used them almost exclusively in the spoken narratives, evidencing an increasing awareness of different properties of speech and writing. To illustrate this point further, Strömqvist (1996) found that 15-year-olds used clearly more transitive constructions in writing than in speaking, while intransitive constructions were more frequent in speaking than in writing. The 9-year-olds included in the study, on the other hand, used exclusively intransitive constructions in speech as well as in writing.

Using direct and indirect forms of speech in writing requires, not only an awareness of the differences between the forms, but also an awareness that what is reported by means of writing is *speech*. Speech represented in the written modality is functionally similar to, rather than equivalent to, speech reported in the spoken modality (this aspect will be further examined in Section 2.4.3.3). Speech reports made in the written modality can thus in a sense be seen as having a double metalinguistic property in that a metalinguistic use of expression (i.e., forms of indirect or direct speech) is made in a metalinguistic mode of expression (i.e., writing). Indeed, the use of these forms in speech and writing is an important and interesting part of the child's metalinguistic and pragmatic development.

2.3 Narrative

Events may be related in different ways. This section deals with how events are linguistically encoded and put into a narrative frame. This is of certain interest since speech events may, and indeed often are, encoded by means of direct and indirect speech.

A review of research reporting on narrative development is preceded by a section where narrative structure and different types of narratives are discussed and defined. Two types of narratives are then examined in greater detail: narratives elicited by means of picture series, and narratives produced within make-believe play.

2.3.1 Structure and types

Let us first make a distinction between three dimensions, which can be referred to as *story, narrative and narrating*. The literary scientist Genette (1988: 13) regards *story* as "the totality of the narrated events" and *narrative* as "the discourse, oral or written, that narrates them". However, I will regard narrated events as being *encoded* in the oral or written discourse (by the narrator), rather than as the events being narrated by the discourse. As a consequence of this, a story is seen as the set of events (whether real or fictional) inferred or reconstructed through the interpretational process of the narrative discourse. Thus, a story "is a series of logically and chronologically related events that are caused or experienced by actors" (Bal 1985: 5) and the narrative discourse is "a version of the core story ... that is actually realized" (Toolan 1998a: 605). Moreover, Genette regards *narrating* as "the real or fictive act that produces the discourse ... the very fact of recounting" (Genette 1988: 13). Here we will define *narrating* as the process of narrating a story.

A narrative typically has a beginning, a conflict and a resolution (see, e.g., Labov and Waletzky 1967), and narrative events are then organized in a sequential way, having a guiding theme knitting the events together. A

29

narrative is structured according to two levels; a *linear level* where events are ordered temporally, yielding the local structure of the plot; and a *thematic level* where episodic units consisting of linear sequences are organized hierarchically in terms of a given goal, yielding the macrostructure of the story (e.g., Bamberg & Marchman 1990, Karmiloff-Smith 1981, Berman & Slobin 1994). A narrative thus is a hierarchical structure that grows out of various sorts of global and local goals and functions (Strömqvist & Day 1993).

A narrative can be of different kinds, and for the framework of this thesis we distinguish two major types: the *personal* and the *fictional/fantasy narrative*, respectively. Uccelli, Hemphill, Pan & Snow (*in press*) state that "[c]ompetent renditions of these two forms of narrative generally share some important features: a focus on a protagonist or protagonists and a set of related actions that these actors carry out; the reporting of supportive detail such as setting or character attributes; the use of a range of strategies for linking events together and tying actions to consequences; and the inclusion of the narrator's evaluative perspective on the reported events." The two types can be distinguished from each other in that the personal narrative is used to report personal experiences, typically in past tense, while the fictional narrative is used by children, for example, in everyday fantasy play. The personal narrative corresponds to what Labov (1972) refers to as "one method of recapitulating past experience by matching a verbal sequence of clauses to the sequence of events which (it is inferred) actually occurred" (1972: 359-360). In addition to these two forms, *scripts* should be mentioned. A script is a narrative about what usually happens rather than about a specific incident. For instance, scripts can be enacted by the little child when playing house or doctor. Scripts do not have a clear plot, however, as Ninio & Snow (1996) point out, they can be seen as a prerequisite to the emergence of real stories. The characterizations of these above concepts can be summarized as follows:[1]

---

[1] Note that the taxonomy of narrative types is not meant to be exhaustive, and that the categories are not necessarily mutually exclusive.

STORY	a series of logically and chronologically related events that are caused or experienced by actors
NARRATIVE	the discourse, oral or written, that encodes them

PERSONAL	reporting of personal experiences
FICTIONAL	narratives produced within the context of fantasy (play)
SCRIPT	a narrative about what usually happens rather than about a specific incident

NARRATING	the process of telling a story

The universal narrative scheme introduced by Labov & Waletzky (1967) and Labov (1972), has had great impact on the study of narratives. According to Labov a "fully-formed oral narrative" consists of six elements: *abstract, orientation, complication, evaluation, resolution* and *coda*. The *abstract* tells the listener what, in a nutshell, the story is about, and in the *orientation*, the listener is informed about who, when, where and what matters. The *complication* consists of narrative clauses that present events, and the *evaluation* of a narrative is that part which reveals the attitude of the narrator towards the narrative by emphazising the relative importance of some narrative units as compared to others. The evaluation tends to immediately follow the complicating action and thereby distinguishes the complication from the *resolution*, in which it is told what finally happened.[2] In the *coda* the story is summed up and the speaker 'bridges' back to the present speaker-addressee situation.

An important feature of Labov's work, thus, is the evaluative function. A narrative should be expressive, have a point, and be deemed to be worth telling, and evaluation consists of all the means used to establish and sustain the point, the contextual significance and tellability or reportability of a story. There are several ways of evaluating and thereby create, as Toolan (1998b: 625) describes it, "a temporary suspension of the action, a brief 'time out' from the telling of the story proper". Evaluations, according to Labov & Waletzky

---

[2] However, Labov (1972) modifies his own scheme from Labov & Waletsky 1967 by pointing out that not only do evaluative devices occur directly after the complicating action but are distributed throughout the narrative.

(1967), range from highly internalized types (a symbolic action or evaluation of a third person) to the most external (a direct statement of the narrator to the listener about his feelings at the time). Important for the present thesis, is that direct and indirect speech are often employed for these purposes. It must be stressed that direct speech is not only used to make a story expressive and to heighten the listener's degree of interest. It is also the case, as Labov & Waletzky state, that in most narratives an evaluation section carries out the function of emphasizing the point where the complication peaks (that is, the break between the complication and the result). Direct speech is thus also an option available to the narrator when informing the listener of the point of a story. However, in addition to this purpose, direct speech may be employed by the narrator to move the story forward, that is to say, a means for advancing the plot. Thus, direct speech is not necessarily only used in order to make an evaluative statement of an event, but can also be a part of the complication. This is particularly true of fictional narratives where speech of direct and indirect form often move the story forward chronologically. However, the narrative function of direct and indirect speech is also salient in personal narratives. Vincent & Perrin (1999) found that 52% of the occurrences of reported speech in a corpus of sociolinguistic interviews had a narrative function, whereas only 19% had an evaluative (or appreciative) function.[3]

In addition, it is possible to present the narrative events from different perspectives, using different voices. A *narrator* moves with events in narrative time; an *author* can regard the events from an outside position; and the events may also be presented from the perspective of the *protagonists*. It is possible to choose only one of these stances, but it is also possible to move back and forth between them, (and in the case of protagonists it is even possible to take the roles of several protagonists), resulting in colorful and multivoiced stories. Wolf & Hicks (1989) studying young children and their replica play (i.e.,

---

[3] In addition to these two functions, it was found that 18% of the reported utterances had an authority function (argumentation appealing to authority), and the remaining 10% had a support function (argumentation by example).

32

symbolic play with small replica-sized figures), found that the children skilfully produced multivoiced narratives early on. A *narrative voice* was used to relate to the main line events (e.g., "...they walked through the forest to find a house, and said:"); with *characters' dialogue* it was conveyed what characters said aloud ("hey, where's my mommy and daddy?"); and finally, when *stage-managing* the children negotiated the literal conduct of their narrative with their audience (e.g., looking to the adult and saying: "that's their house"). Typically, the narrator's voice was marked by past tense and third person, and character's dialogue by present tense and first and second person.

2.3.2 Development

Peterson and McCabe (1983) investigated oral narratives produced by 4- to 9-year-olds and found that, alongside the increasingly common Labov-style story, (with a high point and resolution), a less complex chronological structure was found to be common among younger children, and even persist in the older children's productions. These stories had temporal sequence but little or no sense of contour, resolution, or evaluative point. Karmiloff-Smith (1981), studied narrative development across subjects of a similar age group, 3 to 12 years of age, by means of a picture story task. Several elicitation instruments were used, for example, the *Balloon Story* consisting of a series of six pictures.[4] Focusing on anaphoric devices, Karmiloff-Smith found that the youngest narrators tended to use pronouns as grammatical subject, not realizing that the pronominal reference might be ambiguous (*he* could be used by the children throughout the narrative despite the fact that it sometimes referred to the boy, and sometimes to the vendor). Moreover, the children typically pointed to the pictures describing what they saw, indicating preoccupation with the individual pictures. In the age group of 5 years, the boy was usually introduced with a noun phrase and subsequently referred to with a pronoun. Thus, a global

---

[4] The pictures depict a boy; he buys/gets a balloon from a vendor; the boy lets go of his balloon, and he starts to cry.

structure emerges, by means of a thematic subject (the boy). However, only the older children (8-9-year-olds) typically managed to combine the global and the local strategies. Not only the boy was introduced by a noun phrase, but in the picture where the vendor appears and has a foregrounded role, the vendor was introduced and referred to by a noun phrase. In conclusion, the two above-mentioned studies show that older children, in contrast to younger narrators, manage to construct a narrative with a coherent temporal structure and an anaphoric structure of reference to key characters.

Ninio & Snow (1996: 186), discussing younger children's narrative abilities, suggest that: "Rather than the problem being difficulty in understanding the characters' internal states, we are arguing that the central pragmatic problem facing the producer of narrative discourse is understanding the *interlocutors'* internal states sufficiently to predict the questions they might be asking themselves, to assess their relevant background knowledge, and to be able to influence their perspective." In order to make a narrative comprehensible to a listener and, for example, succeed in introducing and maintaining referents properly, the child narrator needs to put herself in the perspective of the interlocutor and consider the degree of shared knowledge between herself and the receiver of the narrative. Constructing personal and fictional narratives, (see previous section), requires similar abilities of the narrating child. An example of this is setting off the narrative from the surrounding talk and reporting and linking events together. However, the two types also impose different demands on the narrating child. When relating a personal narrative, the child selects a set of happenings from the flow of past experiences that cohere, and in forming a story, considerations of what information is shared and unshared with the listener is required. This is necessary in the case of fictional narratives as well, however, in producing a fictional narrative within make-believe play, joint attention to a play figure or object can substitute for the more elaborate referential strategies that are characteristic of, i.e., personal narratives. Moreover, as Uccelli, Hemphill, Pan & Snow (*in press*) contend, successful fictional narratives also require skill in

plot improvisation and the ability to create tension and interest by, for instance, having the play figures or the story characters speak.

Uccelli *et al* present a longitudinal study where 32 parent-child dyads were observed and video recorded when interacting in a laboratory playroom using toys and materials provided by the investigators. The recordings were made when the children were 20 months and 32 months. In addition, at 5 years of age the children were asked to tell a personal narrative and a fantasy (fictional) narrative using small toys and props. The personal narratives of the 5-year-olds were coded, among other aspects, for the presence or absence of *structure* (including advancing of plot), *orientation*, and *evaluation*, and it turned out that the 5-year-olds of this sample showed individual differences. The "better narrators" produced conventional narratives that followed a sequential structure and were organized around a peak or climax, and the narratives clearly conveyed the narrator's perspective through use of evaluative devices. "Poor narrators" included little evaluation, and the narratives had either a non-sequential structure or an obscure progression of events. In the narratives of fantasy, the "better storytellers" used conventional narrative structure to frame and organize their fantasy performances, and build up a fully-realized story world. There was a story protagonist around whose plans, actions, and reactions the story events were developed, and the perspectives of story characters were displayed by the use of direct speech. The less successful fantasy narrators, in contrast, relied on their adult partner to provide information about the story setting, participants, and plot, and made little evaluation.

Uccelli and her colleagues also found that talk between the children and the adults about nonpresent objects, events and attributes at 20 and 32 months of age, was predictive of the children's skills as narrators of personal experiences and fantasy at 5 years of age. This means that the more the children were engaged in early talk about topics other than "here and now", the more likely it was that they belonged to the group of better narrators at age 5. This correlation was not found between early fantasy talk, (e.g., assigning roles

35

and actions to characters in a make-believe context), and skill in fantasy story-telling at age 5. However, early participation in fantasy talk was found to stimulate the development of specific areas of fantasy narrative, among other things, the representation of character voice (use of direct speech).

The Uccelli *et al* study suggests that the quality of early interaction has implications for later language (and narrative) development. Certainly, it is also the case that the more the children are exposed to narratives and books, the more likely it is that the child grasps what constitutes a good narrative. However, what these narrative elements are, how a story should be told, and what place oral and written narratives have in everyday life, are to a great extent culture dependent. Tannen (e.g., 1982) has shown that mis-understandings can occur between, and even within, societies because of different narrative and literate traditions, and this might have consequences for children's adaptation to the culture of the classroom when they start school (Heath 1986; Nauclér & Boyd 1997; Wolf & Hicks 1989).

2.3.3 Picture-based elicitation

In order to induce children to produce narratives, a common method has been to use a picture story task, e.g., the *Balloon story* (mentioned in the previous section), or the *Cat story* and *Horse story* used by Hickmann (1982). One advantage of this method, among others, is that the memory burden will not be as overwhelming as in other procedures (for instance, when recalling and recounting the events in a particular film), and as Berman & Slobin (1994) contend, pictures have proved to be a reliable means of tapping children's narrative abilities from both a cognitive and linguistic point of view. It should be kept in mind, however, that this works best in cultures with a literate tradition and where picture stories are familiar to the child (see Strömqvist & Verhoeven (*forthcoming*) for cultural issues on this topic).

A frequently employed picture story in developmental studies is *Frog, where are you?* (Mayer 1969), which is also used as an elicitation instrument in

this thesis work. The frog story consists of 24 pictures and is about a boy and his dog who are searching for a missing frog (see Appendix). There are three core components to the story: the boy's realizing that his frog is gone, the boy's search for his missing frog, and the boy's rediscovery of the lost frog.

In a major study using the frog story by Berman & Slobin (1994), five different languages were compared from the point of view of children's development of narrative competence (English, German, Hebrew, Spanish, Turkish; each study made separately and independently (for a full account and references, see Berman & Slobin 1994, Chapter IIA)). Children ages 3-, 4-, 5-, 9 years, and adults were included. Typically, narrators from the two youngest age groups failed to make explicit reference to the three core components, i.e., they had problems with the global structure. The 5-year-olds did better, but the 9-year-olds were much better than the 5-year-olds in this respect. This means that the school children were able to construct a global hierarchical theme or story-line, in contrast to some of the 5-year-olds who constructed their texts in a simplistic additive fashion, utterance by utterance. Moreover, Berman & Slobin contend that "[w]hat pre-schoolers cannot do is embed individual events within a network of associated circumstances that constitute the background events and internal motivations which lead up to a given event and the situations which follow from the event in question" (1994: 57). Berman & Slobin present a developmental continuum representing four phases in the evolution of narrative capacities:

(a) spatially-motivated linking of utterances as picture-by-picture description (3-year-olds)
(b) temporal organization at a local level of interclausal sequential chaining of events (most 5-year-olds)
(c) sequential and/or causal chaining of partially elaborated events (most 9-year-olds)
(d) global organization of entire texts around a unified action-structure (some 9-year-olds, and the adults)

The youngest children sometimes only listed what they saw in the pictures, i.e., they produced descriptions rather than a narrative. Nonetheless, as was the case

with the older children, many 3-year-olds related to the pictures as depicting *events*, evidencing some narrative awareness. Moreover, these children often personalized their accounts, and wanted to tell the story interactionally. The children typically failed to establish an anchoring tense, but manifested a mixed tense usage (shifting from one tense to another without the shifts being thematically motivated). *Here* and *and (then)* were used utterance-initially to mark connectivity. The 5-year-olds proved to be a heterogenous group, some narrators being consistent in tense use, whereas others were not; some using elaborate syntax and rich lexicon while others produced poorer narratives in these respects. These findings are in accordance with those presented in the previous section. In addition, the 5-year-olds telling the frog story frequently connected clauses on a local level with *and*, *then*, or *and then*. The school-age children were more homogenous. They demonstrated good command of complex syntax, had more elaborated descriptions of individual events and of the relationship between one event and another, referred to internal states of the protagonists, and anchored their narratives in one tense.

The adult narrators manifested no problems with global versus local structure. Interestingly though, no particular type of narrative as an adult "model" could be pinpointed, since they produced very different types of stories: "several different profiles emerge across each group of adult texts. Each such 'profile' represents a cluster of different rhetorical options and expressive means chosen by the narrators from a mature repertoire of linguistic forms and anchored in a mature conception of the narrative genre" (Berman & Slobin 1994: 79). Thus, although the 9-year-olds have learned to use many of the linguistic devices necessary to make a narrative coherent and complex, and although they have become more familiar with the "narrative norms" of their culture, it takes even more time to develop narrative skills to the extent that the narrator can freely choose what kind of genre 'profile' to adopt.

Above (Section 2.3.1) it was mentioned that direct speech can be used for evaluation and to make a story vivid. Several pictures in the frog story present events where protagonists can be seen using their voice, (for example, the boy

calling for the frog) and it is therefore to be expected that some narrators produce direct speech in their narratives. Despite this, analysis of direct and indirect speech in frog story narrations is to a large extent neglected (see, however, Bamberg 1991, Reilly 1992, Bamberg & Reilly 1996). In the analyses included in Berman & Slobin (1994), for example, no mention is made at all of direct and indirect speech. (In this volume Bamberg discusses evaluations more thoroughly, but only from the point of view of expression of mental/internal states and motives to act (1994: 234-237)). This might be a result of viewing direct speech as having as its primary function to make the story more dramatic and interesting to listen to, (i.e., an act of performance, rather than a way of relating events). However, as I have argued above, direct speech and indirect speech may serve many different functions, among other things, that of complicating an action, (more functions will be mentioned and discussed in e.g. Section 2.4.3.4). This fact is one reason why direct speech should not be neglected, but carefully analyzed, in the study of the frog story. The frog story picture booklet will be presented and discussed in greater detail in Chapter 3.

2.3.4 Make-believe play

The distinction between personal and fictional narratives was made above, and narratives produced within fantasy were referred to as belonging to the latter type. Narratives like the frog story, as well as narratives produced within make-believe play, fall within this category since they are both a product of pretense. In her book *Play*, Garvey (1990) defines make-believe as "a voluntary transformation of the Here and Now, the You and Me, and the This or That, along with any potential for action that these components of a situation may have" (1990: 82). A make-believe world has its own truth conditions, and the truths are, as Evans points out, "a species of the genus of fictional truths" (1982: 353). Evans continues, arguing that "it is a characteristic of games of

make-believe that one can discover such-and-such to be make-believedly the case" (1982: 353).

In discussing games of make-believe, Evans distinguishes between two types: *existentially conservative* and *existentially creative* games, respectively. In the case of the former, the pretense lies in pretending that something which is there, is other than it is. Evans gives the example of children pretending that globs of mud of a certain shape and size are pies. An example of an existentially creative game, in contrast, is shadow-boxing. In this case, a boxer pretends that there is an opponent he is fighting, i.e., that there is something which in fact is not there. Sawyer (1996), investigating children's role play, distinguishes *direct voicing* (the child's body "becomes" the play character) from *indirect voicing* (the child enacts a play role through the medium of a toy figure). In the context of this thesis, Evans' concept of *existentially conservative game* and Sawyer's *indirect voicing* (which should not be confused with indirect speech), are particularly relevant. Playing with a doll house includes pretending, for instance, that one doll is "Mother" and another "Little sister" (existentially conservative actions) [5], and when spoken utterances are attributed to the dolls, this is an example of indirect voicing. Indirect voicing is also an important feature of story-telling, as when narrating the frog story.

A narrative produced within make-believe play, may, but need not, contain all the components presented in Section 2.3.1. Imagine two little children playing house, and one of the children has the doll, Charlie, going out for a ride in his toy car:

Charlie was out having a ride with his new car
he was going to visit his friend Danny
when he had been out driving for a while, something suddenly appeared on the road
what is that!?!
he cried out and stopped his car immediately
his heart was bumping hard because he was very scared

---

[5] Note however, as Strömqvist (1984) points out, that existentially conservative games of make-believe can be ordered along a scale as more or less existentially creative.

then it turned out that it was just a friendly giraffe
it came towards Charlie and licked him on the face
Charlie started to laugh
what a funny giraffe he was!

In this segment (henceforth "the Charlie-narrative"), we identify a clear
beginning (including an orientation presenting Charlie and that he is out riding
his car), complicating actions (Charlie is driving and has been doing that for a
while when something suddenly and abruptly happens), evaluations (direct
speech,[6] and presentation of internal states: Charlie is *scared* because of the
abrupt event), and finally a resolution (a "happy ending" as it turns out that it
was only a friendly giraffe that would do no harm, and Charlie can laugh in
relief). Shorter narratives may well be embedded in larger narratives. In
relation to this, I would argue that make-believe play structure *as such* has clear
affinities to that of narrative structure. In the case of the Charlie-narrative, we
can imagine that the two children playing, started playing by one of the
children suggesting: "Let's play house!". The participants are then introduced
to the make-believe frame. Further, the children orient themselves in the make-
believe frame and set the make-believe prerequisities: "This is Charlie (holding
up a doll); this is his car (holding up a toy car); let's pretend it's afternoon". In
this way, the sequence of ordered events as presented in the Charlie-narrative,
is preceded by a more general make-believe orientation phase. In the same
manner, the Charlie-narrative may be followed by a coda, signaling stepping
out of the make-believe world, like: "Let's play something else now!". Thus,
using this analysis, the Charlie-narrative is part of a larger play-narrative.

Garvey (1990) does not see play as a narrative in and of itself, however,
she compares narratives and the construction and enactment of pretend
scenarios and finds a number of similar and probably related trends. She

---

[6] The Charlie-narrative contains one clear instance of direct speech (as part of the evaluation and
complication), namely "'what is that!?!', he cried out". "What a funny giraffe he was!" at the end
of the story can be interpreted in two ways; either it is seen as something Charlie expresses in the
form of direct speech (without an explicit speech and speaker introducing clause), or it is seen as
a comment or speech act originating from the narrator role, i.e., the narrating child. In the latter

contends that in both types, the plot thickens and events are linked and related. Garvey points out that already in the two-year-old's play actions, (often verbally accompanied), with dolls and replica objects (like feeding a doll and putting it to bed), a concern with the orderly sequencing of familiar events is revealed. As the children start to learn to integrate narratives into ongoing conversations, they learn to add narratives having different functions to the flow of make-believe play. Narratives and make-believe episodes become more explicitly framed as entities distinct from other ways of talking and interacting that children may command, and both become less subject to interruptions or distractions from the immediate environment. Wolf & Hicks (1989), studying children's replica play between 3 and 6 years of age, found that narrative segments (including dialogue) generally became longer by age. Garvey (1990) also points out a paradoxical aspect in the growth of make-believe: As children become older, their play becomes more "fantastic" and less tied to everyday activities such as preparing meals or putting babies to bed. At the same time, children's use of objects in make believe-play becomes increasingly "appropriate" and realistic, and they represent more and more of the adult world in ways intelligible to us.

Hence, younger children tend to prefer more familiar events in their make-believe play. So, what exactly is the relation between narratives, and for example, playing house? Since the play in this context tends to enact scenes from real life; are we dealing with personal narratives? I do not believe that to be the case. Children as young as three distinguish between the real world and the fantasy world (Garvey 1990; Wolf & Hicks 1989; Strömqvist 1984). The children are *inspired* by familiar experiences, however, they are aware that they are only pretending (this is seen in, for instance, the fact that play situations often are introduced by statements like "let's play house"). Moreover, as Garvey points out, most enactments are most likely not direct imitations of models. Garvey gives the example of a boy who walks into the house

---

case, it may rather be classified as free indirect speech. Free indirect speech will be introduced and discussed in greater detail in part 2.2.

announcing: *Okay, I'm all through with work, honey. I brought home a thousand dollars*, and then hands over the money to his pretend-wife. Most likely the boy has never witnessed this scene in his own home. However, although dismissing this kind of narratives from the category of personal narratives, it is not self evident that they all belong to the category of fictional narratives (as characterized in Section 2.3.1). The younger the children involved in the play, the more likely the narratives are to have little plot and lack several of the features as represented in the Charlie-narrative above. Rather, they are examples of *scripts*. When telling highly-scripted narratives, children typically mark their accounts by use of the present tense, second person pronominals, and sequencing connectives used in a generalized sense (Nelson 1986). As Ninio & Snow (1996) suggest (see Section 2.3.1), narrative scripts are an important preliminary stage to fictional (and personal) narratives.

2.4   Direct and indirect speech

In order to examine children's use of direct and indirect speech, these particular concepts need to be discussed and defined. This section is organized in such a fashion that traditional approaches, terminology and debates are presented and discussed first (Section 2.4.1), and this discussion serves as a base for the definitions and the theoretical model that I present in the next section (Section 2.4.2). In addition, structuring of information in direct and indirect speech is dealt with (Section 2.4.3), and the final section presents previous research on the issue of children's use of direct and indirect speech in spoken and written discourse (2.4.4).

2.4.1   Traditional approaches and terminology

2.4.1.1 Forms of speech

'Direct and indirect speech' is traditionally used to refer to expressions such as Example 2.1 and Example 2.2.

Example 2.1

Robert said, "It's hot in here".

Example 2.2

Robert said that it was hot in there.

Example 2.1 is an example of direct speech and contains two clauses relatively independent of each other. The first is an introducing reporting clause, including information about who is to be quoted (Robert) and a verb of saying (said), and is followed by what is assumed to be a repetition of the actual spoken utterance of that person. The deictic center is then moved from the reporter to the reported/original speaker, (note the shift in tense from the reporting clause to the reported clause: said - is). In written language the orthographic convention of placing quotation marks around the reported speech signals that quotation is occurring, and it marks the quotation as a syntactically independent utterance. In oral language, on the other hand, a variety of prosodic conventions may be employed, to roughly produce the same effects (a further discussion of these matters follows in Section 2.4.3).

In Example 2.2 - the example of indirect speech - the deictic center is where the reporter is and elements such as pronouns, verbs and adverbs from the original utterance must conform to the here-and-now of the act of reporting. Indirect speech is also syntactically characterized by the presence of the subordinating conjunction *that* (although optional), and the complement clause as being subordinate to the clause with the speech act verb. From a grammatical (syntactic and semantic) point of view, this means that what is reported is integrated with the narrative context.

A distinction between 'de dicto' and 'de re' is often made in connection with direct and indirect speech, and refers to the relationship between the reported utterance and the original utterance. A 'de dicto' reading of Example 2.2 gives what the speaker reported on *actually* said, while the 'de re' reading gives the reporting speaker's *interpretation* of what was said. In the former

—

44

case, the original utterance is adjusted to the deictic center of the report situation without changing any other part of its linguistic form, whereas in the latter case, the reporter may alter the form of the original utterance which may include inferences about things of which the original speaker is unaware (importantly though, the reporter and the original speaker have the same reference). Typically, this ambiguity, i.e. whether a reported utterance is a reproduction of the original utterance, or if it has been filtered through the reporting speaker's interpretations and appraisements, is said to be present in the case of indirect speech, whereas direct speech typically has a 'de dicto' reading (see e.g., Coulmas 1986: 4-5). Yet another type, which may be regarded as a special case of 'de re', is 'de intentione'. In this case the reporter reports what he thinks was the original speaker's *intention*. For example, the report in Example 2.2 may be a result of Robert's earlier utterance: "Phew! I really have to take off my down coat!". We will return to these aspects later on.

To report speech seems to be a universal linguistic activity (although only the direct form and not the indirect form is universally distributed (Li 1986)). The distinctions, however, between the two main forms as described above may look and function differently in different languages (Coulmas (ed.) 1986). It is even the case that in some languages, Japanese for instance, is it difficult to determine if the distinction between the two ways of reporting speech exists, (for different analyses and findings of the forms in Japanese, see Coulmas 1985, 1986b, and Maynard 1986). Haberland (1986) points critically to the fact that the model that makes a clear-cut division between direct and indirect speech was primarily developed on the model of Latin and Classical Greek and applied to the study of literature of those languages in certain historical periods. This could certainly cause problems when applying the distinctions to colloquial, oral discourse, not to mention non-European languages. Indeed, although the language Haberland studies, Danish, is a European language, he finds several ways of reporting speech in his material that do not fit the traditional division. One reporting style he frequently finds in his corpus is a

form where direct and indirect speech are mingled. To use the example of Robert again, the mixed type as in Example 2.3 is likely to be found.

Example 2.3

Robert said that it's hot in here.

This example illustrates that all elements for direct speech are present; the word order characteristic of a main clause is preserved, but the report is introduced with the complementizer *that* (which is typical of indirect speech and as an introducer of a subordinate clause). Haberland comments that the complementizer in such cases seems to function as a general report marker. Haberland concludes that he prefers to interpret them as two general tendencies (rather than classifying reports into two mutually exclusive and clearly distinguishable sets).

2.4.1.2 A literary bias?

Direct and indirect speech have been studied for a long period of time, and then often in connection with literary analyses. Has this focus on written literary texts (in which *speech* is represented in the *written* mode), and the fact that many literary scientists have carried out analyses and constructed theories of types of speech reports, had any particular effects on the study of direct and indirect speech as a whole?

A mode of reporting speech in literary texts that started to attract attention in the 19th century, first mentioned by Tobler in 1894 as a "mingling of direct and indirect speech", is the form that is today best known as 'free indirect speech'. (This form should not be confused with the example given in Example 2.3). The essence of this type of stylistic device and the name it should have, has been discussed extensively over the past century.[1] Free indirect speech is an

---

[1] For example, Kalepsky (1899) wanted to see it not as a mixture of direct and indirect speech, but as a third kind, "veiled speech" (*verschleierte Rede*). Bally (1912) regarded it as a special case of indirect speech (*free indirect style*), while Lerch (1919) wanted to treat it as "quasi-direct

indirect quote (or representation) of a character's speech or thoughts typically "free" of a framing clause and characteristically with anaphoric elements shifted to achieve greater integration with the surrounding text. The third-person perspective of the narrator is represented along with a first-person perspective of a speaker. There is typically a double ambiguity involved: confusion between speech and thought, and between character and narrator (voice). Thus, a certain instance of free indirect speech might be intended to be the speech *or* the thoughts of the protagonist *or* of the narrator. It is left to the reader to disentangle these aspects, and in many cases it is not even possible. A free indirect version of Example 2.1 could look as follows:

Example 2.4

It was hot in there.

Structurally, these intertwined points of view are commonly represented by person and tense deictically belonging to the narrator's point of view (as typical of indirect speech) but the sentences are nonembedded, independent clauses showing all the characteristics of main clauses (like direct speech). Questions retain their syntactically direct inverted form (Jespersen gives the example of direct speech "How can I bear to look any of them in the face now?" which in free indirect speech becomes "How could he bear to look them in the face now?" (1922: 298)). Moreover, Banfield (1993) points out that, like sentences of direct speech and unlike the embedded clause of indirect speech, sentences of free indirect speech may contain exclamations or expressive constructions such as exclamatory sentences. Consider Example 2.5, an extract from Virgina Woolf and also an example of free indirect speech, and compare with the indirect speech version given in Example 2.6 (Banfield 1993: 342-343).

---

speech", a special type of direct speech. A term often heard in these contexts is the one Lorck (1921) proposed, "experienced speech" (*erlebte Rede*). Jespersen (1924), however, criticized Lorck since he thought "the writer does not experience or 'live' (*erleben*) these thoughts or speeches, but represents them to us" (1924:291-292). Instead, Jespersen introduces the term "represented speech".

Example 2.5

She would not "give a flower to the gentleman" as the nursemaid told her. No! no! no! she would not!

Example 2.6

*Cam insisted that, no, no, no, she would not.

Jespersen comments: "Represented speech [Jespersen's term for free indirect speech] is more vivid on the whole than the first class of indirect speech. As it is nearer to direct speech, it retains some of its elements, especially those of an emotional nature, whether the emotion is expressed in intonation or in separate words like 'Oh!', 'Alas!', 'Thank God!', etc." (1924: 292).

Literary scientists thus have paid a lot of attention to forms of direct and indirect speech. Let us now return to the more general question stated in the beginning of this section. Has this literary focus had any particular effects on the research area, and more specifically, is free indirect speech seen as a stylistic device reserved only for (written) literary texts?

Some researchers do indeed regard free indirect speech as a stylistic device and a form particularly suited for, and developed for, written literary texts. Polanyi (1982) argues that it is rare in spoken discourse and notices that when it does occur, it is usually in the context of reporting stories in which the narrator only possesses indirect evidence via another narrator (Polanyi gives authentic examples from reports of movie plots (Eisner 1975)). In this way the narrator limits his/her responsibility to what could reasonably be expected of someone to remember of a situation in which the speaker was never a participant. Banfield (1993) is of the opinion that free indirect speech is a literary form where "the expressive function emerges free of the communicative function" and is "one form of (expressive) non-communication"[2] (1993: 339-340). Her claim builds upon her observation that "[i]n no examples which are unambiguously represented speech and thought

---

[2] Banfield (1982) even considers free indirect speech "unspeakable".

does a second person appear, although the first does; hence, the former becomes the essential sign of communication and the SPEAKER/HEARER pair defines communication only" (1993: 352). She claims that writing releases language from the communicative function while in spoken discourse the communicative framework (relation between a speaker and an addressee/ hearer) is inescapable. Consequently, since both writing and free indirect speech are regarded as typically noncommunicative, this is an explanation to why free indirect speech exists almost only in written texts, Banfield argues.

Haberland (1986) does not agree with the opinion that free indirect speech rarely occurs in spoken discourse. He distinguishes free indirect speech as a category of literary stylistics from the same phenomenon as a structural-grammatical category and is critical of the view that free indirect speech is a stylistic means mainly to be found in literature. Haberland found many examples of free indirect speech in a corpus of colloquial Danish he studied (the example below is taken from Haberland (1986: 247).

Hun ville have at jeg skulle være barber. Det var sådan et renligt og pænt arbejde.
'She would have liked me to become a barber. It was such a clean and neat occupation.'

Hence, Haberland represents the view that free indirect speech is not reserved for the written modality and literary contexts and also shows authentic examples of that. This means that this form *is* in fact to be found in at least spoken, less formal Danish discourse. Interestingly, researchers investigating spoken discourse often do not even consider this category and do not include it in their analyses (which mostly include only indirect and direct speech). In this way, the focus on written language and the prevailing view in the literary tradition might be seen to have influenced the investigators of spoken discourse. However, as Sell (1998) points out, today's literary pragmaticists tend to have a view closer to that of Haberland, than that represented by Banfield. Teleman, Hellberg & Andersson (1999) in *Svenska Akademiens Grammatik* (SAG) (The Reference Grammar of the Swedish Academy)

describe free indirect speech (*referatmeningar*) along with the description of direct speech (*citatmeningar*) in Swedish, and provide similar space to both types.[3] SAG, which aims at covering spoken as well as written Swedish, seems to conform to Haberland's view that free indirect speech used as a stylistic literary device should be distinguished from the same form as a structural-grammatical category. Thus, whereas the latter type is regarded as being as frequent in spoken as in written discourse, the former (i.e., the stylistic category) is described as typical to literary texts and treated separately in a relatively short paragraph with the title *utvidgad användning av referatmeningens karakteristika* ('extended use of the characteristics of the reported [free indirect speech] sentence').[4]

A relevant question is if there are forms of reported speech that are typical of spoken discourse rather than written. Consider the example below:

Example 2.7

"It's hot in here."   [uttered with a disguised voice]

Imagining Example 2.7 being uttered by Lisa in a conversation with Tanya, it might have appeared perfectly clear to Tanya that Lisa is retelling what Robert had said in a certain situation since Lisa's modified voice quality indicated that she was impersonating Robert. Thus, information on speaker and speech reported are provided although no concrete verbal speaker and speech introducing clause exists. If Example 2.7 is encountered in a written novel, by contrast, the dimension of voice quality is no longer present, but from the fact that the text is embraced by quotation marks it is clearly indicated that it is a quotation. Furthermore, if a larger discourse context was given, the speaker is likely to have been easily identifiable, at least by inference.

According to the characterization given in the preceding section, Example 2.7 would not qualify as direct speech (since there is no speech framing

---

[3] *Anförda meningar*, 'Quoted sentences'. Vol. 4, Chapter 41, pages 844-876.
[4] Vol. 4, Chapter 41, §18.

clause), and although this type of form is exhibited also in written prose, this structure is traditionally not discussed in the (literary) analyses. If it is included, it is often regarded as a token of free indirect speech, as Haberland (1986:246) critically makes note of: "The lack of a verb of saying is usually attributed to quasi-indirect discourse". Another common way of handling this type of form is to see it as a (less developed) form of direct speech. Direct speech with a framing clause is then seen as the prototypical case. In the previously referred to SAG this structure is mentioned only briefly, and this strengthens the impression that direct quotes are only by way of exception unframed.[5] I claim that this "unframed" direct speech structure should be categorized as a form of speech on its own, on the same level as the direct quote with a framing clause, indirect speech and free indirect speech is in the literature today. I will give more substance to this claim in Section 2.4.3, and demonstrate that the structure shown in Example 2.7 exists and works well, both in written and spoken discourse.

The question is if this tradition of division, built on analyses of written texts, has left its mark on the research area of spoken discourse also, since investigators focusing on spoken interaction often overlook this type of speech (or at least see it as a subordinated type of the "full-blown" form of direct speech that has a framing clause). It is possible, and in that case interesting, since the analyses of speech reported in written texts then influence the way of analyzing speech reported in speech. Indeed, Coulmas (1998) argues that literacy might have shaped the Western notion of the identity of utterance tokens, and calls for more studies of the differences between written and spoken speech reports. I will return to the question of labelling forms of direct speech in Section 2.4.2.1.[6]

---

[5] For a critical review of the chapter on quoted sentences in SAG, see Nordqvist (2000a).

[6] It should be mentioned that in relation to direct speech, a separation of the forms is sometimes explicitly made. Hickmann (1993) and Özyürek (1996), for instance, use the term "unframed direct quotations" in contrast to "framed direct quotations", and Romaine & Lange call cases where dialogue appears without the support of verbs of saying "bald, unframed, or unbracketed reporting" (1991:235).

2.4.1.3 The problematic 'report'

Already the pioneers in the area of direct and indirect speech had, as we could see in the previous section, problems agreeing on a nomenclature. I believe that the confusion is just as persistent today. Up to this point, I have fairly uncautiously used the term 'reported speech'. This term is by far the one most commonly used along with 'direct and indirect speech'. Having dealt with the terminological matters of direct and indirect speech already, this section will primarily deal with the criticism that has been raised on the use of the term 'reported speech'.

There are two main problematic aspects connected with the (traditional) use of reported speech as a term for speech about past speech. The first aspect concerns a terminological confusion, i.e., what types of speech forms are actually referred to as reported speech. The literature does not address this matter as much as the second problematic issue. This issue consists of the relationship between the literal meaning of the word 'report' (and the connotations it gives rise to) and the process of transformation that certain theorists want to ascribe to this particular language activity.

Upon reviewing the literature, it becomes evident that the term 'reported speech' is most often used as a cover term for direct and indirect forms of speech (e.g., Voloshinov 1973; Coulmas 1986a; Hickmann 1993; Polanyi 1982; Goodell & Sachs 1993; Özyürek 1996). However, according to some researchers, 'reported speech' refers only to indirect speech in contrast to direct speech (e.g., Banfield 1993, Kvavik 1986), whereas, quite on the contrary, for instance Genette (1988) refers to direct speech as 'reported speech' and to indirect speech as 'transposed speech'. Thus, no coherent picture of what counts as reported speech can be discerned among researchers addressing the phenomenon. One reason for this inconsistent use of the term in the research community, might be the possible subjective and intuitive interpretation of the general meaning of 'report'. In *The Merriam Webster Dictionary* (1994) there are two definitions: "1) to give an account of: RELATE, TELL; 2) to serve as a carrier of (a message) ... ". The meaning is fairly general and broad. There is,

for instance, no necessary direct and exclusive link between *Robert said that it was hot in here* and 'report', which is not there between *Robert said, "It's hot in here"*, and 'report', and vice versa. The terms chosen by individual researchers might have been chosen only for convenience and stipulated for the purpose of their current study while no greater attention has been paid to how other researchers have chosen to use them.[7] I will return to this, not unproblematic, question of distinctions at a later point.

There has also been an extensive discussion concerning the nature of the speech directly quoted and its relation to the original utterance. The correctness of talking of the speech as actually *reported* has been questioned by certain theorists. Objections have been raised to the line of reasoning that amounts to seeing direct speech as an exact reproduction, or an exact report, of an earlier speaker's talk. According to this exact reproduction view, there is no option to comment on the content of the speech reported, while when you retell what another person said in the form of indirect speech, there is an opportunity to interpret, and express attitudes and emotions in relation to the speech of the original speaker. The listener then understands that, for instance, non-verbal messages originate from the reporter and not the original speaker:

Both direct and indirect speech serve the function of marking a statement as that of someone else than the speaker. The main difference between them can be seen to lie in the speaker's attitude towards the reported speech. In marking an utterance as a direct quotation, he commits himself to faithfully rendering form and content of what the original speaker said; that is what direct quotation suggests. An utterance marked as indirect speech, on the other hand, implies a commitment about the contents but not about the form. (Coulmas 1985: 42)

In this view, the ambiguity of 'de re' and 'de dicto' interpretations is characteristic only of indirect speech, whereas this "kind of ambiguity cannot occur in direct discourse, because no interference of reporter's speech and

---

[7] In reviewing Coulmas (ed.) *Direct and indirect speech* (1986) Susan Thomas concludes: "...I would like to note my own impression that what the study of reported speech needs ... is a standardized terminology. There is a need for standard terms and definitions for nonreported and reported speech and for the varieties of reported speech." (In *Language in Society*, Vol. 18, 1989, pages 102-103.)

quoted speech is possible. Direct speech always has a *de dicto* interpretation."
(Coulmas 1986: 4). This tradition, advocated by, among others, Li (1986),
Palmer (1976) and Leech & Short (1981), however, has been criticized.
Bakhtin (1981), for instance, claims that, in principle, it is impossible to
transform a message from one discourse to another, since an utterance gets its
meaning from the context in which it occurs, i.e., an exact report is impossible.
Instead, there is typically a considerable amount of 'de re' analysis going on.
When examining the speech of Swedish teenagers, Eriksson (1994) as well as
Kotsinas (1994) find that a common way to use direct speech is to exaggerate,
or even invent, speech of another person.

Det vanliga är istället att citaten mycket ungefärligt och ibland ganska överdrivet
återger vad någon sagt eller bara tänkt. Inte sällan är de rent fiktiva och används som en
illustration till det talaren vill berätta om och tycker är viktigt. [Most commonly, the
quotations only approximately and sometimes very exaggeratedly report what someone
has said, or even only thought. Often they are completely fictional and used only as an
illustration of what the speaker wants to tell about and thinks is important.]
(Kotsinas1994:42 *my translation*)[8]

The aspect of invention, transformation and creation and the problems with
regarding this as reported speech, are something that Tannen stresses. In the
volume *Direct and Indirect Speech* (edited by Coulmas 1986) she opens her
chapter, fairly refractorial, since several of her co-authors use the
term 'reported speech', with "[t]he term 'reported speech' is a misnomer"
(Tannen 1986: 311). Elsewhere she writes that she wants to "demonstrate that
taking information uttered by someone in a given situation is an active
conversational move that fundamentally transforms the nature of the utterance.
This is in contrast to the folk wisdom by which the concept 'reported speech' is
taken literally" (1989: 105).[9] Consequently, she follows Bakhtin's lines of
reasoning, and introduces a new term: "I am suggesting, then, that what is

---

[8] To be exact, what Kotsinas here refers to is *pseudoquotation* - "invented quotation". This, by
definition, cannot be said to be an exact report. However, it still illustrates the fact that there is
direct speech that is not exactly reproduced. See also Dubois (1989).
[9] Note, however, as was mentioned in the introduction to this section, that some want to regard
only indirect speech as reported speech.

called 'reported speech,' 'direct speech,' 'direct discourse' ... should be understood not as a report at all, but as constructed dialogue. It is constructed just as surely as is the dialogue in drama or fiction" (1989: 110).

Taking this body of criticism of terminology used when discussing reported speech into consideration, it will soon be evident that 'direct quotation' can also be a problematic term. It would be problematic to claim that someone has been directly quoted in a strict sense, especially in the case of fiction. Did Ulysses ever say: "Son of Atreus, what a word hath escaped the barrier of thy teeth!"[10], i.e., was he correctly and directly quoted? That depends, of course, on whether you consider The Iliad to be a true story or not, and whether you think that the author was there when it actually happened and could make the correct report. I think we can assume this was not an exact quote, unless, of course, Ulysses was a speaker of English.

2.4.2   Towards a new model

The preceding section problematized current terminology and associated theoretical approaches, and several questions have been raised but not answered. This section pinpoints these problematic aspects and sets the theoretical framework for the analyses of direct and indirect speech to be applied in the empirical part of this thesis.

2.4.2.1   Types of structures

As has been discussed above, two types of structures are traditionally distinguished, direct speech and indirect speech. These types will be included in the analyses of this thesis too and the types are illustrated in the examples below:

---

[10] Homer's "The Iliad" with an English translation by A.T. Murray (1988), London: Harvard University Press, p 179.

Example 2.8

Eva Beckman säger dock att man visst kan bli galen av ambition, och tappa omdömet hur lätt som helst.
'Eva Beckman says, however, that one could certainly go mad from ambition, and easily lose control.'

Example 2.9

- Då kommer man av sig, säger Eva Beckman
'"Then you lose the thread," Eva Beckman says.'

These authentic examples are extracted from an interview with the Swedish TV-journalist Eva Beckman, in a Swedish monthly magazine.[11] Example 2.8 illustrates a case of indirect speech whereas Example 2.9 shows case of direct speech. In the first example, the speech reported is in the form of subordinated clauses, introduced by *att*, 'that'. In the latter example, the direct quote is an independent main clause followed by a framing clause containing a speaker identity (*Eva Beckman*) and a verb of saying (*säger*, 'says'). As was established in a previous section (Section 2.4.1.2), there are also cases of direct quotes without framing clauses. I argue that this type (as illustrated in Example 2.10 below) should neither be regarded as a less developed form of direct speech (with a framing clause), nor as a form of indirect speech, but as an independent form on par with the forms in Example 2.8 and Example 2.9. Example 2.10 is taken from the same interview as mentioned above.

Example 2.10

- Tja, andras pappor födde upp grisar, vår skrev böcker.
'"Well, other people's fathers bred pigs, ours wrote books."'

Leaving free indirect speech aside for a moment, this leads us to a taxonomy consisting of three types of speech. The type in Example 2.8 will be referred to as 'indirect speech'. This is in line with traditional terminology. In order to

---

[11] "Jag har en jäkla talang att tassa runt stora egon", *Månadsjournalen, No 5, 1998.* In Swedish, direct quotes are often marked by means of an indented line plus a hyphen, rather than by quotation marks. This will be further described in section 3.3.4.

count as a case of indirect speech, the speech reported is always introduced by a framing clause that may (but need not) include subordinating conjunctions as *att*, 'that' or *om*, 'if', and the clause containing the speech report is a subordinated clause. In order to avoid the confusion that sometimes arises in using the term 'direct speech', (i.e., if it only and by default accounts for direct quotes plus a framing clause, or if it is of no importance whether the quote is preceded/followed by a framing clause), I will hereafter refer to cases like Example 2.9 as 'framed direct speech', whereas the type in Example 2.10 is a case of 'free direct speech'. These two latter types have in common that the mere reported utterance – the quote – consists of an independent main clause that has not been attuned syntactically or deictically to agree with the immediately preceding context (in contrast to indirect speech).[12] The difference between the two types is, as has already been established, that in the case of framed direct speech the quote is framed by a clause containing a lexical speaker marker and/or speech act verb, whereas in the other case, free direct speech, there is no such clause present.[13]

Note in this context that I, by 'indirect speech', 'framed direct speech' and 'free direct speech', primarily refer to the syntactic-structural and deictic properties of these forms, and avoid involving functional aspects as if it is an exact reproduction of earlier speech or not. Moreover, I claim that all three structures are present in written as well as in spoken discourse, and that they can function well in both contexts. I particularly wish to focus on and upgrade the status of free direct speech.

As described in Section 2.4.1.2, free indirect speech has attracted a considerable amount of attention, especially in literary contexts. Furthermore, two main opinions exist, the first one viewing the form as a literary form that is

---

[12] In those cases "quote" are used henceforth, they refer to the *mere speech reported*. No consideration is taken to the presence of a framing clause.

[13] The reason for choosing "free" instead of "unframed" for this latter type, is because "unframed" would be a misleading term since a quote may be "framed" by other means than by a clause with a speaker identity and/or speech act verb. For example, by depicting a certain voice when using free direct speech, the reporter may provide useful information on speaker identity (see e.g. section 2.4.3.1).

almost nonexistent in spoken discourse (e.g., Banfield 1993), and the contrasting one (represented by, e.g., Haberland (1986)) demonstrating that these forms can be found in oral contexts as well. Since Haberland showed this empirically, I see no reason not to include this category in a taxonomy of forms of speech. Particularly interesting, is that his findings were made on spoken Danish, a language closely related to Swedish. Turning Example 2.10 into a free indirect form, it would look like the following:

Example 2.11
Tja, andras pappor födde upp grisar, deras skrev böcker.
'Well, other people's fathers bred pigs, theirs wrote books.'

Notably, the quote indicating dash sign (quotation marks) is gone and the pronoun *vår* ('our') is changed into *deras* ('their') showing characteristics typical of indirect speech. However, the word order is similar to direct speech, and *tja* ('well') would probably have been excluded if it was a purely indirect form.[14]

Although wanting to regard free indirect speech as a form on the same level as the other three forms, I will treat it separately and more in parenthesis than the other forms. This is due to several reasons. One reason is the fact that this thesis concentrates on the report of *speech*. In many cases of free indirect speech it is impossible to decide if it is speech or thoughts that are reported. Another reason is that it is often impossible to decide if we are actually dealing with free indirect speech or not (i.e., the borders can be very vague) and to whom, narrator or protagonist, the free indirect speech is to be ascribed. This makes it more difficult to analyze, and to make, for instance, quantitative analyses comparable to those of indirect, framed and free direct speech. Nevertheless, it is interesting to investigate if this structure is used by children, and if so when it emerges and in what contexts. Is this structure, as many

---

[14] Most often when free indirect speech is described, (especially when it is regarded as a literary form), there is no framing clause present. Note, however, the example from Teleman *et al* (1999:

literary scientists seem to think, a complex stylistic option reserved for written literary works and skilful writers, or is it easily accessible and used in casual speech, even by young children?

2.4.2.2   'Report' revisited

In the previous section I discussed different forms or structures that are available when representing speech, i.e., indirect speech, free indirect speech, framed direct speech and free direct speech. Certain theorists want to assign other terms to these forms, and as has been discussed earlier, indirect speech as well as direct speech each have individually been referred to as reported speech. Apart from the inconvenient aspect of different scholars using different terms, I see a particular problematic point concerning *form* vs *function*. The definition from *The Merriam Webster Dictionary* above indicated that the act of reporting includes a related act where "a message is carried". Indeed, this is what we can use the different forms to accomplish, i.e., we can use the indirect *form* "Robert said that it was hot in there" to relate (which is a *function*) to someone what a certain boy said at a certain time. Thus, in my view, there is a risk involved in naming one category 'indirect speech' (like "Robert said that...") and a contrasting category *on the same level* 'reported speech' (like "Robert said, 'it's hot...'"), (or using 'direct speech' in contrast to 'reported speech'), in the same taxonomy. In this sense I prefer the more common way of using the term 'reported speech', namely as a cover term (and as an indication of what function the forms have) for the different forms of speech.

To regard reported speech as a cover term involves an assumption that the word 'report' itself does not imply the meaning of 'exact reproduction'. This is due to the fact that at least one of the forms cannot be said to be an exact reproduction, namely the indirect form of speech. In the indirect report, the speech has undergone at least syntactical changes and has thereby been transformed. Although I strive to take away the exact reproduction component

---

848) where a framing clause *is* present: *Kunde han inte få gå hem snart, undrade han*, "'Couldn't

in the meaning of 'report', I will need to comment upon the dispute and issue concerning direct quotes and their possible status as exact reproductions of earlier utterances. As was said before, Bakhtin (1981, 1986), Voloshinov (1973) and Tannen (1986, 1989) among others, have argued that it is by no means possible exactly to reproduce what another person has said. The constructed dialogue (Tannen's term) will always occur in a new context, and thereby lose many important resemblances to the original utterance. Clark & Gerrig (1990) claim that parts of the inclination to see direct quotes speech as verbatim reproductions of an earlier utterance,[15] can be explained by a concentration only on the *surface structure* of the *sentence* reported (Partee 1973). It is a fact that actual speech contains hesitations, repairs and pauses, and verbatim reproduction should thus assume that such properties of the original utterance will be correctly and exactly reproduced. That is a difficult task to perform, not least because of the restricted memory capacity of human beings (Hjelmquist & Gidlund 1985; Lehrer 1989). Furthermore, as an exact reproducer, you cannot start to hesitate or make pauses yourself as you will change the utterance quoted. As Clark & Gerrig (1990) note, it is contradictory for speakers to be committed to verbatim reproduction and yet make appropriate repairs. Clark & Gerrig think that the written language bias (Linell 1982) is an important distractor here, hesitation phenomena being rare in written language. However, not only will the surface structure often be affected in the translation between an original utterance and a direct quotation, something also happens with the speech acts:

> In the case of direct speech, we would therefore expect the same linguistic form as, or a linguistic form similar to the one which carried the original speech act, that is, often a full sentence. The propositional content and the illocutionary force of the report are the same as the model's; the difference is that the illocutionary force of the model is only indicated or *displayed* in the report, not performed or *enforced*. ... It is a question of repeating the original (model) utterance without repeating the original (model) speech act. (Haberland 1986: 220)

---

he go home soon", he wondered.'

[15] Cf., "[d]irect speech involves reproducing or mimicking the speech of the reported speaker" (Li 1986:40); the speaker must "reproduce exactly the words of the speaking subject, that he make a copy, so as to render them in substance and form" (Lips 1926:34).

This aspect of displaying, or depicting, will be examined in more detail in Section 2.4.3.

As has been repeatedly shown, a transformation process is going on also when the conditions of modalities remain constant (like in the case of Lisa using the speech mode of expression to relate what Robert had said (also using the speech mode of expression)). An even more obvious translation process is the one wherein a switch of modality and/or mode of expression is involved, like when speech is reported in writing. It is important to keep in mind that indirect and direct forms of speech are used to express (and/or report) *speech*.[16] As was discussed in Section 2.2, certain constraints follow the use of the speech mode of expression, but the fact is that often speech reports and expressions are presented in the written modality and that modality is associated with other constraints. The speech and the conditions tied to the oral (face-to-face) situation are in some sense primary and the norm. Dialogue exchanges in novels are meant to resemble real spoken exchanges, for instance, the examples of indirect, framed and free direct speech I gave in Section 2.4.2.1, were taken from a face-to-face interview reported in a (written) magazine (interestingly, my examples are written reports of a written report of a spoken event). An important issue, then, is what happens in this transformation from the speech mode to the written mode. Indeed, many of the analyses presented in this thesis rest on the assumption that the written mode can take care of speech uttered in oral face-to-face conditions. For instance, I will use written transcripts as an aid for analysis and give examples of authentic spoken instances of indirect and direct forms of speech in the written mode. I will discuss the possibilities and limitations associated with the representation of speech in writing in greater detail in Section 2.4.3.3.

I have already pointed to some facts that make it difficult to make a verbatim oral reproduction of another speaker's speech, and it is by definition

---

[16] Certainly, indirect, direct and free direct structures can also be used to express thoughts (this is especially true for free indirect speech as was discussed above), and they can be used to quote written discourse in writing. However, the primary concern of this thesis is to investigate the possibility of using these structures to express speech.

61

impossible to make a verbatim reproduction of speech in the written mode. Certainly, there are conventional ways of marking off speech (like quotation marks, dashes, indentations) and orthography provides certain options to reflect prosodic aspects (*italics* can be used to show that an item was prosodically prominent, a question mark can be, but is certainly not always, used to indicate an intonational contour associated with a question), but we are still playing with written representations of speech and not speech proper.

In conclusion, I support those who criticize the idea that direct quotations are exact reproductions of earlier speech, and argue that there is always an ongoing transformational process, whether the mode of expression remains the same (like speech to speech) or not (like speech quoted in writing). However, it must be pointed out that it is a matter of degree how faithful the reporter should be to form and content when using indirect speech in contrast to framed and free direct speech. It is by no means possible to reproduce the form, (but only content), in indirect speech, while in direct quotations you might aim at reproducing both the form (or at least something similar to the original form) and content.

2.4.2.3 Speech reporting and speech projection

The preceding section dealt with direct quotations where some kind of original utterance could be presumed, and what the relation between the original utterance and the speech report looked like was discussed. Tannen seems more or less to want to cut the bonds between the two speech events, since the move from "information uttered by someone in a certain situation" to the new utterance in a new situation, is an "active conversational move that *fundamentally* [italics added] transforms the nature of the utterance" (1986: 311). The dialogue, then, is *constructed* and Tannen's choice of term highlights her view on the two speech events as more or less unrelated. It is also the case that she juxtaposes such cases with fiction: "direct quotation in conversation is constructed dialogue, just as surely as is the dialogue created by fiction writers and playwrights" (1986: 311).

I agree with Tannen, however, our opinions diverge on certain points. First, I believe that there still exists an important relationship between the original utterance and the new utterance. In many cases the new utterance can be seen as a *re*-construction rather than as a construction.[17] Secondly, I argue that there are differences between the type of report where an actual source utterance can be assumed, and the cases where no original utterance that can be referred back to exists, (or at least when source utterances are only imagined and invented), like in fiction.

The model I develop here, will cover *(re)constructions of speech which is distinct from speaker time and/or place.* I choose '(re)construction' because it deals with different degrees of invention and creation. In the case of someone attempting to repeat an earlier utterance as carefully as possible, he is reconstructing an utterance to a higher degree than the person who is obviously making up a story containing dialogue, a case which is more towards the "constructed" end of the scale. I include 'distinct from speaker time and/or place' to point out that the current speech event is (in some sense) distinct from the one that is referred to (i.e., the one performed by the original speaker). In the example of Lisa, Tanya and Robert (Example 2.7 on page 50), the speech event of Lisa retelling to Tanya what Robert said is distinct from the original speech event, definitely in time, but also possibly in place. In other cases, the distinct time and place are more "metaphorical". Little children playing with their dolls and letting them speak, are projecting speech onto the dolls who belong to a "make-believe world" (with its own origo and truth conditions) distinct from the real world (to which the child belongs). Consequently, what

---

[17] Also questionable is if 'construction' is a successful term as seemingly restricted to the uses of direct quotations. In a sense, all uses of speech are constructed in that *all* language activities are creative and constructive. However, it should be noted that Tannen talks about constructed *dialogue* rather than constructed speech. She writes that by "dialogue" she does not mean "the exchange of turns that is of central concern to conversation analysts, but the polyphonic nature of all utterance, of every word". Inspired by Bakhtin (1986) she continues that this "polyphony derives from the multiple resonances of the people, contexts, and genres with which the utterance or word has been associated" (1989: 99).

the model covers are cases where two speakers[18] and/or speech events (in some sense) can always be discerned.
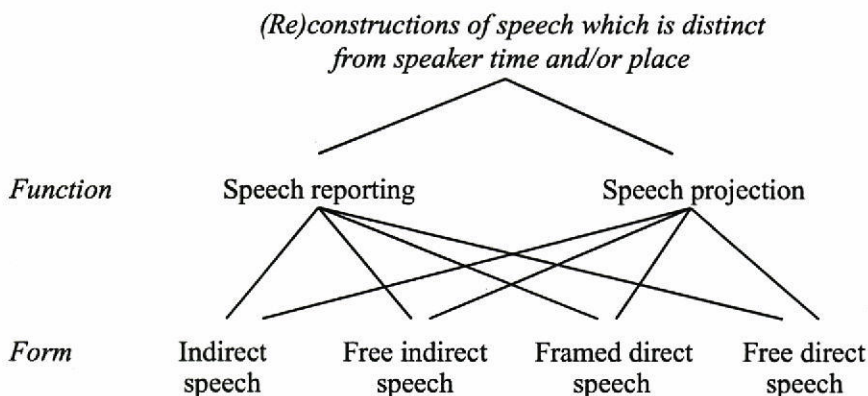
(Re)constructions of speech which is distinct
from speaker time and/or place

Function	Speech reporting	Speech projection

Form	Indirect speech	Free indirect speech	Framed direct speech	Free direct speech

Figure 2-1: (Re)constructions of speech which is distinct from speaker time and/or place; types of speech (functions), types of speech structures (forms), and their possible combinations.

As is illustrated in Figure 2-1 above, two major types of functions are distinguished, *speech reporting* and *speech projection*. The first is about speech that is passing on information contained in earlier actual utterances (to this category belongs what many researchers regard as 'reported speech', but I have chosen to avoid that term due to the problems of terminological confusion (cf. Sections 2.4.1.3 and 2.4.2.2 above)). The second type, speech projection, refers to cases where no original utterance exists or when the original utterance could only be seen as generalized script. Either type can appear in the different structural forms that include *indirect speech, free indirect speech, framed direct speech* and *free direct speech*.

Speech reporting and speech projection can thus be realized by different forms, of which at least two types (framed direct and indirect speech) include a lexical framing clause including a verb of saying. Thus, the terms speech

---

[18] These two speakers may well be the same person. See further the example of free direct speech in Table 2:1 on page 66 below.

reporting and speech projection in these cases refer to more than the mere speech projected or reported (the quoted or described speech). The reason for using these terms in this "less restricted" sense, is that the speech preceding (or succeeding) clause plays a significant role in the act of reporting and/or projecting speech. As will be demonstrated later on, free direct speech may contain as much information as is provided in framed direct speech, yet the information is differently distributed or packaged in the two forms. Consequently, excluding the lexical framing clauses would make comparisons more difficult between framed and free direct speech. Despite this fact, it is indeed of interest in certain types of analyses to distinguish quotes from lexical frames. In these particular cases the separation will be clearly announced.

Imagine an older man, Max, sitting on a bus, when a little girl suddenly enters the bus and points to Max (who has a great white beard):

Example 2.12
"Hey, you look like Santa! You must be tired having traveled by bus all the way from the North Pole!?"
"Yes, it's a looong trip."

The unmarked case, that is, the most common way of constructing and using utterances in ordinary conversation, is the form of free direct speech. It is not customary to start every new utterance with *I say*, for instance, "I say, 'Yes, it's a looong trip'" (although this option may be used for certain purposes). As a matter of fact, that is superfluous and violates Grice's maxim of quantity (Grice 1975) since it is perfectly clear in the actual situation who is uttering the speech.

Turning to the possible case when the speech event in Example 2.12 is reported at a later time, we will refer to it as speech reporting. Typical of this category is the property of *referring back* as there is an *actual source utterance*. Reporting an utterance presupposes that the utterance reported was originally uttered on a previous occasion, and the reporting speaker and the listener are aware of that. Otherwise (i.e. if there is no original utterance and/or speech

event) the listener should be informed about it (cf. speech projection below). Thus, what the reporting speaker is doing, is making a rendition of the earlier speech event. If this rendering, or reporting, is done optimally, the listener will receive information about the original speech (content) as well as about who the original speaker was. (For a more detailed discussion on the process of rendering (demonstration), see 2.4.3.)

In the table below examples of indirect, framed direct and free direct speech are given.

Form Function	Indirect speech	Framed direct speech	Free direct speech
Speech reporting	*In the evening Max is telling his wife what a little girl said to him on the bus home* She said that I looked like Santa.	*Max continues the story* And then she said, "You must be tired, having traveled by bus all the way from the North Pole!?"	*Max reports what he replied to the little girl* "Yes, it's a looong trip."

Table 2:1: Examples of speech reporting; forms of indirect, framed direct, and free direct speech respectively.

These reports refer back to utterances uttered earlier. In this particular case, Max is the reporter and the little girl on the bus the reported speaker (except for the example of free direct speech where Max is reporting what he said himself), the speech reports are found in Table 2:1, whereas the original utterances were the ones uttered on the bus (i.e., the ones in Example 2.12). Furthermore, Max's wife is the receiver, and if they both strive to cooperate (Grice 1975; Allwood 1976), Max is not making up that there existed original utterances nor is he trying to fool his wife. Further, a speaker of English knows that the indirect form cannot be an exact report and that it allows for an amount of freedom in the interpretation, and there is also an awareness of what is valid for

framed direct and free direct reports: "if one wants to quote directly, one has to use the form of direct report, but not every case of direct speech is a direct quotation" (Haberland 1986: 225). A typical context of use of speech reporting is in what we named *personal narratives* in Section 2.3.1, i.e., accounts where personal experiences are expressed. Indeed, Max may have weaved in the reporting of the dialogue in Example 2.12 in a narrative structure as below:

Example 2.13

This little girl, she was around ten years old I guess, sat down and she suddenly said to me that I looked like Santa. And then she said,
"You must be tired, having traveled by bus all the way from the North Pole!?"
"Yes, it's a looong trip."
I started to laugh, surprised by her frankness. And then she too burst out into laughter and we continued laughing for a long time. As a matter of fact, all the way to the next bus stop.

It would also be possible to use the free indirect form. Imagine again the scene with the little girl and Max on the bus, and that the bus driver overheard the conversation. Later on the same day, the bus driver relates to his colleagues what he had heard:

Example 2.14

He looked like Santa. He must have been tired having traveled by bus all the way from the North Pole. Yes, it had been a looong trip.

As is typical for free indirect speech, it is not necessarily clear who the speakers reported are.

The term speech projection is chosen to stress the agency and control characterizing the user of the forms of speech in this category.[19] Most importantly, there is *no original utterance* that is referred back to, or at least no real one. The typical context where speech projection is used is in what was

---

[19] The term projection here is not used in the Freudian psychological sense (i.e., a defense mechanism by which people protect themselves from awareness of their own undesirable traits by attributing those traits excessively to others).

characterized as *fictional narratives* in Section 2.3.1, i.e., in contexts such as novels/fantasy stories and different game and play contexts like make-believe. The components of speech projection are typically three: *a projector, a projected speaker,* and *speech projected.*

Form Function	Indirect speech	Framed direct speech	Free direct speech
Speech projection Ex 1:	*Daddy is making up* *a bed-time story* And the tired little dog said that he very much wanted to go to sleep.	*Daddy is making up* *a bed-time story* The doggie yelps, "I have to go to sleep now so I can play with my friends tomorrow!"	*Daddy is making up* *a bed-time story* "Miaow! I also want to play tomorrow!"
Ex 2:	*Two little boys are* *playing with their* *dolls* And my doll says that he would like to visit yours.	*Two little boys are* *playing with their* *dolls* And then my doll says to yours, "Sure, come in!"	*Two little boys are* *playing with their* *dolls and one of the* *boys gives speech to* *his doll* "You have a really nice house!"

Table 2:2: Examples of speech projection: Forms of indirect, direct, and free direct speech respectively.

The projector (e.g. the writer of a book, the father making up a bed-time story, or the boys playing with their "speaking" dolls) has great freedom and control over who will say what. The projector (like the writer or the father) projects speech onto the protagonists (who become the projected speakers), and by projecting speech, the boys (the projectors) can constitute their game and frame of fiction. In a sense, the projector plays two roles at the same time, the projector role as well as the role of the projected speaker. By using framed direct and especially free direct speech these two roles merge into one; the

projector role is then downgraded, the distance decreased and the narration made more vivid. Similarly, the use of indirect speech upgrades and emphasizes the projector role and increases the distance between the projector/narrator and the projected speaker/protagonist. It is also possible, in analogy with Example 2.14, to use a free indirect form for this type of speech. It has been shown that this is a frequent stylistic device especially in literary texts and, interestingly, this particular form provides the projector/ narrator/author with great control and power over the (projected) speakers as well as the readers. The use of the form can bring about a sense of togetherness between the protagonist and the reader, but, at the same time, as a reader, you seldom know if the speaker is the narrator or the protagonist. In this way the narrator/author can freely play around with parameters of distance and perspective.

The model I have presented here consists of two major categories, for which I have provided and described typical characteristics, prototypical speech, and contextual examples. Admittedly, the borders are not always as clear-cut as might seem to be the case. The border between speech reporting and speech projection should not be regarded as categorical, rather gradual. For instance, it is sometimes impossible to know if an original utterance exists, the phenomenon of pseudoquotation mentioned on page 54 is one variant of this. Haberland (1986) mentions another type, 'hypothetical direct speech', where the speaker reports something which some other person might have said, but which the person actually did not say. Another tricky case is when an utterance in fiction refers back to an earlier utterance in the same fictional frame. I would like to conclude that the crucial point seems to be the speaker and the receiver considering themselves as being in a fictional frame of situation or not, and them being aware of the premises that are connected to the different

situations.[20] Indeed, incorporation, and awareness of these pragmatic aspects, is an important part of the child's language development.

2.4.3 Structuring of information

2.4.3.1 Quotations as demonstrations

Clark & Gerrig (1990) present a theory in which they view direct quotations as demonstrations rather than as examples of description. Demonstrations and descriptions differ in two main ways. Demonstrations are *nonserious actions* rather than serious, they use the example of a person demonstrating a limp to illustrate that this person is not actually or really limping. Furthermore, in demonstration, a referent is *depicted* rather than described, although only selected aspects of the referent are depicted. Clark & Gerrig argue that quotations have all the properties of genuine demonstrations, they are nonserious and selective depictions. Thus, what Clark & Gerrig describe as the heart of their proposal is that quotations, like demonstrations, depict rather than describe (1990:769):

> ... quotations are demonstrations that are component parts of language use. The prototypical quotation is a demonstration of what a person did in saying something. So when Alice quotes George, she may depict the sentence he uttered. She can also depict his emotional state (excitement, fear, shyness), his accent (Brooklyn, Irish, Scots), his voice (raspy, nasal, whiny), and even the nonlinguistic actions that accompanied his speech (gestures, frown, head angle). Furthermore, she can depict nonlinguistic events by themselves.

In order to make clear the demonstration to the recipient, the demonstrater needs to pick out a certain speech event and distinguish it from other possible events. For this to work, Clark and Gerrig (1990: 768) assume four aspects which the demonstrater intends the recipient to recognize. First and foremost there are DEPICTIVE ASPECTS, which are intended to depict characteristic aspects of the referent and to distinguish the referent from other possible

---

[20] Two youngsters engaged in a lively discussion about some recent event, are probably aware that a direct quotation may be a pseudoquotation. That is a natural part of the interaction and will

referents. Yet another includes SUPPORTIVE ASPECTS rather than depictive ones, and are necessary as support in the performance of the depictive aspects. Imagine again the situation where Lisa is reporting to Tanya what Robert said yesterday when he entered the classroom (Example 2.7, page 50). In making a direct quotation, Lisa may depict the sentence he uttered (*It's hot in here*), his accent, and his panting voice. These are depictive aspects. If she had chosen to use a male pitch, that would also be depictive. However, she uses her own pitch which is then a supportive aspect of the depiction. Lisa may also add commentaries on what she is demonstrating and these are then ANNOTATIVE ASPECTS. These are not intended to be recognized as a part of the demonstration itself. Thus, Lisa may giggle while making the demonstration in order to express to Tanya that she thought Robert reacted strangely to the actual low temperature of the room, (i.e., Robert himself did not giggle in the original situation). Finally and fourthly, Clark & Gerrig assume INCIDENTAL ASPECTS, i.e., incidental to the demonstrator's purpose in demonstrating. These refer to aspects that the demonstrator has no specific intentions about and they are what is left over once the depictive, supportive, and annotative aspects have been selected.

When Lisa demonstrates this conversation, it is not sufficient that Lisa herself knows what she has chosen to depict, she also wants Tanya to recognize that she is depicting, and which aspects of her demonstration are depictive, supportive, and annotative. According to Clark & Gerrig three principles are then in operation (1990: 774):

DECOUPLING PRINCIPLE: Speakers intend their addressees to recognize different aspects of their quotations as depictive, supportive, and annotative.

PARTIALITY PRINCIPLE: Speakers intend their addressees to take the depictive aspects to be the quotation proper, the primary point of their quotation.

SELECTIVITY PRINCIPLE: Speakers intend their quotations to depict only selective aspects of the referents under a broad description.

---

in most cases not lead to a break down.

Thus, by the first principle, Lisa intends Tanya to see the division of depictive, supportive, and annotative aspects. By the second, Lisa intends to take only the sentence, the accent and the panting voice to be depictive; and by the third, she intends to depict only these three aspects of what Robert originally uttered.

What, then, is possible to select and depict in a report of a speech event? Clark & Gerrig point out that the possibilities are as many as there are possible aspects of a speech event in general. They provide a set of examples (1990:775):

(i)   DELIVERY: voice pitch (male, female, child), voice age (adult, child, oldster), voice quality (raspy, nasal, slurred), speech defects (lisp, stutter), emotional state (anger, sarcasm, excitement), accompanying gestures (pointing, smiling, frowning)

(ii) LANGUAGE: langauge proper (English, Dutch, Japanese), dialect (British English, Bostonian English), register (formal, informal)

(iii) LINGUISTIC ACTS: illocutionary act (question, request, promise), propositional expression (the proposition expressed), locutionary act (the sentence uttered), utterance act (the utterance issued with repairs, etc.)

Besides choosing aspects of delivery and language in order to depict referents, these markings contribute to making a narrative vivid (e.g. Tannen 1986). Moreover, the use of (changes in) tone of voice helps the listener to keep track of changes in perspective (Nordqvist 1998b). Concerning choice of language, dialect and/or register in the quotation, Clark & Gerrig argue that the expected variant for a reporting speaker is to use their current one. To choose some other language, dialect, or register is to mark it as depictive, supportive, or annotative. Moreover, Clark & Gerrig point to the fact that people perform several linguistic acts simultaneously in speech events. Consequently, it is possible to depict all five linguistic acts described above in a quotation, however, it seems as though most quotations depict illocutionary acts, including the propositions expressed, and treat the other acts as supportive or incidental.

Clark & Gerrig concentrate on direct quoting and only in passing mention indirect speech which is considered to be description rather than depiction. Furthermore, rather than speech projections, the focus is on speech reporting in the sense I described it in the preceding section. Thus, Clark & Gerrig identify two domains, *the source domain* (including an original utterance and a reported speaker), and *the current domain* (including a reporting speaker and the reported speech). However, it is easy to conceptualize a case where the narrator/author in/of fiction (i.e. a projector) is depicting aspects of the referent and his speech (e.g. a protagonist). However crucial it is in the case of speech reporting that the intended referent is distinguished from other possible referents, it is equally crucial to distinguish protagonists, (or narrator from the protagonist), from each other in a fantasy story. Thus, in my opinion, the demonstration theory works well both in the frames of speech reporting and speech projections.

Pre-school children make use of their voice in order to perspectivize and conform to an adopted identity in play (see further 2.4.4.1 below), and this points to the fact that young children are, at least in some sense, already conscious of different aspects of depiction. Yet, for obvious reasons, the child has considerably less experience of speech and speakers than the adult has, and the child has a smaller repertoire of aspects of delivery and language to choose from when, for instance, projecting speech to a character in play. In addition, in order to control the decoupling, partiality, and selectivity principles, certain pragmatic considerations are required. As previously discussed, it takes some time for pragmatic skills to develop and there is reason to believe that awareness about and control over these principles increase over time.

In this section, we have been primarily concerned with depictions in relation to spoken quotations, but Clark & Gerrig's theory applies to written quotations as well. In the following section we will deal specifically with speech reported and projected in the oral mode, whereas Section 2.4.3.3 is devoted to the issue of speech reporting and projecting in writing.

2.4.3.2 Depiction, description and packaging

For reference, I repeat the examples of framed direct speech (Example 2.1), indirect speech (Example 2.2), and free direct speech (Example 2.7), that were given earlier in this chapter.

Example 2.1

Robert said, "It's hot in here".

Example 2.2

Robert said that it was hot in there.

Example 2.7

"It's hot in here".

Setting aside the fact that Example 2.1, Example 2.2, and Example 2.7 refer to oral examples that would naturally include components of intonation and prosody, and that they are uttered in a certain situational context, it is interesting to investigate what linguistic acts (in Clark & Gerrig's sense) can be inferred from only the syntax and wording in these three examples. What unite all three cases are the similar propositional content and illocutionary act expressed, although the framings look different. But Example 2.1 and Example 2.7 also have characteristics in common that distinguish them from Example 2.2. In contrast to the case of indirect speech, the examples of framed direct and free direct speech carry out the same locutionary and utterance act. Yet another difference is what Clark & Gerrig point out, that Example 2.1 and Example 2.7 exemplify *depictions* while Example 2.2 is a *description* of an earlier speech event. Similarly, Hickmann (1993: 65) points to the following differences: "Direct quotations reproduce the quoted speech event as a whole, presenting not only what was said, but also how it was said, [i.e. depictions, *my comment*], whereas indirect quotations typically focus on the propositional content of the original message and incorporate other aspects of the original utterance into the narrative message [i.e. descriptions]".

74

The fact that Example 2.1 and Example 2.2 are lexically framed, but Example 2.7 not, is important. The first two cases contain explicit information that is not present in the last: "The framing clauses minimally announce that what follows is speech originating from some other situation and provide an antecedent ground in the narrative in relation to which the deictics of the framed clause can be interpreted" (Hickmann 1993: 64-65). However, as was discussed in Section 2.2.2 above, packaging can occur on different linguistic levels (lexically, morphologically, syntactically and on a discourse level). If we now permit ourselves to enlarge the context window, it might well be the case that it will become clear, even for Example 2.7, who is the reported/projected speaker: "Some aspects of this antecedent ground need not be specified within the metalinguistic frames themselves, given that utterances reporting speech are often embedded in a larger narrative discourse" (Hickmann 1993: 65). Eriksson (1994) investigated spontaneous dialogues between Swedish adolescents and he reports that each utterance in the form of free direct speech (and who was uttering it) was understood by the listener by virtue of a larger exchange of utterances in which it was embedded. Hickmann argues that "metalinguistic frames, often in conjunction with other narrative clauses, create explicit *boundaries* between the reported and the reporting messages" (1993: 65). Voloshinov stresses the importance of these boundaries: "Between the reported speech and the reporting context, dynamic relations of high complexity and tension are in force. A failure to take these into account makes it impossible to understand any form of reported speech" (1973:119). In addition to metalinguistic frames containing speech act verbs, there are other possible lexical and deictical means of signaling shifts to speech quoting. Sometimes the only clue for the listener is, for instance, a shift from third-person pronouns to first- or second-person pronouns, or from declarative to interrogative sentences.[21]

---

[21] Cf., "And then she met her friend outside the school building. Would you like to play with me?"

75

"Metalinguistic frames" and "reporting context" need not necessarily refer to strictly verbal contexts. Indeed, to treat Example 2.1, Example 2.2 and Example 2.7 without taking into account the fact that spoken utterances are multimodal is to underestimate their natural power of expression. As Clark & Gerrig have illustrated, we can choose several aspects to depict, using modes of voice and gestures. Continuing the discussion initiated above, intonation, for instance, can be used to clarify speaker and speech boundaries (Kvavik 1986; Goodell & Sachs 1992; Nordberg 1992), but also gestures: "An ever so slight change in the tone of voice, a body movement, or simply short pauses can make clear who's speaking" (Haberland 1986: 247).[22] Especially in the cases of free direct speech, the option of packaging information simultaneously becomes important. While aspects of delivery (like voice modifications) in framed direct speech sometimes can serve as a (redundant) support to what is already expressed by words in the framing clause, in free direct speech the information depicted and delivered simultaneously by the mere quote, can sometimes act as the only clue to speech and speaker. One conclusion I would draw already at this point, is that the forms of Example 2.1, Example 2.2 and Example 2.7 in an oral context, although structurally differing from each other, do not *necessarily* differ in how much linguistic information they carry. The crucial point is that the information is structured and packaged differently, and that they are used for different purposes.

As a describer of an event, you can choose how you want to describe it. This means that there is space for subjective interpretations. Hence, in seeing indirect speech as a descriptive mode of expression, you can add, for instance, your own attitude to the speech event (and/or speaker) you report (from). Imagine once again Lisa and Tanya, but this time Lisa is not very fond of Robert since she thinks he always complains about everything. So she makes the following report to Tanya of yesterday's event:

---

[22] In quoting in *sign language*, non-vocal aspects become even more important and effective to display referential or role shift (Padden 1986; Engberg-Pedersen 1993; Emmorey & Reilly 1998).

Example 2.15

As usually, Robert grunted that it was hot in there.

One key to her general attitude is the verb of saying she chooses (in combination with the expression "as is the usual case"). It is not necessarily true that he was actually "grunting" in the original situation, but Tanya chooses to interpret/display it as such. As was discussed elsewhere in this thesis, one tradition of theorists claims that you are only free to add attitudes and make a 'de re' analysis when you are using the form of indirect speech and not when using direct quotations (cf. the quote of Coulmas' on page 53). However, this is not a convincing claim, since Lisa could utter Example 2.7 using a sarcastic and ridiculing voice that does not resemble Robert's original voice at all, but signals only Lisa's attitude to Robert. Since the DECOUPLING, PARTIALITY, and SELECTIVITY PRINCIPLES are in operation, (hopefully) both Lisa and Tanya know that it is Lisa's attitude towards Robert that is depicted, rather than his actual way of speaking (cf. the notion of annotative aspects above). I would like to say that it is not only as a describer of an event you can choose how you want to describe it, but also as a depictor and demonstrator of an event.[23]

2.4.3.3 Representation in writing

Written langauge is monomodal to the extent that only the visual mode is at use. When rendering speech (which is polymodal) in writing, orthographic conventions can be used to represent such things as sentences, words, phonetic segments, and some temporal and intonational information. For instance, quotation marks are used to mark speech boundaries. However, there are no conventional markers to depict tone of voice, pitch, nasality, singing, gestures, or head nods. There does though, seem to be at least some options when depicting aspects of delivery.

---

[23] Caldas-Coulthard (1994: 307) points out that "'quoting' what people say is a dangerous activity. Sayings are transformed through the perspective of a teller, who is an agent in a discursive practice. In this way, social identities and roles are created according to the values of who reports and the institution this person represents".

Example 2.16

'*What did you say?*' Sophie shot up from the chair she had been sitting in. 'You called me Hilde!'[24]

In Example 2.16, the first quote is written in italics to mark that Sophie utters the speech with some kind of emphasis. There are several options that can be used to modify text. Besides italics, boldface, underscoring, uppercase letters and combinations of these can be employed. A skilful writer can most likely use these possibilities to depict quite a few of aspects of delivery.

Different aspects of language, dialect and register might also be depicted and marked (i.e., they are not incidental). Consider the conversational exchange between Bridget Jones and Daniel below.

Example 2.17

'Darling,' said a different drunken voice I recognized.
'Go away, Daniel,' I hissed.
'No. Lemme explain.'
'No.'
'Bridge...I wanna come in.'
Silence. Oh God. Why do I still fancy Daniel so much? [25]

Daniel is drunk, and the author wants to mark and depict this aspect by letting Daniel use a register he would not use otherwise (when not drunk). Nowhere else in the novel does Daniel use words like "lemme" and "wanna".

Although there are various possibilities when depicting certain aspects of the speech event, fewer are available when quoting speech in the written modality. What cannot be depicted and demonstrated then has to be described. In Example 2.17, for instance, "Darling" was described as being uttered in a "drunken voice". Consider also the two following examples:

---

[24] Gaarder, *Sophie's World*, p. 178 (translated English edition).
[25] Fielding, *Bridget Jones's Diary*, p 111.

Example 2.18

I slumped into my seat muttering, 'Shud-urrrrrrrp,' out of the side of my mouth like a humiliated teenager.[26]

Example 2.19

'Now come along, Bridget. I don't want any silliness,' she said in her Genghis-Khan-at-height-of-evil voice.[27]

In Example 2.18 the demonstration in the quote itself ("Shud-urrrrrrrp") is supported by a descriptive verb of saying ("mutter"). In Example 2.19, an excerpt from a telephone dialogue initiated by Bridget's obtrusive mother, Bridget's irritation at her mother (talking in a "Genghis-Khan-at-height-of-evil voice") is, I believe, evident. Thus, verbs of saying not only function as introducers, as established in the previous section, but they also function as descriptives.

The most common verb of saying is, not surprisingly, *say*. Examining factual and factional text (a 5 million word corpus of *The Times* and a 20 million word corpus of the BBC World Service) Caldas-Coulthard (1994) found that *said* was the most frequent type of verb of saying (14,154 instances), followed by *say* (3,634 instances), *told* (1,445 tokens), and *asked* (occurred 1,050 times). All other types of verbs of saying occurred less than 500 times. Tannen (1986) found that *leo* (the Greek word for *say*) was the most frequent type of verb (about 70% of the cases) in the Greek spoken dialogues and written literary texts she investigated. In a comparative American corpus, *say* was also the most common verb, and this verb was used as frequently in the written data as it was in the spoken data. However, there was also evidence that more types of verbs of saying (other than *say*) were used in the American written literary text than in the American spoken dialogues. Page (1988:27) contends that "many writers seek to relieve the monotony of constant 'he-saids' by resorting to elegant variation". In the opening chapter of *David Copperfield*

---

[26] Fielding, *Bridget Jones's Diary*, p 50.
[27] Fielding, *Bridget Jones's Diary*, p 130.

Page, besides instances of *say*, finds speech act verbs like *return*, *ask*, *cry*, *exclaim*, *falter*, *resume*, *repeat*, *reply*, *sobb*, *muse* and *ejaculate*. In more popular fiction, according to Page, protagonists often *gasp*, *moan* and *shriek*, while in sentimental romances they *murmur*, *breathe* and *sigh* their words.[28]

Thus, variations of verbs of saying in written texts are used for stylistic reasons and for the marking of genre.[29] But most importantly, they are used for the packaging of information that cannot be otherwise encoded because of lack of modes of expression like voice and gestures. Fónagy (1986) points out that verbs of saying can denote non-verbal activities as well, like *laugh*, *grin*, *thunder* and *shake one's head*, as illustrated below.

Example 2.20
Sophie shook her head. 'No, I've been gone for more than a thousand years.'[30]

However, it may be discussed whether we are really dealing with a verb of saying in Example 2.20, and the question is what relation the non-verbal act has to the speech act. Imagining this as an oral event, it is as likely that Sophie first shakes her head and after that she starts to speak, as it is that she is performing the two acts simultaneously. It is indeed problematic how to treat these kinds of verbs denoting non-verbal activities; and where to draw the line between what is really speech act verbs and what is not, is tricky. There is a similar type of problem in cases where verb of saying-like examples such as the English *go* (Tannen 1986) and *like* (Romaine & Lange 1991) are used. These have been shown to be particularly common in pseudoquotations. In Swedish *bara* (approx. 'just') is frequently used in a similar way, especially among youngsters (example from Eriksson 1994: 136):

---

[28] For detailed semantic analyses and categorizations of different types of verbs of saying/communication, see Allwood (1980), and Caldas-Coulthard (1987).

[29] Note, however, that the use of descriptive verbs of saying also points to a crucial linguistic assumption about gender relations. When examining English press-texts, Caldas-Coulthard (1994) found that the men *shouted* and *groaned*, while women (and children) *screamed* and *yelled*.

Example 2.21

Ja ba: "gå bort från våran tomt."
'I just, "go away from our garden."'

An example as Example 2.21 is unlikely to occur in (formal) written language, at least today. Several researchers (e.g., Romaine & Lange 1991; Meehan 1991; Eriksson 1995) argue that discourse markers like *like* (Eng.), *go* (Eng.) and, *ba* (Sw.) are becoming grammaticalized, and perhaps it is only a matter of time before these forms show up also in (formal) written texts.

Although free direct speech seems to be more common in spoken dialogue than in written texts (Tannen 1986, Chafe 1982), it also occurs in written media. I have already considered Example 2.20 as such an example, and in Example 2.17 three cases can be found. In Example 2.17, (the conversation between Bridget Jones and Daniel), the speaker identities can be inferred from the fact that only two people are present and that they are exchanging utterances. In addition, two of the projected utterances include the names of the addressees. In cases of framed direct speech, the framing clause may precede as well as follow on a quote. In Example 2.18 the quote '*Shud-urrrrrrrp,*' is introduced by *I ... muttering*, whereas the clause with the speech act verb follows on the quote in Example 2.19. It is an open question whether both types are equally common in written texts, or if either of the two types of construction is more common than the other. There is however an aspect claimed to be unique to written texts, and what Fónagy (1986) calls *enframing*, (the following example is from Tannen 1986: 323):

Example 2.22

"How about," Rhoda suggested, "clearing off the kitchen table so you can have some good old peanut butter and jelly sandwiches?"

In this example, the line of dialogue is interrupted by a clause with a speech act verb, in a way that would be odd in a real spoken dialogue. Not only could

---

[30] Gaarder, *Sophie's World*, p. 157 (translated English edition).

Example 2.22 uttered by a speaker result in the listener loosing the thread, but it also puts cognitive and linguistic demands on the on-line producing speaker.

To conclude, writing is predominantly monomodal and linear, while speech is predominantly polymodal and both linear and simultaneous. This means that the options of expressing simultaneous paralinguistic features in writing are restricted. Consequently, descriptions play a (more) important role in contexts of reporting and projecting of speech in writing than in speech, while in the latter more elaborated multimodal depictions can be used. In one sense, the depictions in the spoken mode provide opportunities to package much information into one utterance. On the other hand, the written mode with its linearity and its frequent descriptions, also allows for effective distribution of information. The discussion of verbs of saying above provides ample evidence of such packaging in that a single verb can, besides marking an utterance as reported or projected, simultaneously provide information on, for instance, voice, gestures and attitudes.

2.4.3.4 Multifunctionality

Berman (1996) discusses "functionality" in terms of how obligatory a particular option is for expressing a given form-meaning relationship, and the availability of other forms for the same purpose. Indeed, direct and indirect forms of speech are *options available* when referring to a speech event, but they are not *obligatory* or the *only* forms available. Consider Example 2.23 and Example 2.24, which illustrate two different ways of narrating/rendering a certain event.

Example 2.23

"Linneaaa! Where aaaare you?"

Example 2.24

Julia called for her friend Linnea.

On the assumption that these examples were produced in oral narratives and the voice quality revealed that "Julia" was the intended speaker/caller in Example 2.23, these two realizations, or structures, serve a similar purpose, namely that of telling that Julia called for her friend. This shows that direct and indirect forms of speech are not the only means available to report on a past speech event. A closer examination of Example 2.23 and Example 2.24 reveals that either of the two structures can be chosen and used for different purposes, or to produce different effects. The example of free direct speech, Example 2.23, can, for example, be chosen by the narrator in order to make the narration more vivid (and thereby foreground certain events), and to decrease the distance between the narrator and the projected speakers, while Example 2.24, by the same token, may be chosen in order to create a sense of distance.

Thus, two different forms may be used to serve more or less the same function, and it is also the case that a form like Example 2.23 works multifunctionally.[31] Figure 2-2 below illustrates how a speech projection in the form of free direct speech can serve several partly overlapping functions and produce different effects simultaneously. Consequently, uttering Example 2.23 may, at the same time as it presents propositional content, provide information about who is the projected speaker (perspectivizing), provide evaluative comments (evaluation), give information on characteristics of the speaker and the speech (depiction/description of referents), contribute to maintaining the listener's interest (vivifying), and specify parameters of distance and role (distance decreasing). [32] In addition to this, a free direct speech utterance like Example 2.23 can, (in the same manner as Example 2.24), be produced in order to move the plot forward (plot advancing). Forms of direct speech are often discussed in relation to evaluative functions, but as has been argued elsewhere in this thesis, speech projections and reports also often serve to move a story-

---

[31] Certainly, Example 2.24 may also work multifunctionally.

[32] Wolf & Hicks (1989) discuss *plurifunctionality* of voices and argue that each voice can, simultaneously, be used, for instance, to portray speech, to describe events, and to offer commentary. In this way, they claim, a narration includes "a network of texts within texts" (1989: 331).

line forward. The underlying assumption in this thesis is that evaluative functions do not exclude narrative functions, and vice versa.[33]

realization of form   *functions*

free direct speech
\+
voice modification

expression of proposition
perspectivizing
evaluation
depiction/description of referents
vivifying
distance decreasing
plot advancing
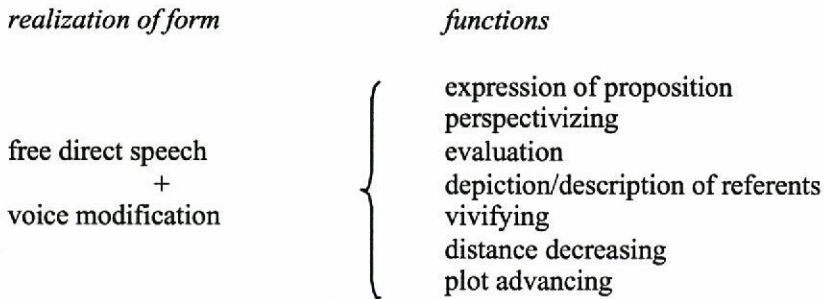
Figure 2-2: Example of multifunctionality in use of free direct speech in combination with change in tone of voice.

Information packaging, form-function relationships and aspects of multi-functionality will be further discussed and empirically examined later in this thesis. In the next section previous developmental studies that concern these kinds of aspects are reviewed.

2.4.4   Previous developmental studies

2.4.4.1   *Children's production*

The review of studies conducted on children's use of direct and indirect forms of speech are here divided according to the two major functional categories introduced above: speech reporting, and speech projections. The overwhelming majority of investigations that have been undertaken deal with children's use of these functional categories in oral discourse, but I will also report on studies that concern developmental aspects of direct and indirect speech in writing.

---

[33] Cf. Tannen (1982: 8): "Direct quotation is a common form of internal evaluation. By putting words in the mouth of the characters, the teller communicates what happened from inside the story. Nonetheless, by deciding what words to put in the character's mouth, the teller is building the story towards the desired point."

Speech reporting

The speech of the little girl Emily was tape recorded over a fifteen-month period between the ages of 21 and 36 months (Nelson 1989). The recordings included Emily's dialogues with her parents before bed, as well as her crib talk to herself when her parents had left her to sleep. Dore's (1989) analyses of Emily's soliloquy crib talk revealed that, from around 23 months (at the same time as she started to produce coherent narrative sequences), she started to reenvoice in her monologues what she had heard in dialogues earlier the same day. She reenvoiced her father's speech, adopting his phrasing and intonation patterns, although she couldn't always separate the speech of her father from her own; "her monologues conflate these two overall processes of imitation and creation. She blends her own 'voice' with that of her father" (1989: 248). However, already a month later she managed to weave together many voices and begins to coordinate more than one genre at a time. An example of this is when she changed her tone of voice when impersonating characters in a book recently read, and used a more neutral voice when narrating other facts of the story. Thus, the analyses of Emily's monologues show that processing of past speech and some degree of awareness of different "voices" are already developed at two years of age.

Studies of infants and their use of speech reporting in spontaneous speech are rare. However, Ely & McCabe (1993) investigated three longitudinal case studies from this point of view (all instances related to book reading or occurring in make-believe story-telling, i.e. speech projections were excluded): Naomi (1;2-4;9) from the Sachs (1983) corpus, and Adam (2;3-4;11) and Sarah (2;3-5;1) from the Brown (1973) corpus. Naomi's first instance of direct speech[34] was found at 24 months, Adam's at 29 months, and Sarah's a month later. There was a preference for reporting immediate and own speech. There were also individual differences between the children regarding the use of direct and indirect forms. Sarah preferred direct speech to indirect (38.1% vs

---

[34] No information is given on what "direct speech" exactly refers to, i.e., if a framing clause is present along the direct quote or not.

23.8%), Adam produced as many direct forms as indirect forms (37.5% in both cases), while Naomi preferred the indirect form to the direct, 42.9% vs 14.3%, (however, it should be noted that Naomi produced only 7 speech reports).[35] In the same study, Ely & McCabe investigated experimentally prompted personal narratives from 96 children at ages 4, 5, 6, 7, 8, and 9 (corpus from Peterson & McCabe (1983)). The results showed that 25% of the 4-year-olds reported past speech while 88% of the 9-year-olds were reporting speech. Direct speech (no distinction is made between framed and free direct speech) was the most common form at all ages but indirect speech increased with age. Moreover, the children at all ages most often quoted their own speech.

Goodell & Sachs (1992) used an experimental design to elicit personal narratives. Eighty subjects in the age groups 4, 6, 8 and adults, were included in this study. Goodell & Sachs coded use of free direct speech as an error [*sic!*] in the retelling of the dialogue heard. They explained it by the fact that none of the children using free direct speech in the study made a change in tone of voice or the like to compensate for lack of a framing clause. A linear age effect was found, with the 4-year-olds using direct speech 82% of the time correctly (i.e. framed direct speech), the 6 year-olds 94% of the time, and the 8-year-olds and the adults 99% of the time. In the case of indirect speech, a U-shaped curve surfaced with a drop in the age group of the 6-year-olds. The cases coded as errors were blends between direct and indirect forms (as in Example 2.3 on page 46) and Goodell & Sachs explain the U-shaped curve with reference to the fact that the 6-year-olds generally used more speech reporting in their narratives and that they attempted a larger variety of constructions than the 4-year-olds. Moreover, "[E]ven though the 4-year-olds used the indirect style 61% of the time, it is not clear that they have really understood that there are two distinct systems. The 6-year olds may be acquiring knowledge about indirect speech, but they may choose the direct style because using it enables

---

[35] Besides direct and indirect speech, instances of narrativized speech were coded for. Narrativized speech represents a summary description of a speech event as in "we were arguing a lot". Thirty-eight percent of Sarah's speech reports were of this type. The corresponding percentages for Adam and Naomi were twenty-five and fourty-three.

them to relay more of the story without heavy processing of the conversationally conveyed meaning. Mimicking is easy, but processing is hard" (1992: 416-417).[36] The typical 6-year-old narrator added one free direct speech utterance after the other without giving the listener any chance to find out which point of view is being taken. The typical adult narrator, in contrast, used indirect speech modes, and framed direct speech only to mark the high points of the discourse.

Hickmann (1993) also used an experimental design where a puppet dialogue (either in the form of a film or in the form of a text read aloud) was to be reported. The subjects included were 4, 7, and 10 years old, and in addition a group of adults. In this study, three main modes of presenting speech were distinguished: The *re-enacting* mode, including a considerable amount of free direct speech, the *reporting* mode containing framed direct and indirect speech, and the *descriptive* mode corresponding to the use of various types of nonexplicit descriptions (such as paraphrases of earlier speech). The tendency was that the 4-year-olds chose the re-enacting or the descriptive mode of narrating, or a mixture of the two modes. The 7- and 10-year-olds used the direct reporting mode (although they sometimes switched to the re-enacting mode in particular situations), whereas the adults either reported in the direct reporting mode or in the indirect mode.[37] In contrast to the two oldest age groups, the 4- and 7-year-olds rarely introduced the dialogue participants appropriately. Furthermore, the 4-year-olds had great problems marking explicit boundaries between narrative and narrated speech. This is explained by the extensive use of free direct speech, but the descriptive mode also caused problems in this respect (cf. the discussion of free indirect speech above).

A general finding of speech reporting, then, is that framed direct and free direct speech are produced at an early age. Not all children reported speech at

---

[36] Note, however, that mimicking is a special case of processing. I will return to this question in section 2.5.

[37] However, note Tannen's (1986) finding that 26% of the quotations found in the narratives of American adults were of the free direct type. Ely *et al* (1995) comment that use of free direct speech might be prevalent in very young children and adults.

the age of 4, however, the form of free direct speech is particularly common at this early stage. As the child gets older and becomes aware of the need to orient the listener and the requirements of a communicative situation, he or she frames the utterances verbally to a larger extent and uses more framed direct speech. At school-age, most children report speech and do it in an "appropriate" way. Indirect speech may also appear at an early age (cf. Sarah, Adam and Naomi as described above), however, Goodell & Sachs (1992) suggest that because 4-year-olds manage to produce this form does not necessarily imply that they possess full control over it. In the Goodell & Sachs study, the two-year-older 6-year-olds seemed to have problems with this particular form, while the 8-year-old school children mastered it. The tendency for indirect speech to appear later and cause more problems, may be due to the greater cognitive demands that are required to transform direct speech into indirect speech. The same trend has been found in Turkish (Özyürek 1996). However, in Özyürek's study, indirect speech appeared much later, at the age of 13. Özyürek explains this with reference to the Turkish construction as being syntactically more complex than the corresponding English one. Common to all the studies of English speech reporting presented above is also that *say* is by far the most common verb of saying in younger age groups (the younger the speaker is, the more common *say* is), while adults use other types as well.[38]

Developmental studies of indirect, framed direct and free direct speech in Swedish are almost non-existent. However, there is one sub-field that has been examined in greater detail, and that is conversations between Swedish adolescents (e.g., Nordberg 1985, 1986, 1992; Kotsinas 1994). These conversations were not experimentally elicited, but naturalistic and spontaneous. The data included a great deal of reporting of personal experiences, and typical to these types of exchanges were pseudoquotations, onomatopoeia, and exclusion of traditional verbs of saying (i.e., a high proportion of free direct speech, and use of framing words like *bara* ('just'), *då*

---

[38] Goodell & Sachs (1992) noted that the 4-year-olds used *say* in about 80% of their direct speech reports. The corresponding number for the adults was 40%.

('then'), and *sen* ('then')). Similarly, Eriksson (1993) investigated spontaneous peer dialogues by Swedish 10- to 15-year-olds and found 394 instances of framed and free direct speech and 33 instances of indirect speech. These proportions reveal that indirect speech is not a preferred means of speech reporting or type of activity in this age group. Out of the 394 instances of direct speech, 65% were of the free direct type, and only 35% of the framed direct type. However, according to Eriksson, the lack of speech framing clauses did not generally cause problems of perspective. The young people used voice modifications to a large extent, and the reported speaker of the utterances could usually be discerned by virtue of their place in the exchange of utterances. Thus, contrary to what the experimentally elicited data of younger English speakers showed, free direct speech was, in this study, shown to be the most frequently employed form, and is even more common than framed direct speech. This suggests that type of context and degree of formality have important effects for choice of form.

A general conclusion from the studies reviewed above is that indirect speech is more rarely used than forms of direct speech in oral speech reporting. Interestingly, direct speech, in contrast to indirect speech, has been shown to be universally distributed (see Section 2.4.1.1). In a study based on 132 interviews with French-speaking Canadians telling about their life, interests, and aspirations, Vincent and Perrin (1999) found that direct quotations were more than six times as common as indirect speech. They distinguished three age groups in their study: young, middle-aged, and old speakers. Their finding was that the older speakers reported speech to a greater extent than the youngest. Moreover, while the subjects in the young group most often used speech reports in order to support and reinforce an argument, the reports by middle-aged speakers typically had an authority function, and for the subjects in the old group a narrative function. Thus, these results suggest that younger speakers not only make use of less speech reporting than older speakers, but also that younger speakers use speech reporting for other purposes.

Speech projections

Andersen (1990) made use of 'controlled improvisation' in order to elicit what children know about the speech appropriate to different roles. Eighteen predominantly white middle-class children aged 4 to 7 took part in three sessions of play-acting with hand-puppets. The three settings consisted of (1) a family or home situation, (2) a doctor's office, and (3) a children's classroom. The results showed that registers tied to the family context seemed to be the easiest, in that every child, irrespective of age, produced some registers tied to specific roles in this context. The doctor session was the favorite of the 5- to 7-year-olds, whereas the youngest children were less comfortable with this. The classroom session was the most difficult for the children, and whereas the youngest children only took the role of a student, the older children additionally adopted the role of the teacher. All children in the study used changes in voice pitch in order to depict the speech of different puppets/roles (e.g., distinguishing male-female, child-adult, doctor-patient). In fact, this was the earliest feature used to mark register. However, there were clear age-related differences in the children's grasp of phrasal lexicon or formulaic speech associated with the different contexts. The oldest children had larger vocabularies and were more skilful in selecting terms appropriate for a given register, although they also often made mistakes in lexical usage.

Although Andersen's study does not focus on use of direct versus indirect forms of speech, or of narrative structure, the results show that young pre-school children to some extent control indirect voicing, to depict voices and to package linguistic and sociolinguistic information into their speech projections. However, the 6- and 7-year-olds do better than the 4-year-olds in this respect, indicating a development of pragmatic and sociolinguistic skills (among others) during these years. Andersen mentions (1990: 80) that 3-year-olds were included in a pilot phase of the study, but these children found the task very difficult. As was mentioned in Section 2.3.1, Wolf and Hicks (1989) studied spontaneous make-believe play in young children and found, not only potentially, but typically, multivoiced narratives from early on. Proportions of

different forms of speech were not calculated in this particular study, however, it was clear that by age 3, separate voices were distinguished in some children's make-believe play narratives, i.e., speech was projected onto the protagonists. The example given from 3-year-old Heather contains character's speech that is imperative in nature, not marked for tense, and free direct in form (e.g., "go back in the house and go to sleep!"). Between the ages of 3 and 6, developmental changes in the marking of separate voices can be seen in the temporal system. The narrative voice might be in past tense, and the dialogue in present tense. Furthermore, the framed direct form shows up. At around age 5, children can also use the narrative voice to present speech indirectly. Thus, the same developmental pattern of forms of speech as is evidenced in reports of actual past speech events, is also evidenced in play.

Uccelli, Hemphill, Pan & Snow (*in press*) found in their study of 5-year-olds' narrative skills that better narrators included a great deal more evaluative aspects than the poor narrators did (see also Section 2.3.2). The personal narratives and fantasy stories of the better narrators included direct and indirect speech, and the narrator's perspective was conveyed in a very clear fashion. The poor narrators of personal experiences exposed very few evaluative components, and rarely included speech reporting. The fantasy performances were not verbally framed and when speech was projected, the animal props were typically held up to indicate that a story character was talking (i.e., the form of free direct speech was used). In addition to these findings, it was shown that girls used evaluative devices in their narratives to a greater extent than boys. These gender differences are in accordance with what Ely and colleagues (1993, 1995, 1996a, 1996b) have found concerning the use of speech reporting.

The studies reviewed above include experimentally elicited speech from make-believe contexts where props and dolls are provided. Reilly (1992) used the frog story picture series as an elicitation instrument. This is another type of task that puts other demands on the child (cf. discussion in Section 2.3). In this study, a group of 3-4-year-olds and a group of 7-8-year-olds were included, and

along with story structure, Reilly was interested in the storytelling performance including elements of affective expression (e.g., quoted speech). Reilly found that there were no significant differences between the groups regarding statements of linguistic evaluation (including, e.g., quoted speech), but the pre-school children used significantly more facial expressions, gestures, and prosodic features than the 7-8-year-olds did. In contrast to the younger children, the 7-8-year-olds produced longer stories of a greater structural complexity. A second study included a group of 10-11-year-olds in addition to 7-8-year-olds, and this time the recipient was not an adult but a pre-school child. Similarly to the first study, the 7-8-year-olds told structurally coherent, but paralinguistically flat stories. The 10-11-year-olds, in contrast, included quite a few affective devices and marked aspects of delivery (like the pre-school children). Thus, the children just having entered school (the 7-8-year-olds) concentrated on structural complexity at the expense of integrating affective devices and making a vivid performance. It is noteworthy, however, that there were no significant differences concerning the inclusion of character speech among the groups.

Writing

From the research reviewed above, it is evident that certain aspects of marking of delivery, (e.g. marking of voice pitch), develop early. What about representing and marking dialogue in writing? Studies on children's beginning use of punctuation are rare and limited (as Hall and Robinson (1996) point out). However, Ferreiro & Zucchermaglio (1996) and Ferreiro & Pontecorvo (1999) analyzed written stories of Little Red Riding Hood from Spanish and Italian primary school children for the use of punctuation marks and direct speech. In both studies more than 1/3 of the texts had no internal punctuation at all, despite the fact that some of the texts contained dialogue. Yet, in those texts where punctuation marks were used, they appeared with greater frequency and variety in quoted speech than in the passages with the narrative text. Importantly, punctuation marks were not often used conventionally

and "correctly" in the dialogues, and in general, "children had difficulty using punctuation marks consistently throughout the quoted speech portions in their texts" (Ferreiro & Zucchermaglio 1996: 194). This was particularly true in those texts containing long portions of dialogue. These results in combination with the fact that Ferreiro and her colleagues often found an overuse of graphic marks in relation to direct speech, indicate that many primary school children are indeed aware that quoted speech should be marked, but not *how* it should be done. Ledin (1998), investigating written argumentative texts by Swedish children 8-12 years of age, found that a common strategy to mark shifts of speaker perspective in dialogues was to simply start a new sentence.

The area of speech act verbs and children's writing also seems to be neglected. The only attempt in this direction found is Tannen's (1987) report on one junior high school girl who received an assignment to write a story in which the dialogue was introduced with other words than *said*. This girl managed very well using words like *shriek*, *exclaim*, *chorus* and *inquire*, but the narrative had a slightly "forced quality" (Tannen's expression) which had no resemblance to how this girl normally wrote, nor how she chatted with her friend.

Gaps in research

The review of earlier studies above is indicative of certain research gaps. Firstly, studies focusing on use of direct and indirect speech in speech reporting are more frequent than examinations of speech projections. Most often the speech reports are elicited experimentally. Furthermore, the category of free indirect speech is neglected. Secondly, most studies are based on English-speaking children, and few studies have been conducted on Swedish-speakers. Finally, developmental studies of speech reporting and/or projection in *writing* seem to be extremely rare.

2.4.4.2 Input

The issues of linguistic input, and the role of conversational and cultural context, generate several important questions with regard to direct and indirect speech. The question of what types of forms, and to what extent they occur, in speech directed to children is relevant, as is the relation between the contents and structure of the input on the one hand and the children's production on the other. Unfortunately, studies on the use of direct and indirect speech in speech directed to children are rare. However, in a longitudinal study of 10 Canadian families, Ely (1993) found that the amount of talk about the past (and about past speech) by parents was positively correlated with their children's concurrent use of reported speech. It was noted that this fact could be explained by the children's high rate of compliance to their parents' prompts about past speech events. Ely *et al* (1996b) showed that there was a positive correlation between parental attention to past speech in conversations with their children (ages 2;1 to 3;7) and children's unprompted use of reported speech with an interviewer at later ages (3;0, 3;7, and 5;0). Hence, although children's speech about past speech often is prompted by the parents to begin with, children use it independently later on. These results support what has been reviewed elsewhere in this thesis (Uccelli *et al* (*in press*), see page 35), namely that children who have, from an early age, been stimulated to talk about distant (speech) events and objects, tend to become frequent (and skilful) users of speech quoting later on.

Further, positive correlations of this type seem to be connected to gender. Ely *et al* (1995), investigating 22 dinnertime conversations, (each including one child (between 2;3 and 5;2 of age), and the father and the mother of the child), found a positive correlation between mothers' (but not fathers') use of speech reporting and children's use of speech reports. Moreover, Ely and colleagues found 268 reports of past speech in total, and 167 of these were made by the mothers, 68 were produced by the fathers, and only 33 were produced by children. Ninety-one percent of the mothers produced a speech report at least once, 86% of the fathers, and 45.5% of the children. Thus, mothers produced

speech reports more often than the fathers and far more often than the children. The mothers also typically used indirect speech while the fathers produced narratized speech (e.g., *I had a long conversation with the guy at the liquor store*), and the children preferred direct speech (including framed as well as free direct speech). The results show that there seems to be a correlation between how much the children hear speech reporting and their own production. Ely *et al*'s study also indicates that mothers' use of speech reporting is of particular significance. Notably, the mothers commonly use indirect speech whereas the children use forms of direct speech. This fact suggests that there is no straightforward relationship between what forms the children hear in their input and what forms they use themselves.

The custom of the parents to talk about past events and to prompt speech reports from the children not only has effects on the child's mere production of these particular forms. It also has important pragmatic and cultural implications: "... reported speech provides listeners with an opportunity to reinterpret past speech, and speakers, in recalling and reflecting on past speech, both redefine and reaffirm the discourse practices of their community" (Ely *et al* 1995: 214). Speech reporting seems to be used differently and to different extents in different cultures. For example, Labov (1972) suggests that middle-class white speakers tend to use more external evaluation in their personal narratives, while inner-city blacks use more internal evaluation (which typically includes forms of direct speech).[39] Wolf & Hicks (1989) compared two 7-year-old girls of different backgrounds narrating a replica play task. Malka, a white middle-class girl and a typical "writer" who came from a home environment where a great deal of emphasis was placed upon literacy-related activities, turned out to recount the events closer to an account than to a story. The events were recounted almost wholly from the narrator voice, and the dialogue voice played only a minor role. In contrast, in the narrative of the

---

[39] The scale of evaluations "range from the most highly internalized type – a symbolic action or the evaluation of a third person, to the most external – a direct statement of the narrator to the listener about his feelings at the time" (Labov & Waletsky 1967: 39).

black working-class girl Rene, the dialogue voice was dominant, the events were recounted from the perspective of the characters, and the speech interaction between the characters had the role of moving the story forward. Rene's way of narrating, including sophisticated use of prosodic devices and internal evaluation strategies, can be explained by her sociocultural background encouraging a rich, oral style.

Many children have stories from books read to them early on in life and these often contain direct and indirect speech. The language of fairytales and children's books differ in several respects from the language used, for instance, at dinnertime conversations. Importantly, although the story is read aloud and delivered through the oral mode to the child, it is *written*, often literary, text that is depicted. When children learn to read themselves, they come in contact with written representations of speech. In analogy with 'child-directed speech' we may now talk of 'child-directed writing'. Baker & Freebody (1989) examined 163 English reading books for young children and found that, taking occurrences of *said, say*, and *says* together, this word family was second in frequency to *the* in their corpus of 84,000 words. Hence, children both hear and see a substantial amount of direct and indirect forms of speech from an early age.

Perera (1996), making detailed analyses of 35 different children's early reading books in English, found a great variety of ways of marking framed and free direct speech. Besides the most conventional marker in English, the double inverted commas (" "), she found single inverted commas (' '), and in some books no quotation marks were used at all. Instead, speech balloons were used, while in other cases the speech was set out just like the rest of the text. In yet other books, the quote was placed on a separate and indented line. Perera also investigated the positioning of the framing clause with the verb of saying. In some books she found no framing clauses at all (i.e., free direct speech), and besides using the method of starting on a new line for each new speaker, using quotation marks, and placing speech within balloons, she found examples of direct quotes printed in italics. Young children thus come across a number of

different ways of presenting direct quotes in their early reading books. In fact, what Perera finds is that the diversity is greatest in the books meant for children younger than 7. After that the books follow adult conventions of punctuation and are relatively uniform. In the 35 books studied, Perera found 160 examples of framed direct speech and in 18 cases (11%) the framing clauses preceded the quote, they followed the quote in 84 cases (53%), whereas 58 (36%) framing clauses were placed in the middle of a quote. Thus, final and medial framing clauses are far more common than initial ones.

Björklund (*personal communication*) surveyed several children's reading books for primary school use in Swedish, and found exclusively dashes as markers of direct quotations. However, in books translated from English, quotations were marked by means of quotation marks. In addition to these findings, she found only one instance where the clause with the speech act verb preceded the quote. In all other cases where a clause with a speech act verb was present, it followed the quote. These results, together with Perera's, suggest that a very common (if not the *most* common) speech model of framed direct speech that school children come in contact with in writing, is the case where the framing clause follows the quote (cf. discussion 2.4.3.3).

2.5   *Research questions reconsidered*

The reviews of earlier research and the discussions that have been raised in relation to them in the previous sections, make possible a more detailed account of the research questions as stated in the beginning of this thesis (Section 1.1.3). The questions – under the four headings *Development of forms, The use of the forms, Input characteristics*, and *Later development in speech and writing* – are thus further elaborated below. Each subsection ends with a number of explicit questions that will be operationalized and explored in the sections to come (Chapter 3 and 4), and answered and discussed in Chapter 5.

Development of forms

The first issue to examine concerns the emergence and the course of development of direct and indirect forms in Swedish children. There are certain cognitive, communicative, as well as linguistic (cf. Section 2.1.1) factors connected to the types of forms that may predict the order of acquisition. *Free direct speech* should be easily accessible to the child, in that speech reporting then involves simply mimicking of earlier speech without shifting deixis (e.g., changing tense or pronouns). In the case of indirect voicing, the child is merely speaking via another object (i.e., enactment). Used in these ways, the free direct form does not put the child to a severe test from a cognitive and linguistic point of view, and this form can thus be expected to appear early. However, unreflected use of free direct speech may result in communicative/pragmatic shortcomings, for example, failing to mark shifts in perspective explicitly. To avoid these possible traps, some kind of marking needs to be done, for instance, by using relevant prosodic cues. Another way is to add a lexical frame, i.e., to use *framed direct speech*. In this case, we are syntactically dealing with two independent clauses. These clauses have particular deictic properties. The clause with the speech act verb departs from the deictic center of the current speaker, and the mere speech is then to be reported from the point of view of the original speaker. This means that relevant pronominal and tense shifts need to be accomplished by the speech reporting/projecting child. This fact in combination with the fact that this structure involves two clauses that need to be coordinated, indicates that framed direct speech emerges later than free direct speech.

Indirect speech lands even higher on the scale of linguistic and cognitive complexity. The deictic center remains where the reporter/projector is (and deictic markers need to be adjusted accordingly), and the construction includes a complement clause. Due to these properties, we may expect indirect speech to appear later than the direct forms. In the case of *free indirect speech*, finally, person and tense deictically belong to the reporter/projector (similar to indirect speech), but in all other respects this form behaves like direct speech. Since this

form is rarely included in developmental studies, it is an open question whether it is at all used by children. If it *is* actually used, the question is when does it appear and in what types of context. If it is scarcely employed by young children, the following questions should be addressed: is this a phenomenon of the form primarily being a stylistic device (used only by professionals in literary prose); is the form too grammatically and syntactically complex, or are there functional explanations (e.g., that the child does not "need" it)?

The review of previous developmental studies above (Section 2.4.4.1), indeed shows that direct speech (free and framed direct speech) emerges first, to be followed by (correct use of) indirect speech only later. More specifically, free direct speech is used to a larger extent in the younger age groups than in the older ones. Pre-school enact rather than report (Hickmann 1993), and fail to make role and speech boundaries clear (Goodell & Sachs 1992). School children, in contrast, often frame their utterances verbally (Goodell & Sachs noticed that the 8-year-olds framed their direct quotes in 99% of the cases, and Hickmann found that 10-year-olds, (but not 7-year-olds), also did so). Indirect speech, finally, is not frequently used by children, but is mastered around 8 years of age in American English (Goodell & Sachs 1992), and around age 13 in Turkish (Özyürek 1996). In contrast to school children, many adults prefer the indirect style (Hickmann 1993). Yet, all these results are based on experimentally elicited data. Ely & McCabe (1993) examined spontaneous speech by three English-speaking children between one and five years of age and found that, in addition to forms of direct speech, indirect speech appeared at an early point in time. Hence, it seems the statement that free direct speech is the first form to be acquired, to be followed by framed direct speech, and only later indirect speech, is a simplification of the developmental picture.

What then remains to be seen in the Swedish data of this thesis is: What forms are the first to appear, what do they look like, and in what discourse contexts do they appear? How frequent are the forms and how does the course of development look?

The use of the forms

It is not solely of interest to explore when the first forms appear and determine the course of development, but it is also important to discover *how* these forms are used and how the *functions* develop. The fact that Ely & McCabe's results differ from the other studies (see above) indicates that the type of activity engaged in is an important factor to consider. The data included in this thesis contain several types of discourse contexts and can thus shed further light on this question.

The longitudinal case studies consist of naturalistic and spontaneous data and both speech projections and speech reporting may appear. As was discussed in 2.4.2.3, these types have different functions in that the former is typically used in frames of fiction where there is no real speech event that can be referred back to, whereas the latter is used in contexts where there is a previous speech event to take a stand on. In the Longitudinal case study data we will be able to determine if both types are used by the children, and in that case, how they are used, and when the various types begin to appear. In the Doll house data and the Frog story data, the primary concern will be speech projections. Yet, the types of contexts and the communicative conditions also differ, and we will be able to examine possible activity related effects in the use of speech projections.

It was evident from the discussion on understanding of minds (Section 2.1.3) that the ability to attribute intentions to others (animate or inanimate objects) is acquired already before the age of three years. Worth exploring are the relationships between this emergent theorizing of mind and the first attributions of *speech* to others. Assigning intentional behavior to others is an important achievement, and a first step, in the process of taking and understanding the perspectives of others. Yet, being able to assess of different perspectives, and to adapt to different conditions of communication, continues to develop for a long period of time. The review of research within narrative development (e.g., Section 2.3.2), indicates that young children have certain problems in considering what information is shared and unshared with the

listener (e.g., 3-4-year-olds providing ambiguous pronominal references (Karmiloff-Smith 1981)), whereas school children do better. The ability to assess and linguistically encode shifts of perspective is indeed central in relation to speech reporting and speech projection.

Linguistic information can be simultaneously packaged into forms of direct and indirect forms of speech, and the use of these forms can serve several functions and create different effects (cf. discussion in 2.4.3.4). Information can be packaged by means of linguistic constructions, word choices, orthographic modifications, gestures, prosody, etc. in order to create an effect (for example depiction of referents, distancing, dramatizing, and speaker perspectivizing). The research reviewed in 2.4.4.1 indicates that it takes some time before children command all these aspects and Berman and Slobin argue that "school-age children show considerably less control of packaging than adults, and pre-schoolers are only beginning to exploit the possibilities of narrative syntax" (1994: 540).

In conclusion, there are several questions to be answered regarding the development of uses of the direct and indirect forms: In what types of conversational contexts do the first forms of direct and indirect speech surface in the youngest children investigated in this thesis (i.e., Harry and Tea), and what are the distributions of speech reports versus speech projections? In what ways, and to what extent, are speech projections used by the three-year-olds playing with the doll house, and their peers narrating the frog story? Do the subjects included in the data of this thesis (children, adolescents, adults) use the forms in such a way that the listener and/or interaction partner can follow the changes of perspective; is it made clear who the reported or projected speakers are? How does the structuring and packaging of linguistic information into the direct and indirect forms look? Since we are dealing with production data only, it is not possible to determine levels of intentionality behind multifunctional use. However, it is possible to study the effects of the use, and explore input data to get an idea of what the parents promote in this respect.

Input characteristics

In Section 2.4.4.2 above, I reported on findings that parents (especially mothers) often reported speech of past events often while young children did it less often (Ely 1993). Moreover, the speech reports given by children were often prompted by the adults. Thus, the parents in Ely's study seemed to want to direct the child's focus to past events and past speech and an explanation for this parental behavior was a desire to make the child reflect on past (speech) events and experiences. In this way, children are socialized into discourse practices (which include how to use the forms) of the particular community.

In the coming analyses, I will explore what the input and the interaction look like in different narrative contexts. To what extent do the adults socialize the children into activity-related use of the different forms of speech? Do the adults provide the children with discourse models, and in that case, what forms do they use? Do the adults prompt their children to use the forms? Are there any relationships between the adults' use of particular forms and the children's own production? These questions will be examined in the cross-activity data from the 3-year-olds and their adult interaction partners. In addition, the data from the Longitudinal case studies will be explored from this point of view, which also makes it possible to analyze whether the adults attune their speech to the children at different age stages. One question to examine in relation to this is if there are any particular relationships in time between the children's' first use of the forms on the one hand, and the adults providing of speech report/projection models and speech prompts on the other.

Later development in speech and writing

Indeed, language input and interaction involve much more than the interaction with and the speech provided by the caretakers or other close family members. Among older children, this fact becomes even more evident as the social sphere is extended and new activities are engaged in. Starting school around the ages of 6-7, for example, means entering a new social context, and this institution

among others will certainly have an important impact on the child's language development for a long period of time.

One aspect of language use that the school focuses on and that is to a greater or lesser extent a new experience for children, is the written language. Many younger children that do not know how to read and write, come into indirect contact with written literary language if and when they have stories read and told to them. As was reported in Section 2.4.4.2, fairy tales and stories for children tend to involve many direct and indirect forms of speech, and direct quotes followed by a verbal framing clause seem to be especially frequent in these contexts. Yet, when learning to read and write, the contact with written language becomes even more direct and the children may start to write dialogues and produce forms of direct and indirect speech in writing themselves.

The design of the data makes it feasible to investigate and compare 3-, 4-, 5-, 9-, 12-, 15-year-olds' and adults' use of direct and indirect speech in spoken narratives. In addition, written data from the school children and the adults are available so that developmental aspects and comparisons between the different modes of production are possible. An investigation of the use of these forms is not only valuable so as to map linguistic development during the school-years, but also from a metalinguistic and pragmatic developmental point of view. When projecting speech, the speaker attributes speech to another "speaker", and when doing this in writing, the speech is rendered in a different modality. Thus, language development also involves distinguishing properties connected to the use of modalities and to different genres. Developing this awareness depends on the amount of practice (not least in school) as well as on cognitive maturity. Earlier research results on modality-specific awareness and narrative production (as reviewed in Section 2.2.3), suggest that this development is not straight-forward, and that an emerging control over differences between the modalities is evident in the early teens.

Questions to be examined in the following sections pertaining to this discussion concern the ways in which school children's use of direct and

indirect speech differs from younger children's use on the one hand, and from adults on the other. Also, are there differences in use among the 9-, 12- and 15-year-olds and the ways they handle the forms when speaking in comparison to writing?

3 Methodology

As the theoretical foundation has now been laid, and the research questions posed, the time has come to discuss the methodology used in the analyses of the data. This chapter begins with a discussion of the fact that long-term language development and later stages of development have not previously been given priority within research on children's language (Section 3.1). In relation to this discussion, I present the data design of this thesis which is argued to be a feasible framework in order to provide insight into questions of long-term development. In the second section (Section 3.2), the subjects and the data included in the corpora are presented in detail. The chapter concludes with a description of the procedures of analysis that are applied to the data (Section 3.3).

3.1 Long-term development and methodology

As was discussed in Section 2.1.1, language development is an open-ended process. Development is always occurring, although there are certainly phases where some aspects of language develop more rapidly than in other phases (cf., for instance, the vocabulary spurt in the child's third year, (Bates *et al* 1988; Bates *et al* 1994; Strömqvist *to appear*, Berglund 1999)). Despite the fact that language development proceeds continuously and is not completed at an early age, studies on long-term development are rare.

There are several possible explanations for the limitations in ranges studied, and for the absence of discussion. First, not everyone agrees that language development is an open-ended, non-linear, and long-term process, and thus there is no reason to extend the analyses over a broader age range. Indeed, there are grammatical forms that appear early and seem established already at an early age, and this is used as an argument not to include older children in such an analysis. However, although a form is found in the young child's language production, forms take on different functions during

development (cf. the discussion in Chapter 2). This is an important reason to study language use over a longer period of time. Second, there are methodological considerations. The data needs to be adequate and comparable between the age groups. In the case of experimental methods, the experiments need to function well with young children as well as with older children or even adults. Moreover, in order to draw theoretical conclusions, the size of the corpus needs to be large, and both data collection and analyses are time-consuming. If children are studied longitudinally, this will certainly take a considerable amount of time.

Eliciting and analyzing narratives from children in different age groups, is a much practiced and well established method for studying language development. Using a picture series like the frog story to elicit narratives, has proven to be a successful methodology (see further Sections 2.3.2-3). Berman & Slobin (1994) state that their database, building on frog story narratives, differs in several respects from other psycholinguistic methods and materials used earlier. Firstly, production rather than comprehension or recall is focused on. Secondly, the sequence of pictures is lengthier and more elaborate than those customarily used in picture-elicitation tasks and this fact allows for investigations of not only anaphora and reference, but temporality and expressions of relations between events as well. Finally, and importantly, narratives have been elicited from children as young as 3 years of age, pre-school and school children, and, in addition to these subjects, from adult narrators. This makes it possible to analyze and make statements about long-term narrative development. Yet, it should be noted that the oldest school-age children in that database are 9 years of age. This means that adolescent narrators are not included, and thus, little can be said about narrative development during the teen age years. The approach to studying very young children's language use, (and comparing it to adult use), is indeed the dominant one within the field of language development. There are also investigations of older children's language use (e.g. Labov 1972, Nordberg 1985), but these are often made without relating the results to those of younger children. Although

the interest in older children's language development has increased recently (e.g., Nippold 1998, Aisenman (ed.) 1999) there is a great deal more to investigate, especially regarding truly long-term development.

However, although an elicitation instrument like the frog story is well suited for studies of long-term development since most children and adults are able to produce narratives from this type of picture book, it is difficult to elicit narratives from children younger than three years of age (and even 3-year-olds may have problems in understanding the task and grasping the story-line (cf. the discussion in Section 2.3.3., and Berman & Slobin 1994: 58-59)). This may be due to factors such as limited attention and memory span, and their emerging sense of narrative structure. Thus, to get a picture of when particular forms of language appear (in this case, forms of direct and indirect speech) and how they develop in early years, other types of data need to be included. In this thesis, data from two children followed longitudinally from around one and a half years of age, will be used as a complement to the frog narrative data. As will be explained in more detail in the next section, the Longitudinal case study corpus consists of video and audio recorded naturalistic interactions between the children and their caretakers, from the time the children are around one and a half years old to when they are four years of age. Since the Frog narrative data included in the present analyses cover the age ranges of three years to adults, the material includes data from the very first words produced by the two Longitudinal case study children, up to adult use of language. This also results in an overlap in ages three and four, since frog narrative data as well as case study data are available at these ages. Moreover, data from three-year-olds playing with a doll house with their mothers also enters the design, resulting in rich data from an important stage of language development. A schematic picture over the data design of this thesis is illustrated below. Note that all these data are production data. The analyses of the use of direct and indirect speech will thus focus on the production aspect rather than exploring the comprehension side of these matters.

longitudinal case studies	make-believe play	picture-elicited narratives
1		
2		
3	3	3
4		4
		5
		- - - - - -
		9
		12
		15
		adults
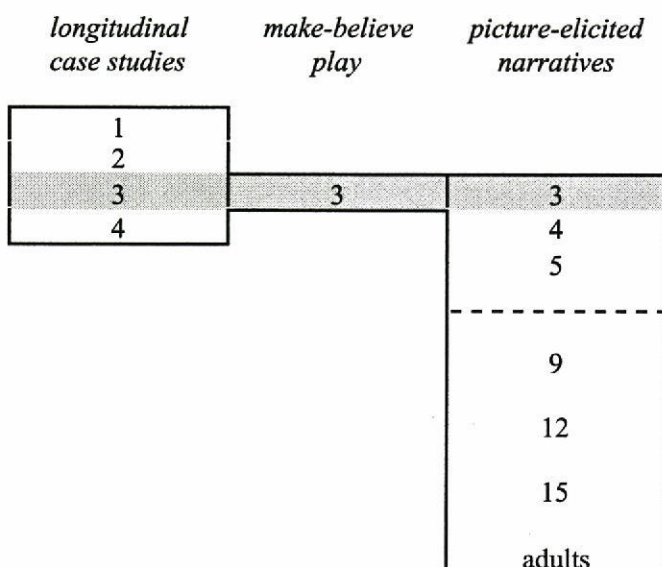
Figure 3-1: Schematic picture of the data design. The numbers represent the ages.

The data was not collected for the purpose of analyzing direct and indirect speech, but the data has been selected in order to explore the research issues as presented in Section 2.5 above. Thus, first, it will be possible to analyze the development of the forms and functions of direct and indirect speech. Second, the data includes different types of narrative contexts that allow for analyses of activity-related use. This applies especially to the age group of three years, where spontaneous data from every-day contexts (the Longitudinal case studies), is compared to a make-believe context (play with a doll house), and the narrative context of telling the frog story. This cross-age and cross-activity component of analysis is represented by the shaded area in Figure 3-1 above. Third, the Longitudinal case studies, as well as the play with the doll house, and the Frog story data by the 3-year-olds, allow for analysis of both child production data and parental input. Finally, the 9-, 12-, 15-year-olds and the adults have produced both oral and written frog story narratives (the fact that this has been done by the older subjects, but not the younger ones, is marked by

a dotted line in Figure 3-1, separating the age groups). In this way, the design allows for cross-modal comparisons regarding the use of direct and indirect forms of speech in speech and writing.

Hence, this is an explorative study, aiming at providing further insight into questions about long-term development. While the data included is rich and diverse, covering several age groups and different narrative contexts, this data-design also introduces certain theoretical and methodological challenges. For example, apart from the Longitudinal case studies and the 3- and 4-year-olds telling the frog story, the data is cross-sectional, i.e., the same children are not followed over time. There is a general theoretical problem with this type of time-saving methodology and it is not restricted to this study. However, all groups to be investigated here contain fourteen subjects (except for the two Longitudinal case studies), and this amount of data will strengthen an indication of general patterns of use. In addition to this, although general developmental tendencies will be the focus of the analyses, individual strategies will also be considered. Another challenge of the diverse material included in this study, is to find ways of comparing between age groups, narrative contexts and subjects. In the sections below, the subjects, data and types of analyses are described in detail including the statement of specific research questions, and description of operationalizations.

3.2   Subjects and data

As was discussed above, the empirical basis of this thesis rests on three main types of data or corpora. In total, the empirical material that will be used consists of close to 275,000 words. A majority of these (175,328 words) are from the Longitudinal case study corpus that will be presented in Section 3.2.1 below; the cross-sectional Frog story corpus, to be described in Section 3.2.2, is comprised of 75,905 words; whereas (approx.) 22,900 words are produced by three-year-olds and their mothers in the Doll house play corpus (see Section 3.2.3).

3.2.1 The Longitudinal corpus

The Strömqvist and Richthoff corpus consists of longitudinal language data
from five Swedish monolingual children (for a detailed account and description
of the entire database, see Strömqvist, Richthoff & Andersson 1993, and
Richthoff 2000). The children have been audio and/or video recorded during
typical everyday activities like free play, bedtime and eating routines, and book
reading, and the material has been transcribed and coded according to CHAT-
format (MacWhinney 1995). There are transcribed data from all children from
around age one and a half years to four years, and the recordings have been
made at approximately three-week intervals. The database, at the present time,
comprises nearly half a million words produced by the children and their
interaction partners (Richthoff 2000).

   The longitudinal corpus referred to in this study, consists of two of the
children in the Strömqvist and Richthoff corpus. Hence, the language
development of the boy Harry and his little sister Tea will be focused on. There
are at least two important reasons for studying these two children. First, there
are comparable recordings available. The recordings have been made at similar
time intervals (around once a month), the datapoints can be matched, the same
age range is covered, the length of the recordings are about the same, and the
activities are of similar types. Thus, the conditions for comparing the two
children are good. Second, the same adults (primarily their mother and their
grandmother on their mother's side) represent the adult input. Their behavior in
relation to the two individual children, (and the impact it has on each child's
language development), is worth examining and discussing.

   Harry is the first-born child of a middle-class family living on a farm in
the countryside on the west coast of Sweden. Tea is three years younger than
Harry and the second child. They speak standard Swedish with a touch of
western Swedish dialect. For each child, 28 recordings, or data points, have
been selected, and cover the age range of 19 months to 47 months. In all, this
sub-corpus consists of 56 datapoints and around 175,500 words. There are

audio recordings available for all datapoints, but many of the interactions have also been video taped which facilitates analyses of, for instance, use of gestures. All recordings were made in the children's home-setting, and only one of their parents and/or a grandmother, and the child were taking part. The typical activities for both children to engage in were activities related to meals, free play, picture-book reading, make-believe play, nursing activities (like looking after their toy animals), puzzles, drawing, and conversation about personal experiences.

Harry can be described as a calm, harmonious, and thoughtful little boy. He very much enjoys playing with a building kit, his toy animals, and he has a great interest in vehicles. He has a normal language development, as is mirrored in the mlu-curve below. The Harry corpus comprises approximately 26,000 words uttered by Harry, and close to 60,000 words produced by his adult interlocutors (mostly Harry's mother, but also Harry's grandmother, and occasionally, his grandfather).
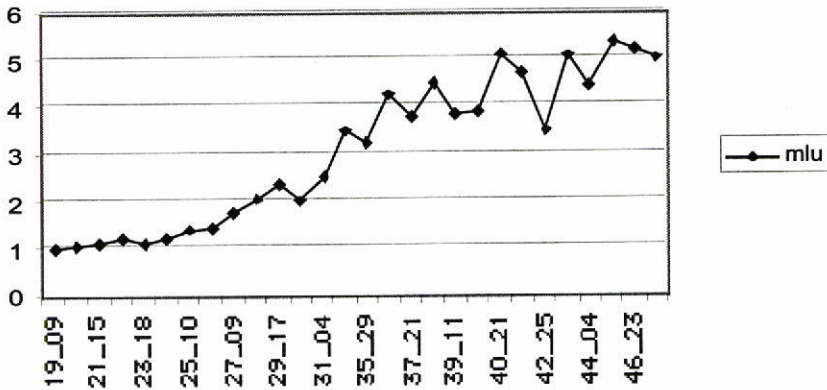
Figure 3-2: Ratio of morphemes over utterances (mlu) in Harry; age range: 19 months and 9 days, to 47 months and 20 days.

Tea is Harry's baby sister. She is a temperamental, enthusiastic and happy girl, expressing rich facial expressions and gestures. She devotes much of her

time to nursing and playing with her dolls and her toy horses. Like her brother, she enjoys looking at and reading picture books. Out of a total of 89,000 words, Tea produces 33,000 words, whereas Tea's mother, grandmother and father stand for around 56,000 words. Tea's mlu development resembles Harry's, although she can be observed to spurt somewhat earlier than her brother.
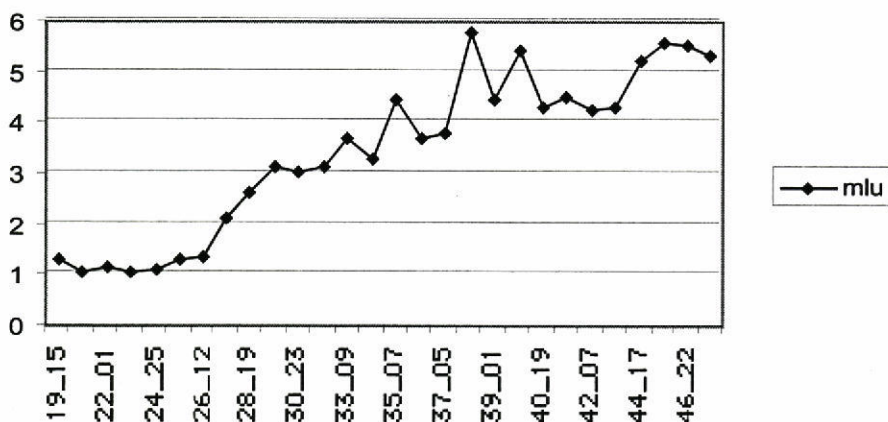
Figure 3-3: Ratio of morphemes over utterances (mlu) in Tea; age range: 19 months and 15 days, to 47 months and 23 days.

Table 3:1 below presents a summary overview of the quantitative data of the Longitudinal corpus.

Data	Age range	No of data points	Total no of child-produced words (utterances)	Total no of adult-produced words (utterances)
Harry	19;09 – 47;20	28	26,347 (10,087)	59,875 (12,494)
Tea	19;15 – 47;23	28	32,966 (10,902)	56,140 (12,630)

Table 3:1: Quantitative data of the Longitudinal corpus.

The recordings are referred to in the text by types of format codes such as "Har28_02.cha" and "Tea24_25.cha", where the first three letters indicate the child being recorded (i.e. Harry and Tea, respectively); the numbers refer to the age of the child by the date of recording (i.e., 28 months and 2 days; 24 months and 25 days); and ".cha" means that the transcript is in the chat computer format (for a more detailed description, see Section 3.3.1 below).

3.2.2 The Frog story corpus

3.2.2.1 *The elicitation material*

Frog, where are you? (Mayer 1969) is a wordless picture book first used by Bamberg (1985). Since then it has been used to elicit narratives in a great variety of languages and ages (see Berman & Slobin 1994, and Strömqvist and Verhoeven *forthcoming*). The booklet consists of 24 pictures and depicts the adventures of a little boy and his dog as they go out into the woods to search for a frog that has disappeared from the boy's room. More specifically, this is what can be seen in the 24 pictures, respectively, (the complete picture series is found in Appendix):

Pict. 1.	A boy and his dog are looking at a frog in a jar
Pict. 2.	The boy and the dog are in bed; the frog escapes from the jar
Pict. 3.	The boy and the dog wake up and look for the frog
Pict. 4.	The boy and the dog search the room for the frog
Pict. 5.	Both look out the window; the dog has the frog's jar on his head
Pict. 6.	The dog falls out of the window, breaking the jar
Pict. 7.	The boy stands outside the window; the boy is carrying the dog which is licking the boy's face
Pict. 8.	They go to the forest, the boy calling for the frog
Pict. 9.	The boy searches a gopher hole; the dog is sniffing a beehive
Pict. 10.	A gopher comes out of the hole and bites the boy on the nose
Pict. 11.	The beehive is on the ground; the boy is searching in a tree
Pict. 12.	The boy has been frightened by an owl and has fallen to the ground; the dog is chased by bees
Pict. 13.	The boy is by a rock; the owl is flying above him
Pict. 14.	The boy has climbed up the rock and is holding onto what he believes are branches, the boy is calling for the frog

Pict. 15.   A deer hiding behind the rock picks up the boy on its antlers
Pict. 16.   The deer runs off with the boy, the dog in hot pursuit
Pict. 17.   The boy and the dog are thrown down a cliff
Pict. 18.   They land at the bottom in a pond below the cliff
Pict. 19.   The dog climbs on the boy's head; the boy puts his hand to his ears
Pict. 20.   The boy tells the dog to be quiet
Pict. 21.   They climb over an old log
Pict. 22.   They see two frogs sitting on a bank
Pict. 23.   They see that there are also a number of baby frogs
Pict. 24.   They take a (baby) frog and leave, waving to the remaining frogs

The story falls within the category of "fictional narrative" (see Section 2.3.1). In one sense, the frog story is a typical children's story.[1] It has a main protagonist and hero (the boy), there is a problem that needs to be solved (find the frog), there is a set of complicating actions (the search for the frog and all adventures related to that), that lead to a resolution and happy ending (they find the frog). Thus, the story frame is very much in accordance with the universal narrative scheme as described by Labov & Waletzky (1967) and Labov (1972) (see Section 2.3.1). In another sense, however, the frog story is atypical, since it describes a series of events mediated only by means of pictures (i.e., there is no text involved).

Although there is one main protagonist, i.e., the boy, there are also two other important figures: the frog, and especially the dog. Besides these subjects, there are a number of acting animals: the bees, the gopher, the owl, the deer, and the other frogs in the frog family. The narrator, thus, has the opportunity to take different perspectives on events and on actors. For this thesis, it is of interest to investigate how speech projections are used in the process of perspectivizing. There are several pictures in the story in which the boy can be seen performing speech acts, for example, calling for the frog out in the woods (picture 8), into a gopher hole (picture 9), and from the rock (picture 14); in picture 20 the boy tells the dog to be quiet, and in the final picture the boy can be seen waving goodbye with his mouth open as if he is saying or calling

something. However, although it is only the boy that is seen explicitly performing speech acts, speech may also be projected onto other role characters. The fact that all the other protagonists are animals that do not have the ability to speak in reality, does not necessarily prevent them from being attributed with the ability to speak in this context. Quite the contrary, in fiction, and especially in children's stories, animals are often ascribed agency and human-like behavior (cf. also the genre of fables). Projecting speech to several role figures, is indeed a powerful device in displaying shifts in perspective.

3.2.2.2 Subjects, data, and procedures

The Frog story corpus used in this thesis includes narratives produced by Swedish monolinguals in the age groups of 3-, 4-, 5-, 9-, 12-, 15-year-olds, and adults. The data has been collected by several researchers in different projects so the subjects and data are presented separately below. An overview of all the data will be given at the end of this section.

3- and 4-year-olds

Berglund and Eriksson (2000) collected Swedish parental report data on children's communicative skills in two randomly selected, nation-wide cohorts.[2] After completion of the study, the families concerned were asked if they wanted to participate in a longitudinal follow-up. Those families accepting to participate had a tape and a copy of the frog story booklet sent home. The parents were instructed to record their child telling the story of the book after having looked through the booklet once. The child was to look at one page at a time and not to turn back. Parents were to minimize their interaction with the child to what was needed to elicit a story. After completion of the task, the tape was returned to the researchers.

---

[1] It should be added that it is a typical *Anglo-Western* children's story (see discussion in Berman & Slobin: 20-22).

[2] The type of parental report used was The Swedish Early Communicative Development Inventory, (SECDI). The SECDI is based on The MacArthur Communicative Development

The children of the first follow-up were either 3 years or 3 and a half years, and consisted of twenty-seven children, whereas subjects in the two-year follow-up comprised twenty children. Fifteen children participated in both follow-ups. (For more details concerning the subjects, see Berglund (1998), and Eriksson (*submitted*).) Fourteen of the fifteen children producing frog story narratives on both occasions have been selected for the investigations of this thesis (the fifteenth child had to be left out due to bad recording quality). The average age in the 3-year-old group is 3;2 (ranging between 3;1 and 3;7), and in the 4-year-old group 4;2 (range 4;0-4;7).

The audio-recordings were sent to the Department of Linguistics, Göteborg University, where they were transcribed and entered into the computer according to CHILDES, that is, in CHAT-format (MacWhinney 1995). The transcriptions were made by students at the department, and checked and coded by myself. The coding unit used in the frog transcripts, is the "clause". A clause is defined as a unit containing a finite verb, although in some cases a clause could contain only an infinite verb (e.g., in the cases of subordinate clauses). As was stated above, the parents were instructed to minimize their feedback, however, some children had problems constructing a narrative on their own, so the parents needed to engage the children by asking questions and the like.[3] The parents' utterances have then also been chunked into clauses. Clauses about subjects other than the story (e.g. discussions about the recording equipment) were removed.

The reason for using clauses as a unit of coding and of analysis, is for the reasons given in Berman & Slobin (1994), namely that it is more "linguistically structured than the behavioral unit of an "utterance" but … less determined by

---

Inventories, CDI, (Fenson *et al* 1993; Fenson *et al* 1994), and has been adjusted to Swedish conditions by Eriksson & Berglund (1999; see also Berglund & Eriksson 2000).

[3] Berman & Slobin (1994) discuss the fact that many of the youngest children do not present a full plotline in their productions. Due to this fact, they refer to the materials that constitute their database as "texts", rather than "narratives". I will nevertheless use "narrative" for the productions of the youngest children also. In section 2.3.1, narrative was defined as the discourse that encodes the story (i.e., the totality of narrated events), and this definition does not necessarily involve the notion of having a clear plot. The issue of parental scaffolding and narrative construction will be discussed later on.

syntactic criteria than a "sentence"." (1994: 26). As is pointed out, by using this kind of unit, it becomes possible to compare relative length of narratives across the sample, and to the results of the Berman & Slobin study. For this particular study, it will also be possible to say something about proportions of clauses of direct and indirect speech in a certain narrative. For details concerning coding procedure of tokens of direct and indirect speech, see further Section 3.3.

Below are two examples from transcripts of 4-year-olds' narratives. In the first excerpt, a girl called CAT ('*CAT') is relating without verbal interference by her parent, while in the second example, the girl called CUS ('*CUS') is discussing what is seen in the pictures together with her parent ('*ADU'). ('xxx' represents unintelligible speech, and '#' a pause).

CAT04SP; girl, age 4;1: SPOKEN

*CAT:	å så ramla Tommy ner
%eng:	and then Tommy fell down
*CAT:	å hunden sprang
%eng:	and the dog ran
*CAT:	så getingarna jagade hunden
%eng:	so the wasps hunted the dog
*CAT:	å så kom en uggla fram
%eng:	and then an owl appeared

CUS04SP; girl, age 4;6: SPOKEN

*ADU:	va gjorde sorken?
%eng:	what did the gopher do?
*CUS:	han an ble arg #
%eng:	he he got angry
*CUS:	han bet han i näsan
%eng:	he bit him in the nose
*ADU:	xxx va gör hunden?
%eng:	xxx what's the dog doing?
*CUS:	vovvar på trädet
%eng:	barking at the tree
*CUS:	men bikupan rasade ner
%eng:	but the beehive fell down
*CUS:	å dom bli väldit arga på vovven #
%eng:	and they get very angry at the doggie

117

5-year-olds

The data from eight of the subjects in this age group were collected in a project lead by Kerstin Nauclér and Sally Boyd, at the Department of Linguistics, Göteborg University. The aim of the project was to study immigrant children's language socialization at home and in pre-school (Nauclér & Boyd 1996, Nauclér & Boyd 1997). Swedish mono-lingual children were used as a control group, and it is these children that are selected for the purpose of this thesis. One of the tasks the children were to perform was to tell the frog story to a research assistant (Nauclér). The remaining part of data collection was carried out by myself and a student. In all cases but one, the recordings were made in the children's pre-school environment. The eight subjects in Nauclér and Boyd's corpus came from working-class families, while the other seven subjects were children from middle-class families. The average age of all subjects is 5;6 (age range 4;7 - 6;8).

The children had a chance to get acquainted with the content before starting to tell the story and the rule of thumb the research assistant had as a listener, was, (as was also the case with the parents in the 3-, and 4-year-old groups above), to minimize their interaction with the child to what was needed to elicit a story. In contrast to the younger children, the 5-year-olds in most cases had little problem with the fact that the research assistant had a fairly passive role, and they managed to construct a narrative on their own. (These matters will be further discussed below.)

All narrations were audio-recorded, and the narratives were transcribed in CHAT-format on a computer. All the material was coded in clauses, in the same way as described in relation to the 3- and 4-year-olds.

School-age children and adults

The corpus comprises fifteen 9-, 12-, 15-year-olds, and adults, respectively.[4] The school-age children were recruited from schools in middle-class areas around Göteborg, and the adult subjects were recruited from groups of undergraduate students taking subjects other than linguistics or psychology.[5] All subjects were monolingual Swedish speakers.

All recording and data collection was carried out at the Department of Linguistics, Göteborg University. In contrast to the younger subjects (as presented above), the school-age children and the adults constructed a written version of the frog story along with the oral one. Every second subject told the story first in speech and then in writing, and every second subject in the reverse order so as to control for order effects. The narrative task was monological in character in the sense that there was not a listener present in the speaking condition (the subject told the story to a camera) or in the writing condition. The spoken narrations were all video-taped which made it possible to keep track of the narrator's turning of the pages in the book. In addition, eight of the adult narrators were video-taped during the writing session (for more details, see Strömqvist & Ahlsén 1998). Before telling the story for the first time (whether in the speaking or in the writing condition), the subjects were invited to look through the picture book, so as to get acquainted with its content. The instruction given was literally; *Här har du en bilderbok utan text. Bläddra igenom den först så att du får en känsla vad den handlar om. Sedan berättar du vad som händer på bilderna,* ('Here is a picture book without text. Before you

---

[4] The data was collected in relation to two projects and research programs. *Tala och skriva i ett lingvistiskt och didaktiskt perspektiv,* 'Speaking and writing in a linguistic and didactic perspective', was funded by the Swedish Tercentary Foundation (Riksbankens Jubileumsfond) 1992-1994 and directed by Sven Strömqvist and Åke Hellstrand (Strömqvist & Hellstrand 1994); *Reading and Writing Strategies of Disabled Groups,* is supported by the Swedish Council for Social Research (SFR) and directed by Sven Strömqvist and Elisabeth Ahlsén (Strömqvist & Ahlsén 1998).

[5] The subjects in the 9-year-old group attended the 3rd grade in school and the average age was 9 years and 4 months by the day of data collection; the subjects in the 12-year-old group were in the 6th grade and on average 12 years and 5 months old; and the 15-year-olds attended the 9th grade and were 15-16 years old. The subjects in the adult group were between 20-30 years of age.

start, browse through it so that you get a sense of what it is about. Then tell what is happening in the pictures').

The 15-year-old and adult subjects, all familiar with using a word processor, wrote their frog stories on a computer, whereas a majority of the written narratives from the 9- and 12-year-olds were written with paper and pencil. A particular computer tool, ScriptLog, was used as a word processor for those writing on a computer. This program has been developed at the Department of Linguistics, Göteborg University, and is a tool for experimental research on the on-line process of writing. For this particular task, a combination of a text window, (where the text was to be written by the subject), and a window where the pictures in the frog story were displayed, one at a time, was used. (For more details on ScriptLog, see Strömqvist & Malmsten 1998; Strömqvist & Ahlsén 1998.)

The spoken narratives were then transcribed according to CHAT-format, as described above. The hand-written narratives were transferred into text and CHAT-format on a computer. A CHAT-format version was also created for those narratives that were written in Scriptlog. Since the CHILDES system has mainly been developed in order to transcribe and code spoken interaction, (rather than written texts), certain adjustments were necessary in order to be able to do automatic computer analyses (with CLAN, see further Section 3.3). Below is an example from a written adult narrative. The first section is an extract from the ScriptLog produced text-file, and the second section follows the CHAT-version of the same passage:

JOT20WR; female, adult: WRITTEN (text)

Picture no 22
Nej, denna gång har de hittat rätt. Bakom stocken sitter hela familjen Kvack. Lite förlägna hälsar Pelle och Plutt på den stora famlijen, som glatt hälsar dem.

'No, this time they have found their way. Behind the log the whole Family Ribbit is sitting. Being a little bit abashed, Pelle and Plutt say hello to the big family, which happily returns their greeting.'

JOT20WR; female, adult: WRITTEN (chat)

*JOT:	nej [,] denna gång har de hittat rätt.
%eng:	no [,] this time they have found their way.
*JOT:	bakom stocken sitter hela familjen Kvack.
%eng:	behind the log the whole family Ribbit is sitting.
*JOT:	lite förlägna hälsar Pelle och Plutt på den stora familjen [famlijen],
%eng:	being a little bit abashed Pelle and Plutt say hello to the big family,
*JOT:	som glatt hälsar dem.
%eng:	which happily returns their greeting.
%com:	picture twenty-two

When transformed into CHAT-format, upper case is turned into lowercase, since sentence case is reserved for proper names in the CHAT-syntax (MacWhinney 1995: 9). Moreover, in order to avoid the risk that the CLAN-program treats the punctuation marks (before and) within a clause as delimiters, these have been placed within squared brackets. Finally, spelling errors and typos have been supplemented with the conventional expression (cf. *famlijen* that has been corrected to *familjen* in the example above). Especially in the case of the younger writers, this facilitates the manual, as well as the automatic, analysis of the narratives.

Summary

In conclusion, the Frog story corpus comprises 84 subjects[6] at different ages and they have produced a total of 154 narratives. What can be added is that the total number of words produced by the children and the adult group is 65,920 (if the adults interacting with the children at the earliest ages are also included, the total number of words is 75,905). In Table 3:2, an overview of the Frog story data is presented.

---

[6] The subjects are balanced for sex; the only exceptions are the 3-, and 4-year-old groups where ten are boys and four are girls. Gender differences will not be focused on in this study.

Age group	No of narratives (N=154)	Mode of narration	Recording of narrative	Total no of produced words
3	14	Speaking	Audio	4,231 (9,917)[3]
4	14	Speaking	Audio	3,914 (7,161)
5	14	Speaking	Audio	3,658 (4,710)
9	14	Speaking	Video	5,032
	14	Writing	Paper and pencil	3,790
12	14	Speaking	Video	5,958
	14	Writing	Paper and pencil[1]	4,852
15	14	Speaking	Video	6,019
	14	Writing	Keyboard	5,257
adults	14	Speaking	Video	10,971
	14	Writing	Keyboard[2]	12,238

[1] Three narratives were written on keyboard.
[2] Eight of the adult subjects were video-recorded during the writing session.
[3] Numbers in parenthesis include parents'/research assistants' production.

Table 3:2: Overview of narrative data in the Frog story corpus.

When examples from individual narratives are given in the thesis text, specific codes with the format as "LIA09SP" are used as a reference. These consist of a three-letter subject ID (e.g. "LIA"), age in years ("09" means 9 years of age), and information about mode of production (either speaking, "SP", or writing "WR"). The age code 20 is used for all adult subjects, regardless of their actual age. In a specific type of analysis when the three-year-old narrators of the frog story are compared to the three-year-olds playing with a doll house, the narratives are referred to as, e.g., "FROG.CAL03SP" for clarity.

Despite the fact that the data has been collected for partly different purposes and in relation to different projects, the corpus is homogenous and suitable for developmental analyses of the use of speech projections. However, I would like to make some comments concerning the behavior of the adults

listening to, and/or interacting with, the children producing frog stories in the youngest age groups. Berman & Slobin (1994) report that in their data collection procedure, the adult interviewers were strictly instructed not to prompt or influence the children as to use a particular form of expression. Moreover, only narratives where a plotline was constructed by the children on their own were included in the database. However, in the database of this thesis, there are cases where it is actually adult scaffolding that helps the child keep going and construct a plotline. This is especially true of the 3-year-old group, where the adults produce more words than the children do (although not significantly more, see Section 4.2.1). In the 4-year-old group, the children produce slightly more words than the adults do, while in the 5-year-old interactions, nearly 75% of the words are produced by the children. The fact that the recordings in Berglund & Eriksson's database were made by the respective parents in the home of the child and with no researcher present, made the amount of feedback and scaffolding by adults difficult to control. Indeed, to not be allowed to interact with her child in the way they both are used to in a book-reading situation, may be a strange situation.[7] Yet, the main focus of this study is not to investigate how children manage to construct a plotline on her own (although some analyses are also conducted on this), but rather to examine children's use of direct and indirect forms of speech. The fact that parental scaffolding occurs in the interactions with the youngest children makes it interesting to investigate if the parents prompt their children to use forms of direct and indirect speech. As will be presented in Section 3.3, child use of forms that have been prompted by adults, will be treated separately from forms produced without prompts.

---

[7] Nauclér & Boyd (1997) show that story-telling activities in Sweden are typically collaborative where the child is expected to participate actively in the construction of the story. Moreover, the adult asks questions for the child to answer, in order to engage the child, and to control that the child possesses and/or is focusing on specific information at a certain point in time.

3.2.3   The Doll house play corpus

These data were collected within the project *Long-term effect on mother-infant behavior of extra contact during the first hour post partum* started in 1976 by Jan Winberg and Peter de Château (Winberg & de Château 1982; for project reports, see de Château & Wiberg 1977a, 1977b, 1984; and Wiberg & de Château 1982). The video-recordings of the three-year-olds playing with a doll house with their mothers, were made by de Château in order to do linguistic analyses of the children's language use. This was done in co-operation with Ragnhild Söderbergh (Söderbergh 1982). In addition, the data were used in a follow-up study of the children (Wiberg, Humble & de Château 1989). The corpus used in this thesis consists of 14 of these 38 video-recordings. All mothers and children participants are monolinguals and have Swedish as their first language. Half of the children are girls and half are boys, and they were all between 36 and 39 months (average 38 months), by the time of the recording. The majority of the families represent a relatively well-educated middle class.

The recordings were made in a laboratory. The mothers and children were instructed to sit on a mattress, three meters in front of the video camera with the doll house between them. The mothers were invited to play together with their children with the doll house for about half an hour. The children were also encouraged to pick out dolls from a basket with various dolls in it, in order to re-create the structure of their own family.[8] The doll house had four rooms: a kitchen, a living room with a fireplace, a bedroom with two large beds and one small one, and a bathroom with a bathtub and a toilet. In addition to different kinds of furniture (chairs, tables, a sofa, etc.), there were two tiny telephones in the house. The house had no roof – a detail which makes it easier for the observer of the video tapes to keep track of the child's manipulation of the details of the doll house. Among the dolls the child could choose from were a woman doll, a man doll, child dolls and even pets.

---

[8] More exactly, the mother and child were given the following instructions: "We want you to play together and show what happens on an ordinary day in your family. A video recording of

The recordings were transcribed according to a format described in detail in Strömqvist (1979, 1984). Swedish orthography is used as a basis for the transcription, but with certain adjustments to render certain qualities of spoken language. Features like pauses, simultaneous speech, and stressed words and word strings are noted, as well as if, e.g., a certain tone of voice is used. The identity of the speaker (mother or child) is indicated at the beginning of the first line of every new turn. All transcripts are either machine typed or hand-written. In this way they differ from the other data in the thesis, since the Longitudinal case studies and the Frog story narrative data are all available in CHAT-format on a computer. Therefore, in order to calculate the size of the Doll house data, the number of words produced in each recording (transcript) were approximated by a sampling procedure (in which a subset of the data was counted manually). The total size of the Doll house corpus is then approximated to 22,900 words (7,400 words spoken by only the 3-year-olds).

Age	No of dyads (recordings)	Total no of produced words (children)	Total no of produced words (adults)
3	14	approx. 7,400	approx. 15,500

Table 3:3: Quantitative data of the Doll house play corpus.

The reference codes have the format "DOLL.010-03" (i.e. subject number 10, age three years, from the Doll house corpus).

3.3 Types of analysis

In the following sections I will present the types of analysis that will be made on the data. This section is divided into five parts, that could also be seen as

this play session is going to be made. If possible, we want you to play as long as the video tape lasts, i.e. for 30 minutes" (Strömqvist 1984: 201).

representing five different substudies. Each of these substudies is based on a particular subset of the data.

From Figure 3-4 below it is evident that there are three main types of data, or corpora, involved in the analyses. The Longitudinal corpus covering the ages 19 months to 4 years of age (two case studies); the Doll house play corpus including interactions between 3-year-olds and their mothers; and the Frog story corpus comprising spoken and written narratives produced by subjects ranging from 3 years of age to adults (20 years and older), (cross-sectional type of data).

	The Longitudinal corpus	The Doll house play corpus	The Frog story corpus		
(a)	1				
	2				
(b) ←	3	— 3 —	— 3 →		
	4 ↓		4	(c)	
			5		
			9	(d)	(e)
			12		
			15		
			20↓ ↓		
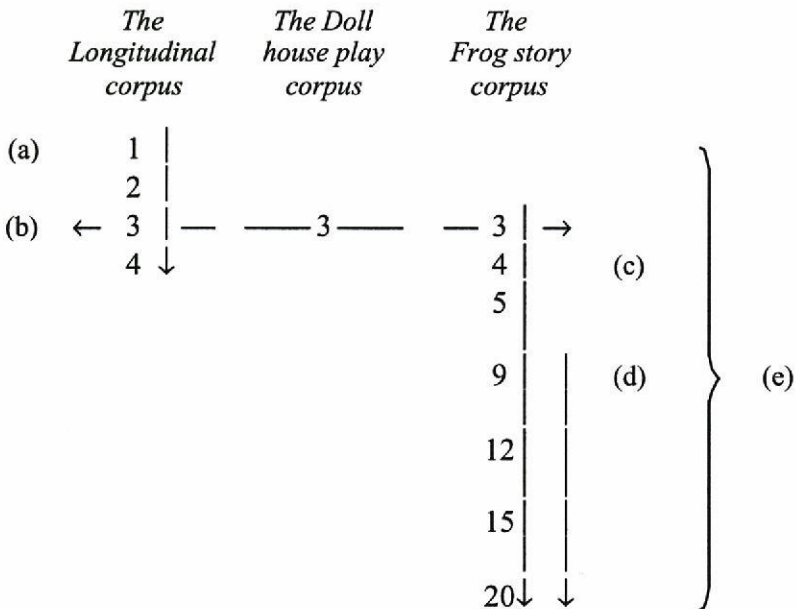
Figure 3-4: A guiding scheme to the substudies. The numbers refer to the ages of the subjects, and the letter indices and arrows refer to the subset of data included in the different substudies.

Index (a), and the arrow leading from age 1 to 4, refers to the substudy that will be presented in greater detail in Section 3.3.1 below. This study has as its base the Longitudinal material. Index (b), and the horizontal arrow, refers to examination of the 3-year-olds and their caretakers. These analyses include

different types of activities (e.g., make-believe play, story-telling), and a presentation of the procedure of analysis is found in Section 3.3.2. In the next section, Section 3.3.3, the types of analysis be applied to the whole corpus of spoken frog story narratives, can be found (see index (c) and the arrow stretching from age 3 to age 20). The spoken frog story narratives by the school children and the adult narrators, will then be compared to their written counterparts, as indexed by (d) and the parallel arrows, (see Section 3.3.4). Finally, in addition to these four substudies, occurrences of free indirect speech in any of the sub-corpora will be investigated (represented by index (e) in the figure). This type of analysis is described in Section 3.3.5.

There will be similar kinds of measurements and operationalizations running through the analyses of the data, resulting in overlaps. This procedure of keeping certain parameters of analysis relatively constant across data types, provides an opportunity to, for instance, compare adults using direct and indirect speech when narrating the frog story to the 3-year-old Tea projecting speech to her dolls. Thus, we will be able to learn what types of forms are used and to what extent, from a long-term development perspective. However, since the types of data, and the specific research questions related to them, also differ to a certain extent, some of the measurements are unique to the particular substudies. By way of example, the Longitudinal case study data are of a different kind and collected under other conditions than the Frog story narrative data. In the first case, two single individuals are followed under naturalistic and interactional conditions, and we are not only concerned with the children's language use, but also with what their caretakers say and what the interaction looks like. In the latter case, the data is cross-sectional, task-oriented, and monological in nature, and there is also written data to examine. Hence, the frog story narratives require a partially different framework of analysis than the case studies. This methodological approach of combining general parameters of analysis with more specific ones, (and making quantitative as well as qualitative analyses), allows for detailed examinations related to different stages of development and different contexts, but in addition, it gives a general

127

and complex picture of the (long-term) development of direct and indirect speech. This is important in order to explore the research questions as presented in Section 2.5.

3.3.1   The Longitudinal case studies

In the analyses of language use by the two young siblings Harry and Tea, and their adult interlocutors, certain questions arise. With respect to the children's language production, we want to know if direct and indirect forms[1] are used at all, and if they are, what kind of forms are used. In addition to this, questions such as when the forms emerge, how they develop, and in what kind of conversational, activity, and narrative contexts they appear are of interest. Are the forms used primarily in speech reporting or in speech projection? Do the children differ in how they use the forms? Moreover, it is important to determine what the particular language input to the two individual children looks like and what the relationship between input and production is. Thus, in addition to examinations of the use by the children, the analyses of the Longitudinal case study material also account for the amounts and types of use of direct and indirect forms by the adult speakers, and attempts at prompting the children to use direct and indirect speech.

The focus of analysis of the children's language use, will be qualitative rather than quantitative. The earliest data of Harry and Tea are from when they are only 19 months old and their mlu-value is close to 1. Gradually, (as is reflected in mlu-curves in Figure 3-2, page 111, and Figure 3-3, page 112), their utterances become longer, the syntax more complex, their pronunciation and discourse skills better, etc. Similarly, the development of direct and indirect speech is regarded as a gradually developing matter. It is hardly the case that spontaneously produced, perfectly well-formed, and distinguishable utterances of direct and indirect speech will suddenly pop up, which would

---

[1] Free indirect speech is not taken into consideration in these analyses. For a specific analysis of these matters, see section 3.3.5 below.

make quantitative analyses straightforward. On the contrary, the forms probably emerge and develop over the course of time. Therefore, careful qualitative analyses of Harry's and Tea's development is carried out, noting when (precursors of) the forms emerge and how they develop, and in what contexts they appear. Authentic examples from the data will be provided for matter of illustration and discussion. The degree to which these candidates of forms are spontaneously produced, or rather imitations of adult utterances produced, (or in some other way prompted or triggered), will also be investigated.

As described above, the data are transcribed according to CHAT-format, which makes possible automatic computer analyses by the CLAN-software (MacWhinney 1995). In the procedure of analysis, all interesting passages (i.e., passages including candidates for speech reporting and/or speech projection and the context in which they occur) are tagged manually in all the Longitudinal data, using the options available within the GEM-program. Running CLAN, then presents lists with passages as in Example 3.1 below, (the example is from the transcript when Harry is 40 months and 21 days old).[2] The qualitative analyses are thus based on what is found in these coded passages (starting with @bg:indir and ending with @eg:indir). The identification and coding of an instance (or candidate) of direct and indirect speech, is based on syntactic and deictic criteria. Indirect speech is operationalized as utterances with the form "X said that __", where the embedded clause has to be a proposition.[3] Framed direct speech typically has the form "X said, __" or "__, X said", where the reported or projected utterance can be a proposition or an animal cry or the like. The same holds for free direct speech ("__!"). (For further prototypical examples and definitions, see Sections

---

[2] The following command was given: GEM +sindir +u *.cha (i.e., all passages coded by means of indir are picked out).

[3] The most common word order of indirect speech in Swedish is as in *flickan sa att hon inte gillade spenat* ('the girl said that she didn't like spinach'). Yet, it should be mentioned that in Swedish, for certain purposes, it is indeed possible to invert the word order: *att hon inte gillade spenat, sa flickan* (*'that she didn't like spinach, the girl said'). However, this construction type is rare, and sounds awkward.

2.4.1.1 and 2.4.2.1). Besides typical linguistic markers, such as the use of a nominal or infinitive clause, transpositions in syntactic person or verb tense, and speech act verbs; changes in tone of voice, gestures and other contextual cues may help determine what is a candidate for direct and indirect speech. This is particularly true in the case of speech projections in relation to make-believe play.

Example 3.1

(*HAR = Harry , *GMM = grandmother on mother's side)

```
-------------------------------------------
***** From file Aspasia HD:Dokument:Avhandling:
Long:03data.Harry:HAR40_21.CHA; line 650.
@bg:      indir
*HAR:     hon ä t [/] tar hon vill ha saker som ja [!] har
%eng:     she ehm takes she wants to have things that
          I [!] have
*GMM:     vill hon de ?
%eng:     is that what she wants ?
*HAR:     ja^a
%eng:     yeah
*GMM:     å va säjer du då [!] ?
%eng:     and what do you say then [!] ?
*HAR:     ja vet inte [=! missmodigt]
%eng:     I don't know [=! discouragedly]
*GMM:     säjer du till Minna då att du vill ha dom sakerna
          själv ?
%eng:     do you tell Minna then that you want to keep
          those things yourself ?
*HAR:     ja^a
%eng:     yeah
*GMM:     å va säjer Minna då då ?
%eng:     and what does Minna say then ?
*HAR:     att hon vill ha dom sakona som ja vill ha
%eng:     that she wants those things that I want
@eg:      indir
-------------------------------------------
```

In the cases of make-believe play, only instances of indirect voicing have been taken into consideration. As was explained in Section 2.3.4, this refers to when the child enacts a play role *through the medium of a toy figure*. Thus, direct voicing, where the child's body "becomes" the play character, has not been included in the analysis. The reasons for leaving direct voicing outside

analysis are theoretical as well as methodological. Theoretically, it may be questioned to what degree direct voicing is really about *projecting* speech (cf. Section 2.4.2.3 where the components of speech projection are described as being typically three: a projector, a projected speaker, and speech projected). Methodologically, problems sometimes arise in contexts of make-believe play when distinguishing utterances in role play (direct voicing) from other types of utterances, i.e., distinguishing cases where the child is playing a role, from cases where they "act as themselves". In analyses of passages with indirect voicing, in contrast, comments about non-verbal acts provided in the transcripts, and inspection of the video tapes, become extremely valuable in the coding process. In order to be coded as indirect voicing, besides having the proper linguistic format of free direct, framed direct, or indirect speech, it should be clearly indicated by the child that the speech is projected onto the doll. This means that the voice should be modified so as to distinguish the doll's speech from the child's, and/or that the child is holding onto the doll, making it move, or something similar. Repetition of utterances in relation to clarification requests has not been taken into consideration (for an analysis of sequences of clarification requests and negotiations of (mis)understandings in recordings of Tea, see Nordqvist 1998c).

As discussed above, it is unlikely that well-formed and complete forms (are the first to) show up, and therefore, forms that may not have (all) these characteristics, but that seem to be forms "under development", are also coded. This may include forms where an element/word is missing, the word-order is reversed, or the like. Moreover, although the reporting clause is left out by Harry in Example 3.1 above, an utterance like *att hon vill ha dom sakona som ja vill ha*, 'that she wants those things that I want', is included in the analysis.

A similar procedure of analysis is applied to the input data. Passages of the adults' own production of direct and indirect forms, and cases where they prompt the children to use these forms, are coded by means of GEM. In the conversation between Harry and his grandmother above (Example 3.1), one instance of what I refer to as 'modeling' is evident, namely, *säjer du till Minna*

131

då att du vill ha dom sakerna själv?, 'do you tell Minna then that you want to keep those things yourself?'. In this clause, grandmother provides Harry with a model utterance (indirect speech). In order to qualify as a 'model utterance', the utterance should have the form as defined above, (i.e., "X said that __", "X said, __", or ("__!")), and be produced by an adult speaker.[4] In the extract above, two examples of 'speech prompting' are also obvious; *å va säjer du då?*, 'and what do you say then?', and *å va säjer Minna då då?*, 'what does Minna say then?', where the second (but not the first) prompt actually results in Harry reporting speech. In order to be coded as a speech prompt, it has to be *speech-related* (Ely et al 1996). This means that the prompt should be an explicit request for a speech report/projection, (thus, *vad **gör** Minna då?*, 'what does Minna **do** then?', does not qualify as a speech prompt, whereas *å va **säjer** Minna då då?*, 'what does Minna **say** then?', does). Speech prompting has a function of engaging the child in conversation, and similarly to modeling, it probably has important implications for the child learning how to report and project speech. Quantitative analyses of cases of modeling and reporting are carried out, and these will be related to the types of conversational context they occur in. Finally, an analysis of when the first instances of model utterances and speech prompts appear is made, and these instances are related to the child's first instance. This type of examination is carried out in order to discuss relationships between the input and the children's own use and development.

3.3.2 Three-year-olds

In these types of analysis the primarily concern is not to follow developmental changes in use of forms of direct and indirect speech by individual children, but to concentrate on a group of children at a particular age stage: three-year-olds. We are then dealing with two main dimensions of analysis; on the one hand, we are interested in relating the three-year-olds' own uses of forms with that of

---

[4] Certainly, model utterances may, and probably are, produced by other speakers than adults, for instance, by siblings. However, in the data of this thesis, the children are only recorded together

the adults' (i.e., the input), and on the other hand, we will further examine the relationship between type of conversational context or activity, and use of particular forms. Due to the general design of the types of analysis (see Figure 3-4 above) there are certain overlaps between the analyses of the individual corpora. In the first type of analysis (as described in the previous section), the Longitudinal corpus is examined. These analyses include the age stages when Harry and Tea are around three years of age and these results are presented in 4.1. The 3-year-olds' telling of the frog story, will be compared to the older subjects performing the same task in a separate process of analysis (to be described in the following section), and these results are reviewed in 4.3. As a consequence of this, the types of analysis presented in the current section (and the results to be presented in Section 4.2) focus especially on the data that are not analyzed elsewhere: the 3-year-olds playing with the doll house and their interaction partners. Still, as will be evident, several comparative analyses concerning use of direct and indirect speech are carried out between the 3-year-olds and the mothers in the doll house play activity, and the 3-year-olds and their caretakers in frog story activity, respectively.

The extent to which speech projections are generally used by the subjects in the two activities will be examined, as will the types of forms. The criteria for what count as speech projections are the same as defined in Section 3.3.1 above. Thus, syntactic/deictic criteria are in operation, as are situational cues (like voice modifications, various non-verbal strategies, and other aspects of the play/narrative situation). Indirect voicing rather than direct voicing is taken into consideration. One example of where this distinction is made has to do with the two little toy telephones that are in the doll house and that the children and their mothers have at their disposal. In those cases where the subjects clearly signal that the dolls are using the telephones and have conversations on the phone, these passages have been analyzed. However, when a child acts as

with adults.

herself and does not act via a toy figure, the passage has been excluded since it is not about *projecting* speech.

The total number of speech projections, and the types of forms, produced spontaneously by the children in the two activities will be calculated, as will the forms produced by the adults (i.e., model utterances). Moreover, adult-produced speech projection prompts will be considered. In order to explore the relative amount of use of speech projections by each subject, the speech projections will be related to the total amount of language produced. Since the types of data are diverse and transcribed and coded according to different standards, (see Section 3.2 on subjects and data), the only measurement that is consistent across the two corpora is number of words. As a consequence of this, *the total number of words included in speech projections* by each subject (children and adults) will be divided by *the total number of words* produced by each subject.

In addition to these quantitative analyses, qualitative analyses regarding the use of the forms is performed. I examine the degree to which the children manage to convey shifts in perspective in relation to speech projections, e.g., between the dolls in the play. Further, I investigate simultaneous packaging of information and marking of aspects of delivery in the speech projections, and how the speech projections are embedded in a narrative play frame. The type of interaction between the children and the adults is also examined in greater detail.

3.3.3 Spoken frog story narratives

General analyses of narrative construction

To serve as a background to the more specific analyses of direct and indirect speech in the spoken frog story narratives, two types of analysis on general narrative structure and development are carried out. The first type of analysis concerns *overall plotline*, that is, to what extent do the narrators construct a story with a global structure. The second one deals with the *use of grammatical*

tense. More specifically, the extent to which the narratives are anchored in a particular tense is investigated. Both types of analysis were also carried out in Berman & Slobin's (1994) cross-linguistic developmental study, and the procedure of analysis applied here is similar to that procedure.

Following Berman & Slobin (1994: 46-50), in order to quantify the construction of the overall plotline of the story, three core components are coded for in each spoken narrative: *I - The onset of the plot*, i.e., the boy's realizing and noticing that the frog is missing. Importantly, the fact that he notices that the frog is missing, must be explicitly mentioned. *II – Unfolding of the plot*; this component concerns the boy's search for his missing frog, and explicit mention must be made of searching (or looking, or calling) for the frog. Moreover, the search for the frog must go beyond the initial start of the search inside the bedroom. *III – Resolution of the plot*, i.e., the boy finds his missing frog. In order for this story component to be coded as present, it is necessary that the frog the boy takes home at the end of the story is explicitly described as the same as or substituting for the frog the boy has lost. In the cases of the 3-, and 4-year-olds, all those instances prompted by an adult where the child mentions either of the components are excluded.

The subjects telling the frog story can choose between telling the story in past tense or in present tense (or to mix them). A narrative is regarded as having an anchoring (or dominant) tense if 75% of all finite verbs in the narrative are inflected in one particular tense (Berman & Slobin 1994: 62).

Quantitative analyses of forms and proportions of speech projections
Which forms of speech appear in the narratives at different ages? What is the total number of instances used in the different age groups, and what proportion of the clauses in the narratives constitute projection clauses? In order to answer these questions, and to conduct appropriate analyses, options within the CHILDES-system (MacWhinney 1995) will be used.

As was described in Section 3.2.2.2, all frog data were transcribed in CHAT. In relation to this description, it was also mentioned that the data was

coded for clauses (defined as "a unit containing a finite verb"). In addition to these codings, the narratives have been coded for cases of free direct speech, framed direct speech, and indirect speech. There are codes for direct quoting included in the CHAT transcription system (MacWhinney 1995: 49). The main code to use is '+"', which is placed at the beginning of the utterance that is directly quoted. A speech introducing utterance (including speech act verb) which precedes the quote, is marked with '+"/.'. If the utterance with the speech act verb follows the quoted utterance, '+".' is placed by the end of the utterance. These codes are employed with some modifications and additions, as will be discussed below.

First, it is important to emphasize again that the utterances are chunked into clauses. Applying the CHAT codes for direct speech to the clause chunked narratives, will yield the following types of codings:

KIA12SP.cha

*KIA:	å så när Snissan kom
%eng:	so when Snissan came
*KIA:	så sa han +"/.
%eng:	then he said
*KIA:	+" schysss tyst på dej
%eng:	hush be quiet
*KIA:	+" få ta de lite lugnt
%eng:	you have to take it a little easy

CAR20SP.cha

*CAR:	+" det här måste undersökas
%eng:	this has to be investigated
*CAR:	sa pojken +".
%eng:	the boy said

In the first example, extracted from a narrative produced by a 12-year-old, the mere speech projected, *schysss tyst på dej få ta de lite lugnt*, 'hush be quiet, you have to take it a little easy', (coded with the '+"' symbol) is introduced by a clause with a speech act verb, *så sa han*, 'then he said', which is marked with '+"/'. Note that the quoted utterance is divided into two clauses. In the second

example, taken from an adult narrative, the clause with the speech act verb (*sa pojken*, 'the boy said'; marked with '+".') follows the quote (*det här måste undersökas*, 'this has to be investigated'; marked with '+"').

Second, the symbol '+"' has also been used for direct quotes that are not preceded or followed by a speech introducing clause, i.e., free direct speech. Below is an example from a narrative produced by a 9-year-old boy.

CHE09SP.cha

*CHE:	och sen hittade pojken en hel grodfamilj
%eng:	and then the boy found a whole frog family
*CHE:	och bland dom fanns den
%eng:	and among them was the one
*CHE:	som han hade haft från början
%eng:	that he had had in the beginning
*CHE:	+" hejdå snälla grodor
%eng:	bye bye nice frogs
*CHE:	+" tack för hjälpen
%eng:	thank you for your help

The quote can either have the form of a clause or not. In the passage above, *hejdå snälla grodor tack för hjälpen*, 'bye bye nice frogs thank you for your help', is chunked into two parts and coded as free direct speech by means of '+"'. Not all quotes in the corpus contain a finite verb, or correspond to a verbal proposition in a strict sense. Cases like *han sa: "grooodan"*, 'he said, "froooggie"', *ha ha sa rådjuret*, '"ha ha", the deer said' exist, and utterances like these have been coded as, and separated into, two parts (i.e., *han sa* and *"grooodan"*; *"ha ha"* and *sa rådjuret*).

Finally, it was also of interest to code for indirect speech. There is no special symbol provided for that in CHAT, but the existing symbols for direct speech were used as a base and applied as in the below example. This means that the symbol '+"' is used also for coding speech that is not directly quoted, but indirectly quoted (in this case, *att dom skulle förvara den bättre så att den inte skulle komma därifrån*, 'that they would keep it in a better way so that it wouldn't be able to escape').

137

3. Methodology

EMI15SP.cha

*EMI:	å då sa dom +"/
%eng:	and then they said
*EMI:	+" att dom skulle förvara den bättre
%eng:	that they would keep it in a better way
*EMI:	+" så att den inte skulle komma därifrån
%eng:	so that it wouldn't be able to escape

The advantage of using codes like those described above, is that all instances coded can be automatically picked out and analyzed by a computer, using the CLAN software (MacWhinney 1995). Moreover, one important reason for chunking the narratives (including the speech projections) into clauses, is to be able to tell how large a proportion in a narrative is made up of clauses involved in speech projections. One fruitful tool in CLAN for these kinds of analyses, is the COMBO. It is possible to use this tool to pick out all utterances/clauses that include the code '+' in combination with ' " '. In addition it is possible to get the context in which the coded utterance/clause occurs:[5]

```
-------------------------------------------
*** File "AspasiaHD:adultcontrols:DATAcontr:ADchat:
SPad.cha:AND20SP.cha": line 63.
*AND:    eh kom närheten av skogsdungen
%eng:    uhm came close to the grove
*AND:    som eh var i närheten av huset
%eng:    that uhm was near the house
*AND:    de [dom] ropade +"/
                      1
%eng:    they called
*AND:    +" grodan var är du
%eng:    frog where are you
*AND:    de [dom] tittade i marken i håler
%eng:    they looked at the ground into holes
-------------------------------------------
```

[5] The command is as follows: COMBO +f +w2 −w2 +s\+\" +u +g1 *.cha. (The +f means that the result of the analysis will be sent to a separate output file; +w2 and −w2 set the size of the context (two utterances precede and follow the key line); +s\+\" picks out all lines/utterances that are coded with a '+'; +u ("unite") processes all files at the same time; +g1 means that the program searches for a string of characters (rather than a word); *.cha means that all files with a name ending with '.cha' are analyzed.)

Hence, what we can see in the clip above is information on filename ('AND20SP.cha'), narrator ('*AND'), the search string ('de [dom] ropade +"/'), the location of the search string in the transcript ('line 63'), and the linguistic context of the search string. In this case, the two preceding and the two following clauses were selected as context by means of the commands '-w2' and '+w2'. Consequently, using this type of procedure, every token of speech projection can be analyzed effectively.

Besides being able to make qualitative analyses of types of forms and speech projection contents, the COMBO-command also generates quantitative information. You automatically get information about the total match of search strings in a file or files, which in this case, more specifically, means that you will know the exact number of speech projection clauses in a file (or files). In order to establish the percentage of speech projection clauses in a certain narrative or narratives, you also need to know the *total* number of clauses. This information will be given when using the MLT (Mean Length of Turn) within the CLAN-software.[6]

Functional analyses of speech projections

The procedures presented above primarily facilitate the quantitative analyses and provide answers to the questions as stated in the beginning of this subsection. However, it is also of interest to discover *how* the forms are employed at different ages, and what functions can be discerned. In order to answer these types of questions, a coding scheme was drawn up using Microsoft Excel. The coding was conducted manually, and as can be seen in Figure 3-5, the scheme consists of ten columns.

In the first column (*Subject*), the subject ID is given.[7] In the *Instance* column the speech projection to be coded is found. An instance of speech

---

[6] More specifically, the command is: MLT +u ("unite all files") +f ("send output to a separate file") *.cha ("process all cha-files").
[7] In the subject ID label 'CAL03SP', (the first cell in the *Subject* column), 'CAL' stands for the name of subject, '03' is the age in years of the subject, and 'SP' gives the information that the

projection is seen as a description of a speech event in terms of framed direct, free direct or indirect speech where the projected speaker and the speech projected are distinguished from the context. Consequently, for example, in the case of framed direct speech, the clause with the speech act verb along with the mere speech projected (the quote) is counted as a speech projection. The demarcation (from the context) of a speech projection is represented by the perspective shifting from/to the current projected speaker to/from another projected speaker, or to/from the narrator's perspective.

Figure 3-5: Scheme in Microsoft Excel used for coding of instances of speech projections in the spoken frog story narratives.

For the youngest age group (the 3-, 4-, and 5-year-olds), in order for a speech projection to be counted and coded as an instance, an additional criterion was that the projections should not be prompted by the adult, but produced spontaneously by the child. In the next column (*Form*), the type of form of the speech projection is coded. As is illustrated in Table 3:4 on the next page, there are seven coding categories to choose from.

example of the speech projection is extracted from the subject's spoken narration (a written narrative is indicated by 'WR').

Coding category	Characteristics	Example
FreeDir	Free direct speech: freestanding direct quotation; no adjacent clause with a speech act verb	"I'm a little mouse."
DirBef	Framed direct speech: direct quotation preceded by a clause with a speech act verb	And then it said, "I'm a little mouse."
DirAft	Framed direct speech: direct quotation followed by a clause with a speech act verb	"I'm a little mouse", it said.
DirEnfr	Framed direct speech: direct quotation interrupted by a clause with a speech act verb	"I'm a little mouse", it said, "and I don't like being disturbed".
Indir	Indirect speech	And then it said [that] it didn't like being disturbed.
Unclear Type	Difficult to decide what type of form (e.g., mixture between forms; unclear if speech act verb is given)	
UnclearIf	Difficult to decide if it is a speech projection at all (e.g., unclear if it is the subject/narrator herself that is making a comment, if the narrator projects speech to a protagonist by using free direct speech; or if it is unclear whether it is quoted speech or quoted thought)	

Table 3:4: Coding categories of forms of speech projections in the frog story narratives.

141

When cases of framed direct and free direct speech in the spoken narratives are analyzed, they are coded for changes in tone of voice/pitch (Voice) or no change (NoVoice) in the *Marking* column. To be coded as Voice, the tone of voice should clearly change pitch in the quotation, so as to mark a shift from the narrator's perspective/voice to the protagonist's perspective/ voice, or vice versa, or mark shifts between protagonists. The reason for including this aspect in the coding scheme, is because the voice can play an important role in certain cases to make clear who the speaker is (see speaker perspectivizing below). This is true especially in cases of free direct speech (as was discussed, for example, in Section 2.4.3.2), although it may play a less important role in framed direct speech, and even less in indirect speech. In addition to this, changes in tone of voice have consequences for the function of vivifying (see, e.g., Section 2.2.2).

Several potential speakers can be seen in the pictures of the frog story. The most evident potential speaker and also the main protagonist, is the boy. However, there are other actors on the stage onto whom speech can be projected, like the dog, the frog, the gopher, the owl, the deer, and even the bees. In those cases where it is possible to decide who the speaker is (by verbal context, voice marking or the like), this identity is coded for in the *Speaker* column. When there is a projected speaker besides those mentioned above, the code Other is available, as is the code Unclear when it is difficult or impossible to decide who the intended speaker is. In most cases the speech projected consists of verbal propositions, however, in some cases animal story characters "speak with animal sounds". Although it may be discussed whether these cases can be said to have propositional content (or rather if they should be regarded as descriptions or interjections), they have been included and coded with Ono (as in "onomatopoeia") in the *Quote content* coding category. In this way, it is possible to distinguish between speech projections consisting of words and speech projections consisting of animal cries (and thereby to exclude all the cases that do not contain verbal content).

Just as there are potential speakers visible in the pictures of the frog story, there are also certain scenes that are more likely than others to result in the production of speech projections. For example, in several pictures you can see the boy calling for the frog (the boy has his hand cupped around his open mouth), in picture 20 you can see the boy holding his finger to his mouth as if he tells his dog to be quiet, and in the final picture you see the boy waving goodbye with his mouth open as if he is shouting something. It is of interest to see which scenes trigger most speech projections, and which ones rarely do. Moreover, it is valuable to see if the narrators are bound strictly to what can be seen in the pictures, or if they go beyond what is depicted and "invent" speech events. These kinds of aspects are coded for in the *Pict. matching* category.[8] There is also a column (column E) called *No of Q-clauses*. In relation to the presentation of the COMBO and MLT procedures above, the possibility of automatically calculating percentages of speech projection clauses out of the total number of clauses, was mentioned. In these calculations, the mere speech projected (depicted/described), i.e., the quotes, are not distinguished from the clauses with the speech act verb. In the column of No of Q-clauses, only the number of clauses representing depicted and/or described speech in each speech projection instance, is noted. Besides being able to distinguish the two types of clauses from each other (i.e. clauses with the speech act verb and clauses with the mere speech projected), this procedure also makes possible the calculation of the average length of clauses involved in quotes per speech turn.

---

[8] In order to pick out those scenes that are most likely to result in speech projections, (i.e., where you can clearly *see* that either of the protagonists is performing a speech act), I asked twenty-one linguistics students to perform judgments. Those scenes/pictures that were judged by 90% or more of the students to visualize a speech act, have been picked out and received a special coding label in *Pict. trigger*. These are, in chronological order, pict 3: The boy discovering the frog's disappearance; pict 4: the boy looking in the boot; pict 5: the boy's call out of the window; pict. 8: the boy's (and the dog's) call out into the woods; pict 9: the boy calling into a hole in the ground; the dog barking by the beehive; pict 14: the boy calling from the rock; pict 20: the boy hushing his dog; and pict 24: the boy and the dog waving goodbye. It should be mentioned, however, that although the judges were told to be strict rather than generous in their interpretations and judgements, almost every picture was judged by (at least) someone to represent a speech act. This means that this story and these pictures are likely to generate dialogues, but also that you often infer more than can actually be seen in the pictures (by virtue of the depiction of an open mouth, for instance).

It may, for instance, be expected that the older the narrator, the longer the quotes.

Two types of functions are coded for and these functions are of particular significance for how the narration will be experienced by the receiver of the story: *Speaker perspectivizing* and *Distancing*. The first type of function is related to Strömqvist's (1996) concept 'perspectivizing' that was described in Section 2.2.2 as an array of information being structured from a certain point of view. In the analyses here, I use perspectivizing in a more specific and narrowed sense in that I examine the extent to which the narrator *makes clear who the intended projected speaker is and/or successfully conveys shifts in speaker perspective* (e.g., from narrator to story character, or between story characters), when using speech projections. Two coding possibilities are employed: success in the speaker perspectivizing process (PerspSucc) or failure in this process (PerspFail). In order to convey shifts in speaker perspective successfully, a framing clause including a speaker identity and/or a speech act verb is not necessary, but the coder listening to/reading the narrative should be able to infer from the context who the projected speaker is (e.g. by means of changes in pitch). When this is not made clear, and when the receiver of the narration has problems in tracking speaker identities, changes in perspective, or even fails to grasp that the narrator is actually intending to project speech, then the speaker perspectivizing function has to be seen as having failed. Thus, when the type of form is coded as UnclearIf (i.e., when I as a coder/receiver have clearly not been able to decide if this should be regarded as a speech projection),[9] and/or when it has not been possible to decide Speaker identity (Unclear), this is coded as speaker perspectivizing failure (PerspFail). In all other cases, the speaker perspectivizing process is seen as successful (PerspSucc).

---

[9] It may be critically discussed whether a case which is unclear if it was at all intended by the speaker to be a speech projection, should be included in the analyses. However, I argue that it is of great interest to note that some narrators may have problems in making these distinctions clear. Admittedly (and this could indeed be an argument against the notion of "failure"), such

Turning to the function of distancing, I discussed in Section 2.4.2.3 that the projector not only has freedom and control over who will say what in a frame of fiction, but that she can also elaborate with matters of distance between the narrator/projector and the projected speakers. On the one end of the scale, in the cases of free direct speech (especially in combination with voice modifications), the distance between the projector and the projected speaker is decreased and the narration made more vivid. Elsewhere (Nordqvist 1998a) I have compared this to the process of acting in a stage play, where the actor plays the part of a fictional character with vivid realization. On the other end of the scale, in the cases where indirect speech is used, the projector role is upgraded and the distance between the projector/narrator and the projected speaker/protagonist increased. The projector then describes or reports, rather than enacts, the events. To continue the analogy with a stage play, by using framed direct speech, the projector takes the director role. Although the direct quoting contributes to vividness and the distance to the characters decreases, the projector is the one controlling and directing the events (via the framing clause). In this way, in terms of distancing, framed direct speech falls in between free direct speech and indirect speech. In the coding scheme there are direct relationships between the coding of distancing and the codings found in the form category. This means that free direct speech is coded with MinDist (as in "minimal distance"), the different forms of framed direct speech with MidDist (i.e., in the middle of the distancing scale), and indirect speech is coded as MaxDist ("maximal distance").

3.3.4 Cross-modal analyses

One of the research objects for this thesis is to compare the spoken narratives with the written narratives, and to investigate language development in school children (cf. research questions in Section 2.5). In order to carry out the cross-

---

uncertainty of perspective may also be a conscious stylistic choice by the narrator (cf. the discussion of free indirect speech in section 2.4.1.2).

modal comparisons, several of the analyses of the spoken narratives are replicated with the written ones. Since the written texts are put into the same transcription format (CHAT) as the oral monologues (with certain adjustments, see 3.2.2.2), this is easily administered. Accordingly, it is possible to make quantitative analyses as to what proportion of the clauses consists of speech projection clauses, and what the distributions of the different form categories are. In addition to this, we can examine the length of the narratives in the two modalities, and if the order of production (i.e., speaking first and writing afterwards, or vice versa) has any effect on use. All these aspects are investigated both from a cross-modal (speech versus writing) perspective, and from a developmental perspective.

Basically, the same coding scheme and criteria that are used for the analyses of the spoken narratives are employed for the written narratives. This means that we will be able, from a cross-modal and developmental perspective, to examine parameters as *Form, Speaker, Quote content, Pict. trigger, Speaker perspectivizing*, and *Distancing*. The fact that the spoken and written production process and conditions differ in several respects, and that the structuring and packaging of linguistic information looks different in speaking and writing, have been extensively discussed above (e.g., Sections 2.2.3 and 2.4.3). From a developmental point of view, it is of interest to see how speech projections are dealt with by school-age writers. The conventional ways of marking quoted (direct) speech in writing in Swedish is by using quotation marks enframing the quote (as in English), or by introducing the speech line with a hyphen. Consequently, the English example given in Example 3.2, could in written Swedish look like either Example 3.3, or Example 3.4.

Example 3.2
The lady asked, "Do you like chowder?"
"Yes I do", I answered.

Example 3.3

Damen frågade: "Gillar du fisksoppa?"
"Ja det gör jag", svarade jag.

Example 3.4

Damen frågade:
  - Gillar du fisksoppa?
  - Ja det gör jag, svarade jag.

Typically, the clause with the speech act verb preceding the quote, ends with a colon.[10] When a line of dialogue is introduced by a hyphen, this line should (but need not) be indented.[11] Both types (as represented in Example 3.3 and Example 3.4) are currently in use in Swedish, however (and importantly) the form as described in Example 3.4 is primarily the preferred one and the form taught in school.

For natural reasons, voice modifications cannot be coded for in the *Marking* column for the written narratives. Rather, since quotation marks and hyphens are the conventional and most obvious means for marking off, and signaling, direct quotes in writing, this is desirable to code for. Consequently, the coding scheme as presented in the previous section for the spoken data, is here adjusted to the conditions of writing. To be coded as Quote in the *Marking* column, the speech projected needs to be marked conventionally, either through hyphens (introducing the speech) or quotation marks. Otherwise the code NoQuote applies.

Another aspect of information structuring to be investigated in the written narratives, is the type of speech act verb used. As established above, there tend to be more types of verbs in written texts than in spoken, and *say* seems to be the most common verb type in both speaking and writing, (Caldas-Coulthard 1994, Tannen 1986; see page 79). In speaking, it has also been found that

---

[10] In cases of free direct speech, the marking of the quote looks the same as in framed direct speech, apart from the clause with the speech act verb which is by definition not there.
[11] *Svenska skrivregler* (1991).

among children, *say* is the most frequent type of verb of saying (Ely & McCabe 1993, Goodell & Sachs 1992, Hickmann 1993). However, since research has rarely been carried out concerning use of types of speech act verbs in children's writing, it is of interest to examine this aspect in the current material. Moreover, attempts to package (and describe) prosodic information in writing, for instance, by means of upper case letters, exclamation marks, reduplication of letters, and the like, are included in the analyses.

Finally, in order to establish if there are any modality-specific traits concerning use of forms of direct and indirect speech, (that are not necessarily age dependent), some analyses are carried out comparing *all* spoken narratives (irrespective of age of the producer) with *all* written narratives. More specifically, three aspects are investigated. First, earlier research points in the direction that free direct speech seems to be more common in spoken dialogue than in written texts (Tannen 1986, Chafe 1982). Whether this is also the case with this data is examined. Second, analyses of the positioning of the clauses with the speech act verb in relation to the quotes is carried out, as are analyses of cases where the quote is *interrupted* by a clause with a verb of saying. This latter type has been argued to be typical to written texts (e.g., Tannen 1986). Lastly, whether or not the proportion of speech projection clauses is greater in speaking than in writing, or vice versa is examined. Since the methodology of comparing a spoken and a written performance of the same material by the same individual is rarely applied, not much is known about this issue. Indeed, the possibility of including this within subject design, provides us with valuable information about how linguistic production and information structuring vary in correlation to mode of production.

3.3.5   Free indirect speech

The fifth and final substudy to be carried out focuses on one type of form, free indirect speech:

Coding category	Characteristics	Example
FreeIndir	Quotation where deictic elements belong to reporter or projector as in indirect speech, but otherwise it behaves as a direct quote.	*With a framing clause:* He was a little mouse, he said. *Without a framing clause:* He was a little mouse.

This form is coded for throughout all data of the thesis. This type of examination makes it possible to relate the occurrences of the forms to a range of variables: the age of the user, the type of activity and narrative context, and the mode of production (speaking versus writing). The coding is done manually, and all candidates are picked out and analyzed separately. Due to the ambiguous nature of the form (see further below), the analyses are qualitative rather than quantitative.

In order to qualify as a candidate of free indirect speech (Y), the following criterion applies:

a)   Utterance Y is likely to refer to a speech event.

The form may include a framing clause (cf. *He was a little mouse, he said* above) that explicitly informs that it is about speech being reported or projected by means of a free indirect quote. The framing clause may, similarly to framed direct speech, precede, interrupt or follow the quote. If there is no such clause, the discourse context (the activity engaged in, the verbal context, changes in pitch, etc.) acts as a clue to determine if the utterance(s) is/are a depiction(s) of a speech event. As has been discussed previously, free indirect speech may be employed by a narrator exactly, in order to leave it for the listener/reader to decide if it is speech that is reported (or projected) or not. As a consequence of this, ambiguous cases have also been coded as candidates of free indirect

speech. Even cases which may qualify as reported/projected thoughts rather than speech are coded for and discussed.

Moreover, in order to be categorized as a free indirect speech form, the candidate form needs to pass the following test:

b)   If the utterance Y is transformed into a direct quote construction, then the word-order of Y should remain intact whereas the deictic elements need to be shifted.

Additionally, in many cases a similar test where Y is transformed into an indirect form results in changes of word-order and/or that lexical elements need to be deleted or added.

By these criteria no consideration is taken of whether the form is a conscious stylistic choice by the narrator, if it is used colloquially and on par with the other forms of direct or indirect speech, or if it is even partly an "incorrect" form where we may suspect that the narrator does not fully control the deictic adjustments (since this study includes very young children it may, for instance, be the case that we find examples that seem to be mixtures of forms).

In addition to analyses of free indirect speech, a minor (i.e. not exhaustive) analysis of mental verbs will be carried out. This type of analysis is motivated by the fact that the use of free indirect speech refers to something in between (reported/projected) speech and thinking, and the borders are vague. Among other aspects, whether a limited use of speech projections in a particular group of narrators can be explained by a more frequent use of "thought projections" in the same group is examined. Moreover, the ascriptions of cognitions (by means of mental verbs) are explored from a developmental point of view, and the occurrence of such verbs in speaking will be compared to those in writing. The verbs *tänka* ('think'; 'intend'), *undra* ('wonder'), *märka* ('become aware of'; 'notice') and *tycka* ('think', 'experience') are included in the analysis. The two first (*tänka* and *undra*) can occur as explicit verbs of thinking (i.e. included in framing clauses) and they are in these cases treated separately and the same

criteria as are true for framed direct speech and indirect speech apply (i.e., "X thought/wondered that/whether __"; "X thought/wondered, __"; see Section 3.3.1, page 129, for coding criteria of framed speech quotations).

4 Results

The preceding chapter dealt with the description of the methodological framework, and this chapter presents the results of the analyses. The outcomes of the analyses will be presented in the same order as given in the previous section (*3.3 Types of analysis*). Hence, in 4.1, the results from the analyses of language use by Tea, Harry and their interlocutors are presented. In the following section, Section 4.2, results from the analyses of language use by the 3-year-olds and their caretakers across activity types, are reported. In Section 4.3, the outcome of the analyses of the spoken frog stories are surveyed, and in Section 4.4 the results from the analyses of the written frog stories are added. The chapter concludes in 4.5 with a presentation of the results of the analyses of free indirect speech.

Discussions of the outcomes of the individual sub-studies are weaved into the above mentioned presentations. A more general discussion of the results, related to the specific research questions (as stated in Sections 1.1.3 and 2.5), then takes place in Chapter 5.

4.1 The Longitudinal case studies

The first study presented concerns the Longitudinal case study material. In the first subsection (Section 4.1.1) below, the focus is on the two children's (Harry and Tea) own production and use of forms of direct and indirect speech. This analysis leads to an attempt at describing general developmental phases the children go through. The following subsection, 4.1.2, discusses the langauge use by the adults interacting with Harry and Tea, and in the final subsection, 4.1.3, the children's and the adult's use of direct and indirect speech are discussed in relation to each other.

4.1.1   The children's production

It is striking how similar the developmental paths are when comparing Harry to his sister, Tea. However, there are also intriguing differences between the two and their use of direct and indirect forms, as will be evident from the survey of the results below.

In the children's early phase of development, i.e., when they are around 20 months of age and the mean length of their utterances is close to 1, there are hardly any candidates for speech reporting and/or projection. However, at this stage of development, the adults and the children are often occupied with looking in picture books and discussing what they see in the pictures. These books are often about different kinds of animals, and the adults typically ask the children to name them and to imitate the particular sounds or animal cries connected to them:

Example 4.1

HAR22_18.cha

*HAR:	**dä:: [=! härmar]**
%eng:	baa:: [=! imitation]
*MOT:	ja så säjer lammen # bä:: säjer la /.
%eng:	yes that's what the lambs say # baa:: the lambs say
*HAR:	**bää**
%eng:	baa::
*MOT:	jaha bä:: # säjer dom
%eng:	yeah baa:: # they say

In this example from Harry at 22 months and 18 days, Harry is imitating a lamb in a book. Mother attentively confirms Harry's utterances by sentences such as *ja så säjer lammen bä:: säjer la*, 'yes that's what the lambs say, baa:: the lambs say'. The latter sentence is an example of modeling – Harry is provided with a framed direct speech form – which we will come back to in the next section. Interestingly, Harry's mother uses the speech act verb *säga*, 'say', rather than *låta*, 'sound like', implying that animals do in some sense speak. However, in turning to Harry's contributions in the exchange above, we can hardly say that

154

the expressed animal cries are full-blown forms of free direct speech. Instead, they can be seen as some kind of proto-forms to free (or framed) direct speech, and they occur in the material of both Harry and of Tea until age 22 months.

By age 2, with mlu-values between 1 and 2, we find more obvious candidates of framed and free direct speech in the children's language production. However, they are short, incomplete and prompted by the adult.

Example 4.2

HAR24_16.cha

*MOT:	va gjorde dom då trollen ?
%eng:	what did they do then the trolls?
*HAR:	**haj**
%eng:	hi
*MOT:	sa du hej te dom ?
%eng:	did you say hi to them

Example 4.3

TEA27_26.cha

*MOT:	va säjer du då?
%eng:	what do you say then?
*TEA:	**ont**
%eng:	hurts
*MOT:	säjer du de att de gör ont?
%eng:	do you say that it hurts?

Example 4.4

HAR24_16.cha

*MOT:	va gjorde vi då?
%eng:	what did we do then?
*HAR:	**sa fy**
%eng:	said shame
*MOT:	fy sa mamma ja
%eng:	yes shame mummy said

The child utterances in the examples above are responses to questions by the adult, although the adult question in Example 4.3 is the only explicit *speech* prompt (*va säjer du då?*, 'what do you **say** then?'). Moreover, all three

(candidate) speech reports produced by the children are followed by adult confirmations.

These dialogues are conversations about Harry's and Tea's own personal experiences, rather than play. However, in the material there are several examples of different kinds of play, and especially in Tea's case, there are many episodes with make-believe (toy) play. In Section 2.1.3 (*Understanding of minds*) it was established that by age two children may ascribe passive agency to replica objects within make-believe play. This is evident in Tea, since she makes the dolls move around, however, she does not project speech onto them. Generally, indirect voicing on behalf of the children is not common at this age. On several occasions, the mother projects speech onto dolls and disguises her voice, but the children rarely respond to that behavior.[1] Instead, Tea often speaks *to* the dolls (i.e., Tea herself is a speaker).

At 28 months, Harry produces his first well-formed framed direct speech utterance in the recordings. Mother and Harry are talking about when his friend, Sanna, visited him, and Harry and Sanna were throwing toys on the floor in a room upstairs:

Example 4.5

HAR28_02.cha

*MOT:	då ropade Jonni upp
%eng:	then Johnny called upstairs
*MOT:	hörru ni där uppe!
%eng:	hey you up there!
*MOT:	nu får ni inte slänga sakerna!
%eng:	now you can't throw the things around!
*MOT:	ropa Jonni
%eng:	Johnny called
*HAR:	ha
%eng:	uhum

---

[1] In an episode of the recording at 26_18, Harry and his mother are playing with a toy train driven by a doll. His mother occasionally weaves in speech projections in her little chat with Harry, whereas Harry does not. However, suddenly Harry says *hello* (without disguising his voice), and then points to the figure sitting on the train, indicating to his mother who the (projected) speaker is.

*MOT:	ja
%eng:	yeah
*HAR:	**m Jonni gopa så nej kacka saka!**
%eng:	yeah Johnny called like that no toss things!

As can be seen in Example 4.5, Harry forms a quotation (*nej kacka saka!*, 'no toss things') preceded by a clause with a speech act verb (*Jonni gopa [=ropa]*, 'Johnny called'). The utterance is similar to a preceding maternal utterance: *nu får ni inte slänga sakerna, ropa Jonni*, 'now you can't throw the things around, Johnny called'. Thus, there is a model utterance, and although Harry's utterance is not a complete imitation of that utterance, (note, for instance, that Harry uses a synonym to *slänga* ('throw around'), namely, *kasta* ('toss'), and that Harry uses the feedback negation *nej* ('no') where it should be the syntactic negation *inte* ('not')), there are clear affinities between the two. In this phase of development, child-produced direct speech is generally rare, and when it occurs, it is often contingent upon a model utterance or a (speech) prompt by the adult. This is true also for Tea as is exemplified in Example 4.6 below.

Example 4.6

TEA29_12.cha

*MOT:	å va gjorde pappa då?
%eng:	and what did daddy do then?
*TEA:	päng vä # prang väg Bella
%eng:	ran away # ran away Bella
*MOT:	ja för å ta fånga \<stanna stanna Bella ropa han> [=! med hög röst]
%eng:	yeah in order to catch \<stop stop Bella he called> [=! in a high voice]
*TEA:	**tanna Bella p ropa** nu Bella dä min häst pang väg
%eng:	stop Bella d called now Bella there my horse ran away

What is emerging at this age, in both children, is an ability to tell personal narratives, and to report exciting experiences. These accounts are often made in co-operation with the interlocutor, as shown in the extract above. The particular segment in Example 4.6 is a part of a longer narrative about when Tea's own

pony Bella one day got loose and started to run away, and Tea's father in a fairly dramatic way tried to catch the pony. Tea reports this story to, and together with, her mother and the girl uses rich mimicry, lively gesturing, and talks loudly and avidly. As Tea and her mother reach the climax of the complication (cf. Labov and Waletzky, Section 2.3.1), they raise their pitch level and use framed direct speech.

Example 4.6 illustrates yet again what was described in relation to Example 4.5. Tea's framed direct speech utterance very much resembles the immediately preceding mother-produced utterance of direct speech. However, not all cases of framed direct speech are (partly) copies of adult utterances. For instance, both Tea and Harry spontaneously report what the cows sound like by using direct speech (*däje säje bö:ö*, 'that one says booh', HAR30_10; *den tossan säjer muh::*, 'that cow says moo', TEA34_21). At 33 months, Tea tells her mother that porridge has a nice taste: *ja inte säjer de ä ne äckli, de ä dott ja säjer*. Literally translated, this sentence reads 'I not say it [the porridge] tastes bad, it tastes good I say', and includes two examples of direct speech ("X say, __", and "__, X say").[2] However, Tea has problems with the word-order. The correct Swedish word-order should be *ja säjer inte de ä ne äckli, de ä dott säjer ja*, 'I say not it tastes bad, it tastes good say I'. That Tea falls short in this grammatical matter, is not restricted to direct speech, but is typical for her speech at this age.

Wolf et al (1984) report that children begin to ascribe intentions to figures in play with small replicas around the age of 31 months (cf. also the discussion on the development of children's understanding of minds in Section 2.1.3). In Tea, it is evident at this age that she is still restrictive in projecting speech onto dolls. In the recording at 31 months and 22 days, there is a long episode where Tea and her mother are furnishing and playing with a doll house. Tea and her mother hold onto the dolls and make them move. Mother occasionally projects

---

[2] The first case could also be seen as an incomplete variant of indirect speech, where the subordinating conjunction *att* ('that') is missing: *ja inte säjer* att *de ä ne äckli*, 'I not say *that* it tastes bad'.

speech to the dolls and modifies her voice (e.g., speech projection to a toy Santa: *å den här soffan vill ja ha [=! grov röst]*, 'oh, I'd like this sofa [=! with a rough voice]'), and prompts Tea to project speech (*va säjer flickan då?*, 'what does the girl say then?'), but Tea herself rarely performs indirect voicing. However, in Example 4.7 an instance of Tea projecting speech onto a doll is given.

Example 4.7

TEA31_22.cha

```
*TEA:    de hä:: ä en thlicka
%eng:    this is a girl
*MOT:    jaha
%eng:    uhum
*TEA:    tack ett tet
%eng:    thanks a present
%act:    pretends that the Santa gives a Christmas gift to the girl doll
*TEA:    tack ja ta denne # i tack
%eng:    thank you I take this thank you
```

Between the ages of 3 and 4 years, language development proceeds quickly in both children. Harry sometimes uses indirect speech and weaves in these forms in rich personal narratives. He also enjoys telling stories of events that he has made up. Below is first an example of indirect speech where the subordinating conjunction *that* is missing, whereas the second example is a full-blown indirect form. Both utterances are from an episode recorded at 36 months, where Harry tells his mother about the adventures of Harry's make-believe friend and Harry himself. There has recently been a burglary in the house of the make-believe friend and Harry reports:

Example 4.8

HAR36_26.cha

```
*HAR:    sa till lisen våran tjuv äm ä tog alla tj m alla tjuffona
         [= smörgåsarna] all en deke
%eng:    said to the police our thief ehm took all the sandwiches all a ladder
```

*HAR: **ä han sa att tjuven va [/] va m va dum**
%eng: ehm he said that the thief was mean

(Yet another example of Harry's use of indirect speech is found in the dialogue excerpt in Example 3.1, page 130). Tea, in contrast, has some problems with this particular form. She seems to avoid indirect speech and instead prefers to use framed and free direct speech. At 35 months she omits, among other things, the necessary subordinating conjunction *om* ('if'):[3]

Example 4.9

TEA35_07.cha

*TEA: **dom sråga häst sett nt n halsbandet**
%eng: they asked horse seen the necklace

At 41 months she responds to an adult speech prompt with a grammatically correct indirect form, and at 42 months the first spontaneously correct indirect form is recorded in Tea (*ja sa att dom e arga dä*, 'I said that they are angry there').

Forms of direct and indirect speech more frequently appear in this phase of development than they did in earlier phases. In addition to this fact, there is an increasing degree of sophistication with which the forms are employed. It has already been mentioned above that Harry enjoys telling personal narratives and stories about his make-believe friend. To this should be added Tea's extraordinary capability by the age of three to organize make-believe play. She shifts perspective between the real world including Tea and her mother, and the dolls' world; she switches perspectives among different characters (by indirect voicing); and she narrates within the play. The short extract in Example 4.10 below illustrates several of these aspects. In this dialogue, Tea and her mother play with little figures and dolls and doll house furniture including a little boat. What is evident, among other things, is that Tea controls indirect voicing (by

---

[3] An alternative interpretation may be that Tea uses a framed direct speech construction rather than an indirect: *dom fråga häst: "(har du) sett halsbandet?"*, 'they asked horse, "(have you)

means of changes in tone of voice, gestural behavior, and linguistic forms), switches between the play-stage (indirect voicing) and the real world (some speech is directed to her mother), and she shifts verb tense in the final utterance to signal "narrator's voice". Moreover, Tea is the one controlling and organizing the play, while Tea's mother is fairly passive and non-prompting.

Example 4.10

TEA35_07.cha

*TEA:	**ja vi åka mä^ä**
%eng:	I want to go too^oo
%act:	tries to place the figure behind the doll in the toy boat
%com:	indirect voicing
*TEA:	**man måste sn man säja man vill åka mä då # man inte säjer t då man kan eh få inte åka**
%eng:	one must say one (if) one wants to join the boat # (if) one doesn't say then one cannot go
%com:	comments about the play, directed to MOT in reality; indirect speech with elements missing
*MOT:	nä man får fråga fint
%eng:	no you have to ask nicely
*TEA:	ja^a sint
%eng:	yes nicely
*TEA:	**a # ja så xxx [=! med klen röst]**
%eng:	a # I so xxx [=! in a weak voice]
%act:	pretends that the figure asks the girl in the boat
%com:	indirect voicing, changes in tone of voice
*TEA:	nu hon fick inte åka mer
%eng:	now she wasn't allowed to ride anymore
%act:	takes the figure away
%com:	comments about the play, note the change of verb tense

Thus, Tea skilfully switches perspective in her play by means of speech projections at age three.

Harry more often reports speech than projects speech. He uses framed direct and indirect forms in personal narratives, and he also manages to shift perspectives between reported speakers. In the passage below (Example 4.11),

---

seen the necklace"', rather than *dom fråga häst (om han hade) sett halsbandet*, 'they asked horse (if he had) seen the necklace'.

Harry tells his mother about when he visited an amusement park together with his father, and they tried a roller coaster. The utterance marked with bold includes a reported dialogue exchange between a couple of boys getting on the roller coaster (who are asking Harry's father if Harry is frightened), and Harry's father (who confirms that Harry is indeed frightened). Some (deictic) elements are missing or wrong (*sa*, 'said', probably should have been *frågade*, 'asked', and *den här pojken är rädd*, 'this boy **is** frightened' is more eligible than *den här pojken var rädd*, 'this boy **was** frightened'), but he otherwise manages the speech reporting and the changes in perspective well.

Example 4.11

HAR39_11

*HAR:	\<å\> [\<] [/] å då titta nåra pocka ba [/] bak
%eng:	and and then some boys looked back
*MOT:	titta dom bakåt ?
%eng:	did they look back?
*HAR:	**ja å [/] å sa till min pappa om å Harry va pa rädd å [/] å då sa pappa ja^a ä Harry ä ha de hä pocken va rädd för tåg**
%eng:	yeah and and said to my dad if Harry was scared and and then Daddy said yes Harry is he is yes this boy was frightened of trains
*MOT:	jaha å då så berätta din pappa de
%eng:	uhum and then your daddy told them
*HAR:	ja^a
%eng:	yeah
*MOT:	att du var rädd
%eng:	that you were scared

At 43 months, Harry reports dialogue in a similar manner when he reminds his mother of what happened last Christmas:

Example 4.12

HAR43_01

*HAR:	**m ä å så pekade du på soffan: #**
	- hä ä min plass
	- nä de e min [!] plass

%eng: and then you pointed to the sofa,
              "here is my place"
              "no that is MY place"[4]

Strictly speaking, the speech reports in Example 4.12 are of the free direct type (the first clause does not include a speech act verb). Note that Harry makes a short pause (marked by "#") before starting to quote.

Three months later he makes a similar speech report and he clearly and sophisticatedly marks by means of clauses with speaker identities and verbs of saying (*säjer hon*, 'she'll say', and *säjer ja*, 'I'll say'). Note, however, that Harry's introductory clause is contingent upon an adult model.

Example 4.13

HAR46_23

*GMM:	va tror du Camilla säjer då då ?
%eng:	what do you think Camilla will say then?
*HAR:	**då # säjer hon:**
	- vems e den mommo
	- den e min, säjer ja då
%eng:	then she'll say,
	"whose granny is that"
	"that's mine", I'll say then

Elsewhere I have discussed the fact that direct and indirect speech are actually speech-about-speech, and speech-within-speech. Thus, they have a "built in" metalinguistic component, and in using the forms, a metalinguistic act is carried out in a sense. As has been shown in this section, the children control this metalinguistic activity from a fairly early age. In addition to this, there is evidence of conscious and explicit reflections over speech use in the data of three-year-old Harry and Tea. At age 36 months, Harry has a long and serious discussion with his mother (within play) concerning what a tractor probably can and cannot say. Similarly, at age 46 months, Tea reports to her grandmother that she has been to an animal store where there was a speaking

parrot. Tea reports that she had indeed planned to say something to the parrot, but she could not find anything suitable to say at that moment, and when she *did* find something to say, then they had to leave the store, so she was not able to talk to him, etc. More evidence of a budding awareness of the role of speech reporting and projection comes from Harry at 37 months. During this recording Harry and his mother "read" three different fairy-tale books, and Harry continuously poses the question: *å va säjer X då?*, 'and what does X say then' (Tea also has a similar strategy at 37 months). It seems that he has figured out, not only that the typical structure of fairy-tales contains much dialogue and character speech, but also that speech projecting and reporting has an important informative, narrative and plot-advancing function.

From the qualitative analyses above, it is clear that Harry and Tea develop similarly in many respects. This is clear also from the summarizing Table 4:1 below, where I distinguish four main developmental stages that both children go through. Yet, there are some differences as well, one the most important of which is that Tea gradually becomes a very proficient organizer of, and narrator within, make-believe play and controls indirect voicing, whereas Harry often makes (speech) reports about his own personal experiences. Since Tea uses quite a bit of indirect voicing, she also produces a considerable amount of free direct speech. However, she avoids indirect speech. Harry, in contrast, often uses indirect speech. Taking into account the fact that Harry is a boy and Tea a girl, these results conform to what Ely and McCabe (1993) found about English-speaking 4- to 9-year-olds where girls were more likely to use direct speech and less likely to use indirect speech while boys used less direct and more indirect speech.

Thus, despite the fact that there are similarities in the language used by the two siblings Harry and Tea, there are some differences in the children's production. In the recordings with Harry and Tea, the same adults (i.e., their mother and their grandmother) appear and a question raised is whether the

---

[4] The utterances are marked by means of quotation marks here in order to make clear the dialogue exchange.

adults employ the same strategies in relation to the two children. It is then of interest to turn to the input aspect of the longitudinal study.

	HARRY 19-47 months	TEA 19-47 months
I	*19-23 months (mlu 1,03 – 1,14)*	*19-23 months (mlu 1,09 – 1,31)*
	• proto speech projection/reporting (mimicking animals)	
II	*24-27 months (mlu 1,24 – 1,8)*	*24-27 months (mlu 1,14 – 2,15)*
	• emerging direct forms, but short and incomplete • needs prompting	
III	*28-35 months (mlu 2,08 – 3,27)*	*28-34 months (mlu 2,63 – 3,31)*
	• sometimes well-formed framed direct speech, but often imitations • first personal narratives, prompted	
		• prompted projections of speech onto dolls
IV	*36-47 months (mlu 4,25 – 5,05)*	*35-47 months (mlu 4,50 – 5,33)*
	• well-formed framed direct speech • dialogue exchanges	
	• indirect speech • speech reporting • narrative construction (especially in personal narratives)	• free direct speech • speech projections • voice modifications • narrative construction (especially in make-believe play)

Table 4:1: Summary overview of developmental phases in Harry (19-47 months), and Tea (19-47 months); framed direct, free direct, and indirect speech, and contexts of use.

165

4.1.2 Language input

The recorded and transcribed input data to Harry and Tea, respectively, are of a similar size. The number of adult utterances is around 12,500 in the recordings of Harry (exact figure: 12,494), as well as in the recordings of Tea (12,630). From these figures together with the total number of speech models (i.e., instances of adult produced forms of framed direct, free direct or indirect speech) and speech prompts (i.e., explicit requests for speech reports/ projections), it is evident that speech modeling and prompting are relatively infrequent:

	models	prompts	total
HARRY	44	28	72
TEA	60	36	96
total	104	64	

Table 4:2: Number of speech report/projection models, and prompts for speech projection/reporting in the input to Harry and Tea.

This table indicates that there are more instances of speech report and/or projection models than prompts for speech reports and/or projections in the language input to Harry and Tea (104 vs 64). Moreover, there are more instances of speech modeling in the input to Tea than in the input to Harry (60 vs 44), and Tea is also more often prompted for speech reporting and/or projection than is Harry (36 vs 28).

Concerning the types of forms of the models, a form in which a quote has a word order characteristic of a main clause (similar to direct quotes), but is introduced with the complementizer *att*, 'that' (similar to indirect speech), is found to be used by the adults in addition to framed and free direct speech, and indirect speech. [5] In the summarizing figure below over number and types of

---

[5] Four cases of this type were found in the data: *då sa pappa att de får du inte göra*, 'then Daddy said that you can't do that', (mother to Tea, 29 months); *eller också säjer dom att # de e inte vår hund*, 'or maybe they say that # that is not our dog', (grandmother to Tea, 40 months); *men då sa Tomten att den säcken den finns julklappar i ti Harry*, 'but then Santa said that in that sack there

forms in the input to Harry and Tea, this form is regarded as a variant of framed direct speech (cf. discussion in Section 2.4.1.1, Example 2.3) and included in this category.

Framed direct speech:	62
Free direct speech:	30
Indirect speech:	12
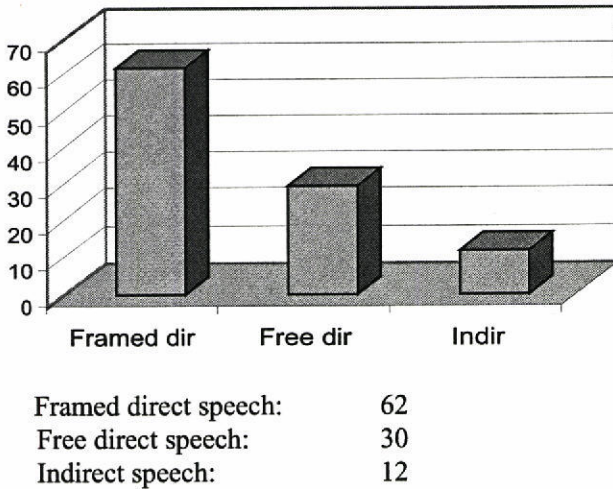
Figure 4-1: Speech models: the total number of instances of types of forms of speech in the input to Harry and Tea.

Framed direct speech is the most common form in the input to Harry and Tea (n=62), to be followed by free direct speech (n=30), and indirect speech (n=12).

The Longitudinal data contain a wide range of different types of activities and it makes little sense if the forms are not related to contexts of use. When examining the occurrences of speech modeling and prompting in the data, they are all found to fall within either of four different types of conversational contexts. The types of contexts – *make-believe (play)*, *book reading*, *personal narratives*, and *habitual and hypothetical topics* – are described and exemplified in the table below.

are presents for Harry', (grandmother to Harry, 31 months); and *de brukar du allti säja när vi e ute å åker mamma att ja vill ha en egen kyrka*, 'you always say like that when we are out in the car Mummy that I'd like a church of my own', (mother to Harry, 37 months).

Type of context	Characteristics of the type of context	Typical function of speech	Authentic (translated) examples
Make-believe (play)	Playing make-believe with toys or dolls; production of fantasy stories.	Speech projections	"More food!", the doll says.
Book reading	The adult reading a book to the child; the child and the adult reading together; discussions related to a particular book or story.	Speech projections (reporting)	"Have a look at my nice car!", Max says.
Personal narratives	Reporting of (specific) personal experiences.	Speech reporting	Then Daddy called, "pull yourself together now!"
Habitual or hypo-thetical topics	Coversations related to habits of the child, and hypothetical reasoning. Abstract reasoning where personal experiences and knowledge about the world are discussed and generalized, without referring to a *specific* event experienced.	Speech reporting and/or Speech projections	Sometimes you even say, "thank you dear Mummy!" [related to habits of the child] If Santa comes, are you going to ask him if he has a fire engine? [hypothetical reasoning]

Table 4:3: Types of conversational contexts; characteristics of the contexts and of speech function, and authentic examples from the (input) data of the Longitudinal case study data.

The narratives produced within the first category, make-believe (play), are fictional (see Section 2.3.1). Instances of direct and indirect speech are typically of the speech projection type. The next category, book reading, is a more complex matter from the point of view of speech projections and reporting. Although children's books are typically a product of pretense from the point of view of the author of the story, the status of the *reading aloud* of the story may be discussed. In a strict sense, the adult or the child reading the

book can be said to report the speech projections. However, since the reader of the story also becomes the narrator of the story in a sense, the reader (narrator) also has an active role as a projector of speech to the characters. Use of direct and indirect speech in the third category, personal narratives, is typically about speech reporting (although the degree of faithfulness as regards the form and content of an earlier (speech) event may vary (cf. discussions in 2.4)). In the final category, habitual or hypothetical topics, there is typically no *specific* (speech) event to relate to. In this sense it can be said to be about speech projections. Yet we can talk of speech reporting since the conversations are grounded in real experiences of the child (rather than in frames of fiction). This shows (as I pointed out in 2.4.2.3) that the division of speech reporting and speech projections should be seen as gradual and overlapping rather than definite and mutually exclusive.

Relating the forms to this four partite division, the forms used in the model utterances are distributed as follow:

	Make-believe (play)			Book reading			Personal narratives			Habitual & hypothetical topics		
	framed dir	free dir	indir	framed dir	free dir	indir	framed dir	free dir	indir	framed dir	free dir	indir
Har	2	0	0	12	2	0	8	0	2	9	3	6
Tea	11	24	0	14	0	3	3	0	1	3	1	0
tot	13	24	0	26	2	3	11	0	3	12	4	6

Table 4:4: Distribution of forms of speech over four types of conversational contexts; input to Harry and Tea.

From Figure 4-1 above (page 167), it was clear that framed direct speech is the most frequently employed form in general, and Table 4:4 shows that framed

direct speech is used by the caretakers in all four types of conversational contexts. Figure 4-1 also showed that model utterances with the form of free direct speech are fairly frequent. In the table above, however, we see that almost all of these cases (24 out of 30) are from one type of activity – make-believe (play) – and in the input to one of the children (Tea). All these 24 cases of free direct speech are from make-believe play contexts, in which Tea and her mother or grandmother are playing with dolls or toy figures. Free direct speech is rarely used by the adults in situations of book reading, relating of personal experiences, or in discussions of a habitual or hypothetical nature. Indirect speech, in contrast, is found in exactly these three latter contexts, whereas no instance is found in make-believe (play).

The numbers presented in the table above indicate that there are some differences between the types of contexts in which the children are provided with model utterances. If prompts for speech report/projection are also included in the analysis of contexts of use, the following figures appear:

	Make-believe (play)		Book reading		Personal narratives		Habitual & hypothetical topics	
	models	prompts	models	prompts	models	prompts	models	prompts
Har	2	4	14	5	10	11	18	8
Tea	35	8	17	20	4	6	4	2

Table 4:5: Distribution of instances of speech report/projection models, and prompts for speech reports/projections in four activities; input to Harry and Tea.

From this table, it is clear that there are more models and prompts in the input to Tea in the activities of make-believe (play) and book reading than is the case with the input to Harry in these activities, (the higher number in every column is marked with bold face for clarity). In the categories of personal narratives and habitual and hypothetical topics, in contrast, there are more occurrences of

model utterances and prompts for speech reports in the input to Harry than to Tea.

Tea devotes much of her time to playing with her dolls and her toy horses and in the preceding chapter I provided examples indicating that by age three, Tea is a skilful narrator within make-believe play. The tables above show that the adults produce many utterances of framed and free direct speech in these types of interactions with Tea. A closer look at the recordings indicates that Tea and her caretakers are engaged in make-believe play already in the first recording at 19 months, and passages of play are forthcoming in almost every recording after that. From a developmental point of view, it can be observed that to begin with, only the adults produce speech projections (indirect voicing) during the play sessions (however, a handful of cases represented by Tea occur, cf. Figure 4-2 (a) below). At 31 months Tea is continuously prompted to project speech onto the dolls and she responds accordingly. This conforms well with the claim made by Wolf *et al* (1984) that children begin to ascribe intentions to figures in play with small replicas around the age of 31 months, and points to a sensitivity in Tea's mother to adjust to Tea's level of development. From 35 months on, Tea frequently makes use of speech projections and indirect voicing herself (without being prompted).

Projecting speech within make-believe play with dolls is rare in the interactions with Harry, and after 30 months there is no instance of modeling or speech prompting found at all. However, Harry often makes up fantasy stories and tells his caretakers about his make-believe friend. The four instances of prompts found in the make-believe (play) category above are produced within that type of narrative.

Both Harry and Tea enjoy "reading", listening to and discussing stories from books. Table 4:5 shows that Tea receives only slightly more model utterances than Harry in her input (17 vs 14), however, she is prompted to produce direct and indirect forms of speech to a greater extent than Harry (20 vs 5). Generally, when Harry and Tea are prompted to report or project speech, they respond to the prompts, and in the case of Tea, she and her interlocutors

171

often collaborate and co-construct their telling of stories when reading books. Several of the books appearing in the interactions show up in more than one recording and the stories are well-known to Tea. Typically, the mother or the grandmother direct the narrating, but every now and then Tea is asked to contribute. In Example 4.14, a story about "Dirty Harry", (Harry is a dog), and in

Example 4.15, a story about a rabbit who is afraid of a lot of things, Tea is specifically asked to deliver script-like utterances that contain speech of the story characters.

Example 4.14

TEA40_19

*GMM:	å så kommer han hem
%eng:	and then he comes home
*GMM:	**å va säjer barnen?**
%eng:	and what do the children say?
*TEA:	**tha smutsi han ä**
%eng:	he's so dirty
*GMM:	**ja så säjer dom nog**
%eng:	yeah that's probably what they say
*GMM:	**eller också säjer dom att # de e inte vår hund**
%eng:	or maybe they say that # that's not our dog

Example 4.15

TEA40_19

*GMM:	**å va säjer han te björnen då?**
%eng:	and what does he say to the bear then?
*TEA:	**<ja lill smaka> [=! förställd röst]**
%eng:	<I want to taste> [=! disguised voice]
*GMM:	**på ditt kex**
%eng:	your biscuit

Thus, Tea is prompted for speech, and in the first example Tea's contribution is confirmed by GMM, (*ja så säjer dom nog*, 'yeah that's probably what they say'), and then grandmother makes another suggestion of what the children in

the story might have said (*eller också säjer dom att de e inte vår hund,* 'or maybe they say that that's not our dog'). These three-part sequences of dialogue, i.e., adult prompt - child response - adult confirmation, are very typical to the interactions during the book-reading activity in the recordings with Tea. In Example 4.15 the grandmother completes Tea's utterance according to what is probably explicitly spelled out in the text of the book. Notably, Tea changes her tone of voice in order to depict the voice of the bear.

Harry does not respond to any of the five prompts given by the caretakers in the book-reading situations. However, the type of interactions between Tea and her mother and grandmother as explained above frequently occur when Harry narrates about personal experiences and when habitual and hypothetical topics are discussed.

Example 4.16

HAR46_23

*HAR:	å så såg vi en ledopajst som skulle äta upp en liten ä get
%eng:	and then we saw a leopard that was about to eat a little goat
*MOT:	mm
%eng:	uhum
*GMM:	va synn de va om den geten
%eng:	was a poor goat
*HAR:	mm
%eng:	uhum
*MOT:	**va sa du då Harry?**
%eng:	what did you say then Harry?
*HAR:	**sluta ledopauden**
%eng:	stop it leopard
*GMM:	**så sa du**
%eng:	that's what you said
*GMM:	å du va gjorde leoparden?
%eng:	and what did the leopard do?
*HAR:	inte så som ja sa till honom
%eng:	not as I told him
*GMM:	0 [=! kort skratt] de va illa att han inte gjorde de du
%eng:	0 [=! laughter] that's too bad he didn't do that

173

Example 4.16 below shows how Harry's mother prompts Harry to tell what the leopard said the day Harry visited the zoo, and grandmother then confirms or concludes his response by saying *så sa du*, 'that's what you said'. (For a similar example of prompting in discussions of a hypothetical nature see, for instance, Example 4.13 on page 163.)

In the section below, the relationship between the adults' language behavior and Harry's and Tea's own use of direct and indirect speech is further examined.

4.1.3 Modeling, prompting, and children's use

In order to shed light on developmental aspects and to illustrate the relationship between input and production, the following is plotted in Figure 4-2 below: the first occurrences in the data of an adult model utterance; the first occurrences of a speech report/projection prompt on behalf of an adult; and of the child's first production of a form (that is not prompted by an adult). To facilitate comparison between the children and their input, the pattern of the Harry data, and the occurrences in the Tea data are shown separately. Moreover, the division of types of conversational contexts has been maintained.

From these figures it is evident that the first forms produced by the children show up soon after their second birthday. The first forms produced by Harry in the personal narrative category and in make-believe (play) have the free direct form as has Tea's first form. This can be expected since their mlu-values at this point in time are only 1.4, and their capability of combining two clauses is restricted. However, the first forms to emerge in the recordings of the children's speech in relation to book-reading (at 30 months in Harry's case, and at 33 months in the recordings of Tea) are framed direct speech. Still, as was discussed in the preceding section, these early forms are unstable in nature, and it is sometimes difficult to draw a line between the types.

●	adult model utterance
■	adult speech prompt
✕	child use of form

a)   *Harry*

Play   ●   ✕   ■

Book   ●   ✕
       ■

Pers   ●   ■ ✕

Hab.   ■   ●   ✕

Age ───▶

20 mon. 22 mon. 24 mon. 26 mon. 28 mon. 30 mon. 32 mon. 34 mon. 36 mon.

b)   *Tea*

Play   ●   ■   ✕

Book   ●   ■   ✕

Pers   ●   ■

Hab.   ● ■

Age ───▶

20 mon. 22 mon. 24 mon. 26 mon. 28 mon. 30 mon. 32 mon. 34 mon. 36 mon.
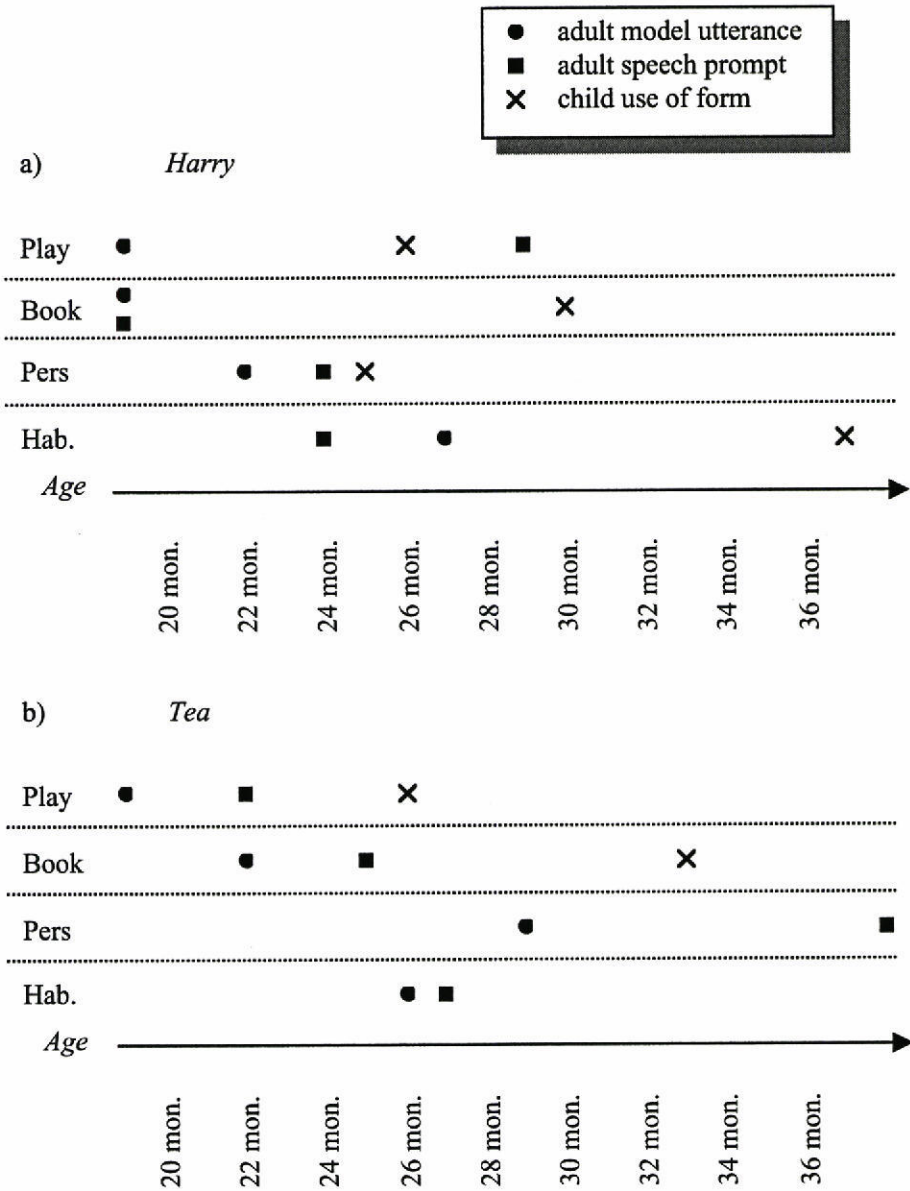
Figure 4-2: Schematic figures of the first occurrences of an adult model utterance (marked by ●); of a speech report/projection prompt on behalf of an adult (marked by ■); and of the child's first production of a form (that is not prompted by an adult) (marked by ✕); in the data of Harry (a) and Tea (b) respectively. Four types of conversational contexts.

When evaluating the input, attention should be drawn to the fact that there are model utterances present already in the first recordings at 19 months in the input to both children and by the time of the recordings at 27 months, model utterances and speech prompts can be found in connection with nearly all types of contexts.[6] Notably, model utterances and speech prompts typically precede the child produced forms. This is particularly obvious in Tea's case, where, for all categories, model utterances are the first to appear, to be followed by a prompt, and then a form produced by Tea. Hence, a plausible and somewhat simplified picture of the development is that in a first phase, the child is provided with linguistic examples; in the next phase, and as an adjustment to the child's linguistic level of ability, the adult prompts the child to produce forms of her own; and in a final phase, the child produces forms without being prompted to do it (cf. the above description of Tea's development in make-believe play).

From the figures above, it is also clear that the first speech reports and speech projections in the child directed speech are made within play and book-reading contexts, whereas forms within discussions about personal narratives and habitual and hypothetical topics emerge only later. This type of order of emergence can be expected since at an early stage of development in the child, it is more meaningful to be engaged in activities and discuss topics related to the here-and-now (like playing with toys, looking at pictures in a book, and listening to stories), than to talk about non-present topics, earlier events, or even hypothetical events. The differences between the children regarding personal narratives should be noted however. In the input to Harry, there are both model utterances and speech report prompts present before age two, and Harry's own first instance of speech reporting is found at 24 months (see Example 4.2 and Example 4.4). Yet, in the material of Tea, the first model utterance is from the recording at 29 months, the first prompt at 39 months, and Tea herself makes no speech report at all in this type of context (although she

---

[6] Except for the category of personal narratives in the input to Tea, and there is no prompt present for fantasy (play) in Harry's input until the recording at 29 months.

replies to prompts, the first response showing up in the recording at 39 months).[7]

All these results should be considered keeping in mind that the total number of model utterances, speech prompts, and child produced forms is low. Moreover, the data are only samples of the children's (and the adults') language use, collected under very special conditions, and the only possible claims concern what is actually found in the recordings. Still, I think there is a relationship between the children's language production and development, the language the adults use, and the way the adults and the children interact. The section where the results of the children's own language use (Section 4.1.1), indicated that at age three, Harry proficiently weaves in direct and indirect speech in his narratives of personal experiences whereas Tea elegantly handles indirect voicing within make-believe play at the same age. As I see it, part of the explanation of their skills within these respective domains can be found in the results of the analyses of the input. Consequently, Harry receives more forms of direct and indirect speech in the input in relation to personal narratives and discussion of habitual and hypothetical topics (28 models in total in Harry's case, and 8 in Tea's case), and for personal narratives at an earlier point in time (22 months (input to Harry) versus 29 months (input to Tea)). Furthermore, Harry is prompted to report speech in relation to personal narratives and in conversations about habitual and hypothetical topics to a greater extent than Tea (19 vs 8). In addition to this, Harry is prompted to report speech within personal narratives from an earlier point in time than Tea is (24 months vs 39 months). Tea and her caretakers, in contrast, are from an early point in time devoted to make-believe play and to project speech onto toy figures. Although the first models and self-produced forms show up at about the same time in the data of both children, speech models and prompts are clearly more frequent in the input to Tea (35 model utterances and 8 prompts in the Tea data, versus 2 models and 4 prompts in the input to Harry).

---

[7] See, however, Example 4.6 on page 157 for a case of imitation.

In her introduction to the description of the Longitudinal Swedish corpus, Richthoff (2000: 7) writes that the ambition of the data collection has been to catch the children's language within conditions as natural as possible. Therefore, the children and their caretakers were recorded when they were engaged in activities which commonly occurred in their home environment. From these data, there is reason to believe that discussing personal experiences is an activity that frequently occurs in the everyday context of Harry. He seems to be encouraged exactly to do so and also gets practice. In the same manner, Tea is encouraged to play and develop her skills within this area, including projecting speech onto dolls and play characters. These discourse habits are confirmed by Richthoff (*personal communication*) who reports that Harry at these ages gladly narrates about exciting events that he has experienced and he tells stories about his make-believe friend. In addition, he enjoys playing with his toy cars and although he occasionally includes figures or dolls in his play, he never makes them speak in the way that we so frequently find in the interactions between his sister Tea and her caretakers. Richthoff reports that Tea too enjoys narrating about personal experiences, for example, Tea conducts long reviews of films she has seen previously. Hence, similarly to her brother, Tea is a frequent narrator of experienced events. Yet, she neither includes direct and indirect speech in these accounts nor is she prompted to do so.[8]

In Section 2.3.2 on narrative development, a study by Uccelli *et al* (*in press*) was reported that showed that there was a positive relationship between the amount of discussion between the adults and the children about nonpresent objects, events and attributes before age 32 months and overall performance of personal and fantasy narration at age 5. Elsewhere (Section 2.4.4.2) I reported on a study by Ely *et al* (1996b) where a positive correlation was found between parental attention to past speech (speech reporting) in conversations with their

---

[8] It should be repeated that the methodology used for the coding of the data was to distinguish sequences of dialogue where forms of direct or indirect speech were found and then to conduct a detailed analysis of only these extracts (see 3.3.1 for a description of the procedure of analysis). Thus, narrative passages where speech reports or projections are not included, were not considered in the analysis.

children at an early age, and unprompted use of speech reports at a later age. Thus, from the results of the analyses of the Harry data above, there is reason to believe that the stimulation of Harry's narrative attempts will have an effect on narrative performance and speech reporting at an age beyond what is included in this study (47 months). Moreover, Uccelli *et al*, also found a positive correlation between early participation in fantasy talk and the representation of an evaluation marker as character's voice in fantasy narratives at age 5 (see Section 2.3.2). Tea is specifically privileged is this respect. (For a discussion on these results from a gender perspective, see Nordqvist (2000b)).

Both Harry and Tea are encouraged to report and project speech and to construct narratives, and their caretakers actively engage the children in conversation. The examples from Tea and her grandmother reading a story together are clear examples of conversational co-construction. This type of interaction is typical of Western classroom conversation (Nauclér & Boyd 1997), and in this way the children become socialized into the practice of school. The frequent engagement in book-reading activities also contribute to an increasing awareness of narrative structure, and what constitutes a typical narrative of a particular culture (cf. the examples given in the preceding section where Harry and Tea were listening to a story and kept asking *å va sa X då?*, 'and what did X say then?').

Before concluding this section, there are certain aspects worth pointing out that have implications for the analyses in following sections. Of particular interest for the analyses of the three-year-olds narrating the frog story and the ones playing with a doll house (Section 4.2 below) is the fact that there seem to be connections between types of forms and type of activity. From the analyses of forms in the input in relation to book reading above, we know that framed direct speech is the most common form used in this context (cf. also Baker & Freebody (1989), and Perera (1996) reviewed in 2.4.4.2), and the first forms produced by Harry and Tea are indeed framed direct speech. In the same manner, free direct speech is the most common form to be employed in the make-believe contexts, both by the caretakers and in Harry's and Tea's own

use. Common to both types of activities is that they are typical for interaction from an early point in time.

Yet another important point to keep in mind for the analyses of written narratives in a later section (Section 4.4), is that book reading (in the way as is often performed in the data of Harry and Tea, i.e. the adults read a story from a book) is about reading a written, often literary text aloud. Thus, children who do not yet know how to read and write themselves, indirectly come into contact with characteristics of written discourse when and if they have stories read to them. The commonality of framed direct speech in these contexts is surely (partly) a result of the type of modality, since it is not possible to make use of changes in tone of voice in order to convey shifts of speaker perspective in written text and speaker identity and verb of saying then need to be explicitly spelled out.

Finally, an important conclusion from the analyses of Harry and Tea and their use of direct and indirect speech, is that, despite the fact that the general course of development is similar, there are also individual differences, and the caretakers use their language in partly different ways with the two children. This knowledge should be borne in mind in the analyses to come.

4.2   Three-year-olds

This section primarily presents the results of the analyses of the 3-year-olds playing with the doll house with their mothers, but these results are also compared to the 3-year-olds narrating the frog story with their caretakers. The first section below (Section 4.2.1) deals with the frequency of speech projections in the two types of activities, and points out differences and similarities between the children and the adults. The second section (Section 4.2.2) explores what forms are used; the third section (Section 4.2.3) is devoted to how the forms are used (especially in relation to the Doll house corpus); and finally, in section (Section 4.2.4) the results are summarized.

4.2.1 Proportion of speech projections

In discussing the Longitudinal data for Harry and Tea in the preceding section, I distinguished four types of conversational contexts. The activities of relating the frog story and playing with the doll house, which are analyzed in this section, have clear affinities to two of these categories: *Make-believe (play)* and *Book reading* (cf. 4.1.2). Playing with a doll house typically falls within the first type of category, whereas narrating of the frog story has characteristics typical of both types of categories. This is due to the fact that *Frog, where are you?* is indeed a book with pages to turn, and events that are chronologically ordered are depicted. However, it contains no text to read aloud and the "text" then in a sense needs to be made up. In this way, it is more similar to *Make-believe (play)* than to *Book reading*. We know from the analyses of Harry and Tea above (Sections 4.1.1 and 4.1.3) that speech projections produced within make-believe play and book-reading typically occur from an early age. Before age three, both Harry and Tea produce forms of speech projections within these types of activities.

The total number of speech projections coded for in the Doll house data amounts to 300, and 166 of these are made by the children. In the frog story data the total amount of speech projections is 28, and 26 of these are produced by the children. However, these figures are only informative if they are related to the sizes of the corpora and since the types of data are diverse and transcribed and coded according to different standards, (see Section 3.2 on subjects and data), the only measurement that is consistent across the two corpora is number of words. Figure 4-3 below shows the total sizes by means of included words of the Doll house corpus (children versus adults), and the part of the Frog story corpus with the three-year-olds and their parents. In Table 4:6 further information is given regarding the range and mean number of words in the corpora.
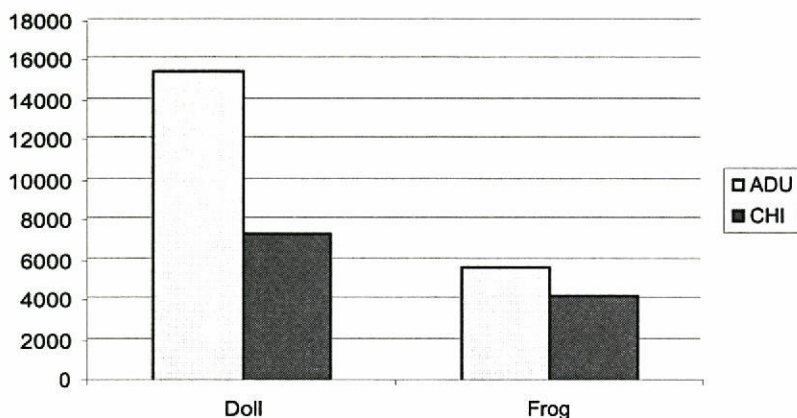
181

Figure 4-3: Total words produced by children (CHI) (N=14) and their mothers (ADU) (N=14) in the Doll house corpus, and by children (N=14) and their parents (ADU) (N=14) in the Frog story corpus.

Data	No of subj.	Total	Range	Mean	Std. Dev.
Doll (CHI)	N=14	7,357	310—720	526	136
Doll (ADU)	N=14	15,478	480—1,700	1,106	350
Frog (CHI)	N=14	4,231	147—503	302	120
Frog (ADU)	N=14	5,686	180—1,241	406	295

Table 4:6: Range and mean number of words by corpus; children vs adults.

Thus, from these data it is evident that the mothers in the Doll house corpus produce about twice the number of words in comparison to their children (15,478 versus 7,357), and this difference turns out to be statistically significant $(t(13) = -4.995; p < 0.001)$. The adults in the Frog story corpus do not produce significantly more words than their children $(p = 0.12)$. For both corpora it is true that the standard deviation value of the adults is greater than for the

children. This means that the individual differences among the adults are greater than among the children.

By calculating the total number of words included in speech projections by each subject (children and adults) and dividing these figures by the total number of words produced by each subject, information on proportional use of speech projections is provided.
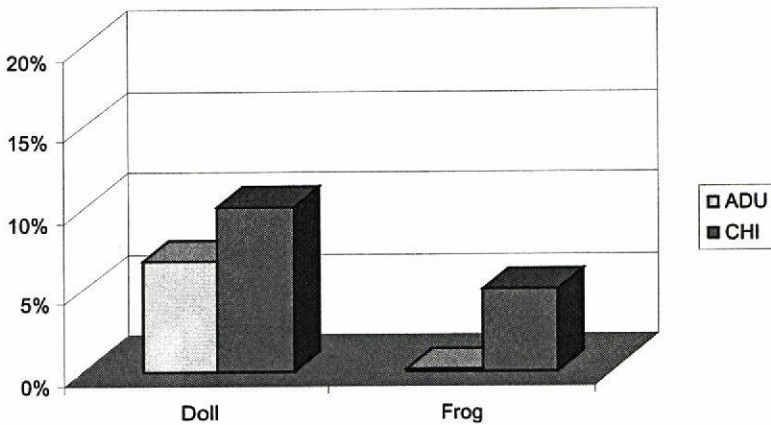

Figure 4-4: Proportions of speech projection words; mean values of percentages of words included in speech projections (related to total number of produced words); children and adults in the Doll house corpus and the Frog story corpus.

What is evident from the mean values is that the values are higher for the children than for the adults in both types of activities. Around 10% of all words produced by children in the doll house context are produced as part of speech projections, while they accounts for nearly 7% of the words by the mothers (for further descriptive statistics on the data, see Table 4:7 below). This difference is significant on a 0.05 level (paired t-test). The mean values for the frog story activity are approximately 5% for the children and close to 0% for the parents, and this difference is also statistically significant ($p = 0.01$). Thus, this suggests

that the option of projecting speech to toy figures and to story characters is especially appealing to young children.

Data	N of subj.	Mean	Std. Dev.	Min.	Max.	Median
Doll (CHI)	14	10.2	10.7	0.2	34.8	8.2
Doll (ADU)	14	6.8	8.1	0	29.8	4.8
Frog (CHI)	14	5.1	5.0	0	17.6	3.7
Frog (ADU)	14	0.2	0.5	0	1.9	0

Table 4:7: Descriptive statistics of the data; proportion of words involved in speech projections in relation to total number of words (percentages).

The results also indicate that speech projections are generally employed to a greater extent in the doll house activity than in the narrative activity of the frog story. However, a statistical comparison (unpaired t-test) between the two groups of children results in non-significant values, i.e., the children playing with the doll house do not proportionally make use of more speech projections than the children telling the frog story. However, the same type of analysis reveals that the adults in the doll house activity use significantly more speech projections than the adults in the frog story activity ($t(26) = 3.044$; $p = 0.005$), suggesting that playing with a doll house with children is more conducive to adult speech projections than being engaged in the frog story activity.

A more detailed examination of the data shows that all fourteen 3-year-olds playing with the doll house produce speech projections, while the corresponding number of mothers is eleven. Thus, most of the subjects (children as well as adults) make use of speech projections in the doll play. As is shown in Table 4:7 above, it is also the case that the individual differences within this group are great. For one of the mothers and one of the children as much as around 1/3 of all words produced in the play consist of speech

projections (cf. Max.-value), whereas three of the mothers never use speech projections.

In the narrating of the frog story, twelve out of fourteen children project speech whereas only two adults produce one speech projection/model each. Hence, the adult interlocutors in the doll house play provide their children with speech projection models whereas the adults narrating the frog story jointly with their children rarely do. On the other hand, five out of the fourteen adult interlocutors in the frog story activity use speech prompts.

These different patterns of use by the adults in the two activities, (i.e. more speech projection models in the doll house play activity than in the narrating of the frog story, and more speech projection prompts in the latter than in the former), can partly be explained by methodological considerations. The mothers of the 3-year-olds playing with the doll house were invited to *play together* with their children, whereas the parents in the case of the frog story were instructed to *minimize their interaction* with the child to what was needed to elicit a story from the child (see Section 3.2.3 for details). In the latter case, the adults, rather than telling the story *to* the child (which would most likely have resulted in more speech projection models), encouraged their children to narrate as much as they could themselves (e.g. by speech prompting). Still, although the adults were instructed to have a relatively passive role, the figures in Table 4:6 show that the adults produce as many words as their children do. Nauclér & Boyd (1997) have shown that it is typical of Swedish adults (i.e. parents and teachers) narrating the frog story together with pre-school children, to ask questions in order to make the children participate actively in the construction of the story. Nauclér & Boyd claim that the questions have two functions: the adults encourage their children to display their knowledge of the world (knowledge that is already known to the questioner), and the adult checks that the child possesses and/or is focusing on specific information at a certain point in time. Indeed, by using speech projection prompts like the following, *va säjer pojken där, va säjer han?*, 'what does the boy say there, what does he say?' (input to FROG.CET03), the adult

185

checks that the child is focusing on a certain piece of information, and in this particular case, information which is relevant to the plotline.

4.2.2 Types of forms

Playing with a doll house is an example of what was referred to in Section 2.3.4 as an existentially conservative game, i.e., involving things which are pretended to be other than they are (Evans 1982). Consequently, the dolls are assigned roles and are made to act within a make-believe world. Commonly, dialogues between, e.g., the mother doll (enacted by the real mother) and a baby doll (enacted by the three-year-old) take place, and these dialogues contain exchanges of utterances of free direct speech. This partly explains the frequency of adult model utterances as indicated above, and it also has implications for the distribution of types of forms used by the children and their mothers. The figure below shows the types of forms used by the children, in the Doll house data and in the Frog story data.

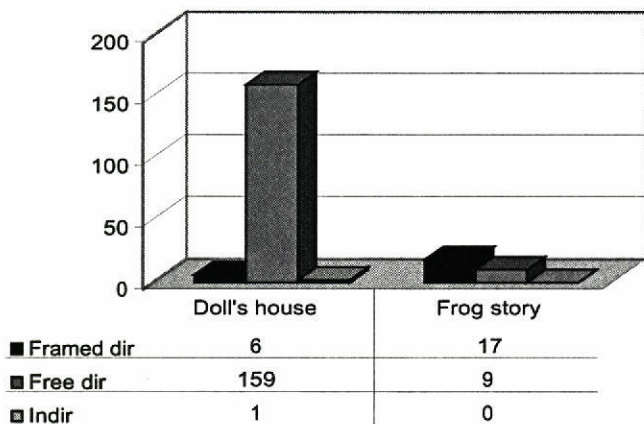|	Doll's house	Frog story
■ Framed dir	6	17
■ Free dir	159	9
▨ Indir	1	0

Figure 4-5: Distribution of types of forms of speech projections produced by the children in the Doll house data (n=166) and in the Frog story data (n=26), respectively; absolute numbers.

The overwhelming majority of forms used in the play with the doll house consists of free direct speech (n=159; 96%), and in the Frog story data framed direct speech is the most commonly occurring form (n=17; 65%). It should also be noted that indirect speech is extremely uncommon in the two corpora. As far as the frog story data is concerned, it should be mentioned that only clear cases of speech projections have been included in the calculations. As will be further illustrated and discussed in relation to Speaker perspectivizing in Section 4.3.4, the 3-year-olds telling the frog story sometimes fail to make clear if they are indeed speech projecting or not. These utterances typically have a word order characteristic of a main clause (cf. direct quotes), but due to their ambiguous status have not been included in the figure above.

A closer examination of the individual children shows that 9 out of the 11 children telling the frog story and producing speech projections, frame their direct quotes at least once. In the case of the doll house as many as 13 children produce free direct speech, but only two of these children, one girl (DOLL.010-03) and one boy (DOLL.040-03), also use framed direct speech (three times each). One girl (DOLL.042-03) uses an indirect construction once (*fråga om han kan koka mat*, 'ask if he can cook') in this corpus.

The distribution of forms of the mothers in the Doll house corpus looks similar to the production of their children. This means that free direct speech (n=116) is far more common than framed direct speech (n=15) and indirect speech (n=3). One of the mothers (DOLL.037-03) produces all three types of forms; one mother uses free direct constructions and one indirect form (DOLL.040-03); seven of the mothers make use of framed direct and free direct constructions; and two of the mothers use only free direct speech. It is worth noting that all three children mentioned above using framed direct or indirect constructions, also have mothers using direct or indirect speech models. The total number of speech projections in the adult input in the Frog story data amounts to only two. These two instances are of the framed direct type and produced by two different adults. Both children of these adults, use framed direct speech (CIT03SP and CIS03SP). From the analysis of forms in

187

the data of Harry and Tea in Section 4.1 above, we know that free direct speech is the most common form to be employed in the make-believe contexts, whereas framed direct speech is relatively more frequent in book-reading contexts. This pattern seems also to be true of the Doll house data and the Frog story data.

In the previous section (Section 4.2.1), I showed that there is no difference in the number of speech projections used by the children in the doll house play and in the frog story activity (as measured by number of words in speech projections divided by the total number of words). This indicates that the relative amount of use of speech projections is not a factor distinguishing the two types of activities. The analyses conducted here regarding the types of forms, suggest that, rather than being a question of the number of speech projections used, it is the way in which they are used, that is crucial. It is thus relevant to inquire how the forms are employed, and for what purpose?

4.2.3   Use of the forms

In a previous section (Section 4.1.1), we noted that Harry as well as Tea start to construct projected (and reported) dialogue exchanges without adult prompting around age three. In Example 4.7 on page 159, for instance, Tea manages successfully to display the shifts of perspective of projected speakers (in this case from the narrator's point of view, to a Santa doll, and then to a girl doll) at 31 months of age. She uses free direct speech and holds onto the dolls, and clearly distinguishes who is speaking by directing her gaze to, and by moving, the current doll. Indeed, this points to a fundamental difference between the two activities of playing with the doll house and narrating the frog story. There are concrete, physical and three-dimensional objects that can be held onto and made to move in the former activity, whereas in the latter, the figures are only depicted visually and one-dimensionally. In projecting speech onto the dolls, the fact that speech is projected and onto whom it is projected can be made clear by, for instance, the projector wiggling the doll. In projecting speech to

the characters in the frog story, in contrast, the projection often needs to be more explicitly described by linguistic means, in order to inform the listener about shifts in perspective.[1]

The children narrating the frog story do not use much free direct speech. However, one passage of dialogue exchange by means of free direct speech utterances is found in the data:

Example 4.17

FROG.CET03SP.cha

Situation: CET and his mother are looking at pictures 18-19 in which the boy and the dog are sitting in the pond, the dog can be seen climbing the boy's head, and the boy puts his hands to his ears.

*ADU:	de va ju inga fiskar
%eng:	there were no fish
*ADU:	som va dumma i vattnet heller
%eng:	who were mean in the water
*CET:	**nä dom skulle bara säja +"/**
%eng:	no they just wanted to say
*CET:	**+" hej [!]**
%eng:	hello
*ADU:	mm
%eng:	uhum
*CET:	**+" va bår ni [!] då ?**
%eng:	where do you live ?
*CET:	**+" vill ni leka [!] moss ?**
%eng:	do you want to play with us ?
*CET:	**+" ja de vill vi**
%eng:	yeah we do
*ADU:	ja säkert sa dom så +"
%eng:	yes I'm sure they said so

These types of dialogue occur more frequently in the Doll house data. Below is an example from a three-year-old girl:

---

[1] One way of showing shifts in perspective is to point to the actual protagonist, something that is likely to occur at this age. Unfortunately, there are only audio-recordings of the 3-year-olds available which prevented me from doing this type of analysis.

189

Example 4.18

DOLL.010-03

Situation: The child is arranging the dolls in the bedroom of the doll house. The child is holding onto the child doll and the father doll and makes them speak to each other.

CHI:	**mammas säng**
eng:	mother's bed
CHI:	**i MAMMAS SÄNG VILL ja gå å lägga mej**
eng:	in Mummy's bed I want to lie down
CHI:	**JA'HA**
eng:	uhum
CHI:	**sa pappan då**
eng:	the daddy said then
CHI:	**vänta tills VI ha // ha / gå LÄGGA oss**
eng:	wait until we're going to bed
CHI:	**JA'HA**
eng:	uhum
CHI:	**sova en liten STUND**
eng:	sleep for a little while

Although it is fairly clear in these two examples that shifts in perspective of projected speakers occur, it is not perfectly obvious in all utterances who the intended projected speakers are. To whom do *they*, *you*, and *we*, respectively, in Example 4.17 refer? Onto whom is *sleep for a little while* projected in Example 4.18? In this latter case, an analysis of the video recording and eye gaze, gestures, etc., gives no clues, and the child makes no particular changes in tone of voice in order to signal speaker identity. As will be further discussed in 4.3.4 below the 3-year-old narrators do have certain problems to succeed in making shifts in perspective clear.

In other cases information about the projected speakers and the projected speech is encoded and packaged multifunctionally and simultaneously. In Example 4.19 below, the child repeats the same grammatical construction and proposition: *ja vill också ha en korv*, 'I would like a hot dog too'.

Example 4.19

DOLL.034-03

Situation: The play is situated in the living room of the doll house where there is a fireplace. The mother doll, the boy doll, the father doll, and the baby doll are roasting hot dogs.

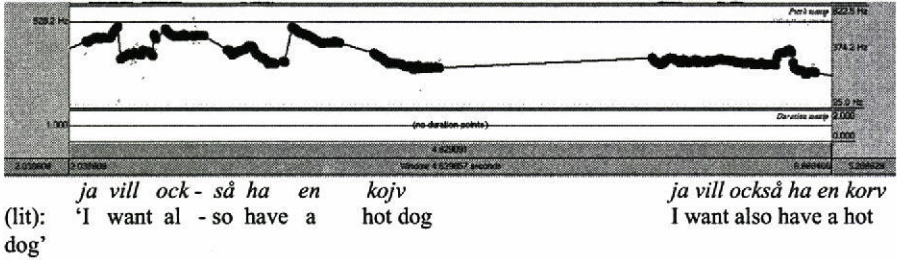|	
MOT:	får han TA en korv här ti pappan
	så å så gå han ti pappan så här
eng:	he may TAKE a hot dog here to his father
	and then he goes to his father like this
act:	makes the dolls move
MOT:	**varsågod säjer han**
eng:	here you are he says
CHI:	**ja vill OCKSÅ ha en kojv**
eng:	I would ALSO like a hot dog
act:	holds onto the baby doll; very squeaky childish voice
CHI:	**ja vill OCKSÅ ha en korv**
eng:	I would ALSO like a hot dog
act:	holds onto the father doll; deep voice

In the first child utterance, the child disguises his voice to depict a baby's voice (very squeaky), whereas in the contingent utterance, the child uses an deep exaggerated voice in order to depict a father's voice. Figure 4-6 (a) below shows the F_0 (fundamental frequency) correlates of these two utterances. A sample of the child's "normal" speech is given in (b) for the purpose of comparison. [2] The total temporal extension of example (a) is 4.5 seconds and of (b) 1.8 seconds.

In utterance (b), picked from a passage in the recording where this particular boy is talking with his mother in his assumedly ordinary voice pitch level, the mean F_0 value is 380 Hz. In the first utterance in (a), (the utterance that is projected onto the baby doll), the mean pitch value is 370 Hz. This value is fairly close to his ordinary pitch level. Still, the "doll voice" is squeakier in quality. Moreover, the baby impersonation is slower (0.31 seconds per syllable) than the sample from the child's own "normal speech" (0.22 seconds per syllable). The second utterance (which is projected onto the father doll), has a

mean F_0 value of 203 Hz, and this impersonation of the father's speech is faster than the two other utterances (0.16 seconds per syllable).
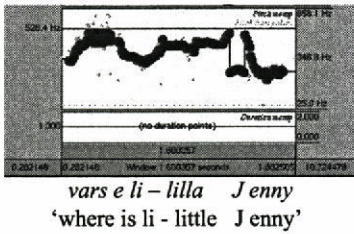
(a)

	ja vill ock - så ha	*en*	*kojv*		*ja vill också ha en korv*	
(lit):	'I want al - so have a		hot dog		I want also have a hot	
dog'						

(b)

vars e li – lilla   J enny
'where is li - little J enny'

Figure 4-6 (a) and (b): F_0 (fundamental frequency) correlates of three utterances by a 3-year-old in the Doll house corpus (DOLL.034-03).

Hence, the change in pitch, voice quality, and manipulation of temporal characteristics between the two first utterances, facilitates tracking the shift in perspective. In addition to these changes in pitch, the boy uses a childish pronunciation of *korv*, 'hot dog', namely *kojv*, in the first utterance, while in the second utterance *korv* is pronounced according to the adult standard.

Thus, by using a certain grammatical construction (including a deictic element as the first person pronoun) when expressing a speech act, simultaneously adding a particular voice quality, pitch and temporal extension,

---

[2] The F_0-analyses are conducted within PRAAT. Copyright by Paul Boersma and David Weenink 1992-2000. See http://www.fon.uva.nl/praat/.

modifying the pronunciation of a particular word (*korv-kojv*)), and using gestures (i.e., wiggling the dolls), the child provides information on (shifts in) speaker perspective and indicating aspects like age, gender, and emotional state. In addition to this, the combined acts also provide a vivid account of what is happening on the play scene, and the plot is moved forward.

Besides modifying the pronunciation of a word, there is also the possibility of selecting a particular word to signal speech register and social roles. In the example above, the child chooses to depict voices belonging to people very familiar to him (i.e., father and little sister/brother). In the example below, the same girl as in Example 4.18 depicts a conversation between two ladies:

Example 4.20

DOLL.010-03

Situation: The child holds onto two dolls, one in each hand, places them outside the doll house and makes them speak to each other:

CHI:	**gomorron gomorron har du sovi gott inatt ?**
eng:	good morning good morning did you sleep well last night?
act:	the child glances at the doll she is holding in her right hand; uses a squeaky voice
CHI:	**ja // de har ja VERKLIGEN gjort**
eng:	yes // I did INDEED
act:	gazes at the doll she is holding onto in her left hand; uses a squeaky voice
CHI:	**jaha**
eng:	uhum
act:	glances at the doll in her right hand; uses a squeaky voice
CHI:	visst var dom TOKIGA?
eng:	they were pretty SILLY, weren't they?
act:	the child turns to and looks at her mother and makes a comment on the dolls' conversation; normal narrator's voice pitch

The child uses a disguised and affected voice for both speakers, and notably, projects onto speaker number two the word *verkligen*, 'indeed'. This word probably does not belong to the child's own ordinary speech register, but is

used in order to depict the speech of an older distinguished woman (not necessarily her mother). This ability to change register points to an early sociolinguistic and pragmatic awareness. Andersen (1990) found that 4-year-olds had generally adopted speech registers tied to family and home situations in play (see 2.4.4.1). The examples given above illustrate that even 3-year-olds control such register adjustments, but in addition, registers tied to contexts outside the immediate home context are also controlled to some extent.

The final utterance in Example 4.20 is directed to the child's mother. It is uttered with a normal pitch and seems to be a comment about the events (i.e., the ladies' conversation) in the play. Hence, the perspective is successfully shifted from a projected speaker to a narrator's, or rather, a commentator's role. Strömqvist (1984), basing his study on partly the same corpus as the Doll house corpus of this thesis,[3] showed that the three-year-olds had generally managed to mark switches between the 'O-domain' (i.e., the "real" world outside the make-believe world of the doll house), and the 'D-domain' (i.e., the domain constituted by the make-believe world of the doll house). Indeed, Example 4.20 above, not only provides an example of perspective shifts within the play (the 'D-domain'), but also a shift from the doll's world to the "real" world (from the 'D-domain' to the 'O-domain').

The short passage in Example 4.20 is a monologue in the sense that only the child speaks. However, often the play (narrative) is jointly constructed, that is, both the child and the mother make contributions to the play. In Example 4.21 below, the child, acting as a girl doll, has a telephone conversation (using the toy telephones provided in the doll house) with the grandmother doll (enacted by the child's mother). The telephone conversation is concluded after a while, and the girl doll and the mother doll starts to talk with each other:

---

[3] The entire Doll house corpus (see section 3.2.3 for a detailed description) originally consisted of 45 recordings of mother-child dyads. Strömqvist (1984) examined 24 of these. Four of the recordings analyzed in the current study are the same as in the Strömqvist study.

Example 4.21

DOLL.042-03

CHI:	**vi // då kan du komma till MEJ i morr // du kan komma till mej NU idag**
eng:	we // then you can come to ME tomorr // you can come to me NOW today
MOT:	**ska ja // FARMOR komma ti dej ida?**
eng:	can I // GRANDMOTHER come to you today?
CHI:	**ja**
eng:	yeah
MOT:	**jaha**
eng:	uhum
CHI:	**hej da**
eng:	goodbye
MOT:	**ja får se om jag har TID**
eng:	I have to see if I have the TIME
CHI:	**hej da**
eng:	goodbye
MOT:	**hej då**
eng:	goodbye
com:	closing up the telephone conversation
MOT:	**nu ropade mamma // Maria vem vare som RINGDE?**
eng:	now mother called // Maria who was it that CALLED?
act:	calling
com:	change of deixis
CHI:	**de va FARMOR**
eng:	it was GRANDMOTHER
act:	shouting
MOT:	**va ville FARMOR da**
eng:	what did GRANDMOTHER want then?
CHI:	**hon skull // VI skull fara DIT**
eng:	she was // WE should go THERE
act:	shouting
MOT:	**skull vi fara till FARMOR ?**
eng:	should we go to GRANDMOTHER?
CHI:	**jaa**
eng:	yeah

Strömqvist (1984), investigating telephone conversations made by the children within the play (i.e., monologically, without interference by the mother), found that the children were aware of many pragmatic rules connected to this particular activity. We also find evidence for this in the passage above, among

other things, in that the child is the one initiating the conclusion of the conversation by saying *hej då*, 'goodbye'. This passage also has a narrative structure and a storyline:

(1)   a telephone conversation takes place between the child doll and "grandmother";

(2)   the telephone conversation finishes;

(3)   MOT (mother) acts as a narrator and provides a narrative clause in past tense: *nu ropade mamma*, 'now mother called';

(4)   the mother switches perspective to the mother doll and she asks the child (doll) to provide her with information about the recent telephone conversation in which she was not a participant: *Maria vem vare som ringde?*, 'Maria who was it that called?', and *va ville farmor da?*, 'what did grandmother want then?';

(5)   the child (doll) replies according to these conditions (i.e., the fact that the mother doll does not know about the contents of the telephone conversation, at least not who the child doll was talking to and what this person said); and makes an appropriate report with appropriate deictic changes (pronouns and tense): *de va farmor ... [hon ville att] ... vi skull fara dit*, 'it was grandmother ... [she wanted that] ... we should go there' (note also that this latter utterance is, in fact, a speech report within a speech projection).

Thus, not only do the child and her mother jointly construct a narrative, but this example also indicates a pragmatic awareness and theorizing of mind on behalf of the child.

Söderbergh (1973, 1980) contends that the language and rules of play are a fairly complex matter. The rules mean that roles are assigned to the dolls (*This is Mummy!*), the stage is set (*This is their bedroom*), the story planned and stage-managed (*Now the baby must go to bed*), and then a dramatic play (including playing the parts of the dolls) is performed (*She is eating now. "I'm hungry!"*). Analyzing children's interactions between one and a half years to four years, Söderbergh (1973) found that the children did not master the language of the game until they were between 3-4 years of age. This is in accordance with what was reported in Section 2.1.3 on children's development of theories of mind. Thus, the time between 3 and 4 years is a phase where a considerable amount of development takes place from this perspective.

In the analyses of Harry and Tea in the previous section we learned that they both, by the age of three, started to construct narratives on their own and that the children gradually improved their narrative skills as they were approaching their fourth birthday. In constructing their narratives, however, they also often needed conversational support from their adult interlocutors. Indeed, the play narrative given above in Example 4.21 is a joint product between the mother and the child. Example 4.22 below illustrate yet another type of interaction where the mother dominates the interaction whereas the child has a more passive role.

Example 4.22

DOLL.022-03

MOT:	här
	dom ligger å SOV ännu
	vicka SJUSOVARE
eng:	here
	they're still sleeping
	such LIEABEDS
CHI:	a ska vi vakna då?
eng:	yeah should we wake up then?
MOT:	ja de TYCKER ja att vi GÖR säjer vi
eng:	yeah let's say we DO that
CHI:	**VAKNA nu då** [skriker]
eng:	WAKE UP now [shouts]
MOT:	**VAKNA nu äre dags å stiga OPP** [ropar]
eng:	WAKE UP now it's time to get UP [shouts]
MOT:	DÅ stiger PAPPAN upp
	å så ha han klätt PÅ se redan seru
	de gick snabbt
	å så stiger MAMMAN upp också
	men den HÄR han sover visst ÄNNU va?
eng:	THEN the FATHER gets up
	and he has gotten DRESSED already you see
	that was quick
	and then the MOTHER gets up too
	but THIS one is STILL asleep huh?
CHI:	JA
eng:	YEAH
MOT:	**måste vi gå å säja // VAKNA nu de e morron**
	dags å stiga upp nu [rör dockorna; modifierar rösten]

eng:	we have to go and tell him // WAKE UP now it's morning it's time to get up [makes the dolls move; modifies her voice]
MOT:	gå ti SKOLAN va? troru de?
eng:	go to SCHOOL huh?
CHI:	ja
eng:	yeah
MOT:	eller ti DAGIS
eng:	or to PRESCHOOL
CHI:	ja
eng:	yeah
MOT:	eller ti DAGMAMMAN kanske
MOT:	or to the CHILD-MINDER maybe
CHI:	ja
eng:	yeah
MOT:	vicket ska vi TA?
eng:	what should we CHOOSE?
CHI:	DAGMAMMAN
eng:	the CHILD-MINDER
MOT:	m: ti dagmamman å så stiger han UPP den här killen här också å gäspar å e så trött
eng:	right, to the child-minder and then he gets UP this guy too and he yawns and he is so tired
MOT:	så bäddar vi sängarna litegrann
eng:	and then we make the beds a little

The mother in this example invites the child to participate by stating questions but still the mother is the one controlling the scene and the one embedding the events (including the speech projections) in a narrative frame. Thus, in the data of the 3-year-olds playing with the doll house with their mothers, we find a whole range of degrees of participation by the children. On the one end of the scale there are children creating their own play dialogues and narratives with little interference by the mothers, and on the other end of the scale we find mothers directing the play.

4.2.4 Summary of the results

The results in Sections 4.2.1-3 showed that the mothers in the Doll house corpus produced about twice the number of words compared to the children, a

difference which was statistically significant. The adults and the children in the Frog story corpus produced about the same amount of words (no significant difference). For both corpora, it turned out that the adults displayed greater individual differences than the children did.

When relating the number of words in speech projections to the total number of words produced, subject by subject, the results displayed that the children's mean values (percentages) were higher than the adults' in the doll house activity (10% versus 7%) as well as in the frog story activity (5% versus close to 0%). These differences were statistically significant and indicated that the children relied on speech projections to a greater extent than the adults did. A comparison between the groups of children engaged in the two types of activities showed non-significant values, whereas the difference between the two groups of adults was significant. Thus, the relative amount of use of speech projections by the children did not differ between the doll house activity and the narration of the frog story, but there were clear tendencies that free direct speech was the most frequently employed form when playing with the doll house, whereas the quotes were more often framed in the frog story activity. This is in accordance with what was found in the data of Harry and Tea (free direct speech typically occurred in the category of *Make-believe (play)* whereas framed direct speech was found in *Book reading*, see Section 4.1 above). Indirect speech was extremely rare (one case in total), and the few children making use of framed direct speech in the doll house play did so with mothers who also used this form.

In contrast to the mothers in the doll house play, the adults in the frog story activity rarely produced speech projections, i.e. speech projection models. However, speech projection prompts were used by five of the adults relating the frog story. This should be compared to the adults in the Doll house data, where no prompts were found. These prompts typically had a function of directing the child's attention to specific events crucial to the plotline (e.g., the boy's calling for the frog), and engaging the child in conversation.

4. Results

An examination of dialogue exchanges where free direct speech was produced by children in both types of activities, revealed that changes in speaker perspective were not always made clear. However, in other cases in the doll house play this type of switch of perspective was successfully carried out, and information about the projected speakers was packaged into the form of free direct speech in a rather sophisticated manner. I found evidence of some awareness of speech registers associated with close family members, but also of people from outside the three-year-old's closest social context, and the children were found manipulating voice quality, pitch and duration, segmental features and lexical traits, in order to depict a certain speaker and speech register. In addition to these results, both children and adults were found to use speech projections to move the plot forward in make-believe play. The extent to which the mothers took the lead and controlled the play events and put the speech projections into a narrative frame, varied.

These results, together with the results of the Longitudinal case studies in 4.1, indicate that children at the age of three years are in a fairly dynamic phase of language development. With regard to narrative abilities, the 3-year-olds begin to relate sequences of events and to command the language and rules of play, yet, the narrative sequences are short and adult scaffolding is often needed. We find evidence at this age that speech projections are woven into narrative segments, dialogue exchanges and shifts of speaker perspective occur, and different types of information are packaged into the speech projections by means of phonology, lexicon and gestures. However, the boundaries between projected speech, narrator's voice and the child's own commentary are not always clear (this will be further discussed in the next section), indicating that the 3-year-old still has a way to go in the development of narrative and pragmatic skills.

4.3 Spoken frog story narratives

Some results from the 3-year-olds narrating the frog story were presented in the previous section. The narratives from this age group will be further explored in this section and compared to those of the older children and adults. The first subsection (4.3.1) concerns the construction of an overall plotline in the frog story and examines the choice of tense in the narratives. Subsection 4.3.2 deals with the number of speech projections used in the narratives, and subsection 4.3.3 explores the types of forms of direct and indirect speech employed. In 4.3.4, how the forms are used by the narrators is discussed and the section is concluded with a summary of the results (4.3.5).

4.3.1 Overall plotline and tense use

There are three components in the frog story that can be distinguished as being particularly important for the global structure of the plot, namely the boy's realizing that his frog has disappeared (the onset of the plot), the boy's search for the missing frog (the unfolding of the plot), and the boy's finding the lost frog (the resolution). Each spoken narrative was coded for explicit mention of the three elements, respectively, following the procedures of Berman & Slobin (1994; see also Ragnarsdóttir 1992).

Component	3 yrs (N=14)	4 yrs (N=14)	5 yrs (N=14)	9 yrs (N=14)	12 yrs (N=14)	15 yrs (N=14)	Adults (N=14)
I	2	6	10	8	12	13	14
II	7	8	9	12	14	14	14
III	2	3	11	11	12	12	14

Table 4:8: Number of narrators making explicit reference to each of three core components, by age.

4. Results

What is evident from the table above, is a developmental trend such that the older the narrators are, the more components they mention. In the youngest age group the numbers are very low, whereas in the adult group of narrators all three core components are explicitly mentioned by all narrators. Moreover, the component that is included by most narrators is the boy's search for his frog (II). This occurrence may be explained by the fact that the act of searching plays such a salient and crucial role in the story, and in several pictures you explicitly see the boy calling and searching for something (i.e. the frog). In Berman & Slobin's (1994) study, (where 268 narratives by English, German, Spanish, Hebrew, and Turkish 3-, 4-, 5-, and 9-year-olds were analyzed), the scores were lower for the youngest age groups than for the older groups on this component (II). A reason for this discrepancy may be the fact that the youngest narrators in this study tell the story together with a parent. Although explicit adult prompted mentionings of the search are excluded from the present analysis, it is likely that the discussion around the story leads the child to realize that the boy is searching for a frog/the lost frog.[1] Notably, only a few of the youngest narrators (the 3- and 4-year-olds) linguistically encode the resolution of the plot (component III; the boy's finding the lost frog), whereas eleven out of fourteen 5-year-olds do. This indicates an ability by some of the older children to connect the events depicted in the final scenes to the events depicted in the beginning of the story, an ability that is not present among the younger children.

One peculiarity evident in the table above, is the fact that only eight of the fourteen 9-year-olds make explicit reference to component I. Interestingly, in this age group several of the narrators point out the boy's discovery of the frog's disappearance implicitly rather than explicitly. Consequently, the following example extracted from a narrative produced by a 9-year-old girl,

---

[1] Moreover, since the recordings were made by the parents in the children's home without the presence of a research assistant, is it difficult to know for what might have happened before the tape-recorder was turned on. It should also be noted that the 3-year-olds and the 4-year-olds are the same children at different age stages. Although the children have not seen or been "reading" the frog story for a year when they tell it at age four, they may remember some parts of it.

(EME09SP), was *not* coded for component I, although the discovery of the frog's disappearance may be inferred by the listener/reader:

Example 4.23

när pojken låg å sov rymde grodan ... och när pojken vaknade blev det helt kalabalik, pojken letade och letade...

'when the boy was sleeping, the frog escaped ... and when the boy woke up, everything was in an uproar, the boy was searching and searching...'

The youngest narrators only rarely mention all core plot components. The problem with constructing an overall plotline is also evident from an analysis of the usage of tense in the 3-year-old group. For a narrative to be counted as anchored in a particular tense, 75% of all finite verbs in a narrative should be of the same tense (Berman & Slobin 1994: 62). Figure 4-7 below provides a summary overview of anchoring tense in the spoken narratives.

Typically, the narratives by the 3-year-olds are told in present tense, (eleven out of fourteen). Two narratives are of the mixed tense type, and only one single narrative is anchored in the past tense (although in the narratives with dominant present tense there are also past tense forms scattered throughout). The "narratives" have the character of conversations between the children and their parents about what can be seen and what is happening in the pictures, rather than narrations organized around a plotline. The children keep moving into and out of the picture describing mode, and shifts are typically not thematically motivated but based on local considerations. As was also the case with the youngest children in Berman & Slobin's study, the 3-year-olds treat the task as interactive and personalized and they require far more confirmatory prompts and encouragement from their interlocutors than other age groups. The 4-year-old narrators behave in a similar manner, although they do not require as many prompts from the interlocutors as the younger children. Moreover, at age four, the children anchor their narratives in present tense to a lesser extent

than at age three; instead, some of the narratives are anchored in past tense (n=4) and some in a mixed fashion (n=4).
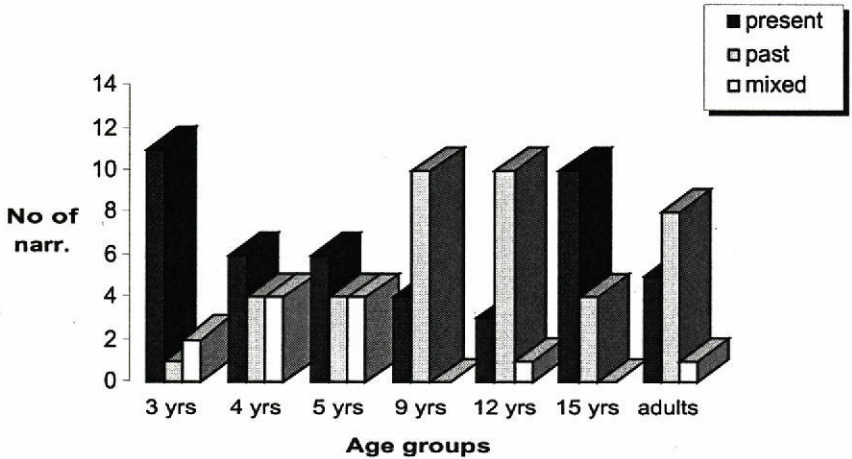

Figure 4-7: Tense anchoring in the spoken narratives, by age.[2]

The 5-year-olds manage to relate without support from the interlocutors, and when switching between tenses, they do it in chunks (i.e., the narrator may start out telling in past tense, then switch over to present tense for several pictures, and then return to the past tense for the rest of the story), rather than shifting tense sporadically (which is more typical of the 4-year-olds). Six of the fourteen narratives are anchored in present tense and several of the 5-year-olds describe what is happening rather than relating to the associated causal circumstances motivating the events. However, it is indeed a heterogeneous group, in which some narratives are globally structured whereas others are not.

The picture that emerges from this survey of the results is very much in accordance with what Berman & Slobin (1994) and Karmiloff-Smith (1981) described. As in those studies, this study illustrates that the 3-year-olds describe what they see in the pictures, they need adult scaffolding (cf. the results from the analyses of the doll house in the preceding section), they use

present tense, and they fail to construct a global structure. The 5-year-old narrators, in contrast, do better on the construction of plotline and typically need not rely on adult prompting. The 4-year-olds have certain aspects in common with the younger narrators in that they do not typically manage to construct a global structure of the frog story, and they need adult scaffolding (although less than at three years). Yet they also share characteristics with the 5-year-olds in that some narratives are anchored in past tense. Past tense is typical of Swedish fairy tales and several of the older narrators choose to use it. Interestingly, it shows that *within* the group of 4-year-olds, it is the four oldest 4-year-olds by the time of the recording that anchor their narratives in past tense (age of all four subjects, 4 years and 6 months). This suggests that by around the fifth birthday important narrative practices and pragmatic abilities are typically acquired.

In their crosslinguistic study, Berman & Slobin found that past tense is favored by most English and Hebrew-speaking 9-year-olds, while present tense is used more often as the anchoring tense by German, Spanish, and Turkish-speaking 9-year-old narrators. These results indicate that different cultural traditions of narration (including the typical tense used in fairy tales in a specific culture) and practices of school have an impact on the children's narrations at this age.[3] Moreover, only one (out of a total of nearly fifty narrators in the Berman & Slobin study) moved from one tense to another. The authors comment that these school children produce texts which give evidence of a sustained narrative mode, and they also point out that it is remarkable how similar the narratives are to one another, across and within languages. They conclude that the 9-year-olds, having been exposed to narrative texts in the context of classroom and home reading, have a fairly stereotypical and homogeneous way of narrating. These aspects are confirmed to a large extent by the Swedish 9-year-old narrators. For instance, they never mix tenses, but

---

[2] N=14 in each age group. If nothing else is mentioned this is the case henceforth (for the figures in section 4.3).
[3] However, since I have little knowledge about the narrative traditions of these cultures (except for the Anglo-saxon culture) I cannot explore this further.

typically anchor their narratives in past tense (ten out of fourteen narratives). As mentioned above, past tense is often used in fairy tales, and indeed, nine out of the ten Swedish narrators telling in the past tense, start out with classical storybook openings like *det var en gång...* ('once upon a time...'). The 9-year-olds generally tell more complex narratives than the 5-year-olds, they are not bound to picture descriptions, and they organize the story around the central theme.[4] However, although many of the Swedish 9-year-olds are fairly stereotypical narrators (and some even produce proficient but "dry and businesslike" (Berman & Slobin 1994: 74) stories), there are, interestingly, other narrating styles and narrative types present in this age group. I will return to this fact later on, but it can be said already at this point, that the other narrating styles are characterized by the use of a great number of speech projections.

The 12-year-old Swedish narrators also prefer past tense (ten out of fourteen narratives are anchored in the past tense), and they too often start their narrating by *det var en gång...*, ('once upon a time...'). In analyzing coherence and choice of verb types (transitive versus intransitive verbs) in relation to one of the scenes of the frog story (the boy's fall from the cliff), Nordqvist & Strömqvist (1995) found that the 9- and 12-year-olds used similar strategies, whereas the 15-year-old narrators showed other patterns of use. Hence, something seems to happen between 12 and 15 years of age. Indeed, from the point of view of tense use, the 15-year-olds, (in contrast to the younger school children), prefer to anchor their stories in the present tense, (ten out of fourteen). Instead of *det var en gång*, the stories begin with formulas like *det här handlar om en pojke och en hund*, 'this is about a boy and a dog', and *lille Peter sitter vid sin säng...*, 'little Peter is sitting by his bed...'. The narratives are not as elaborate as they are in the adult group of narrators and in terms of content the narratives are fairly tightly connected to what is happening in the pictures. Yet, the narratives have a clear global structure.

---

[4] Note, however, that not all 9-year-olds *explicitly* mention the boy's realization that the frog is gone (cf. discussion above), which can be seen as the actual onset of the plot.

Berman & Slobin did not include adolescent narrators in their study, but they studied adult narrators. The narratives produced by this group of narrators were heterogeneous, and this is also striking in the Swedish corpus. Indeed, the adult narrators in both the Berman and Slobin study, and mine, have a range of different narrative and rhetorical styles to choose from, and they do it successfully. The narratives range from very complex and elaborate, to shorter and more concise and there are different beginnings present. Still, although the group is heterogeneous in these respects, there are certain characteristics that the adult narratives have in common that distinguish them from the narratives of the younger subjects. The adult narratives have a complex syntax, a rich vocabulary, and are longer than the narratives produced by the younger narrators (see further Section 4.4.1). In the use of anchoring tense, more than half of the narrators, (eight subjects), anchor in past tense. Thus, the trend among the 15-year-olds to use present tense as the dominant tense, is not carried over to the adult group. In this sense, the adult narrators are more similar to the younger school children.

4.3.2 Proportion of speech projection clauses

The above overviews, now makes possible to explore how, and to what extent, the speech projections are integrated in the narratives. The whole corpus of spoken narratives (N=98) is comprised of a total of 7,801 clauses. 540 of these clauses were coded as speech projection clauses (i.e., clauses with speech act verbs, and clauses consisting of quotes and speech depictions/descriptions).

The proportion of speech projection clauses out of total clauses was calculated for each narrative. The figure below shows the mean values for each age group.
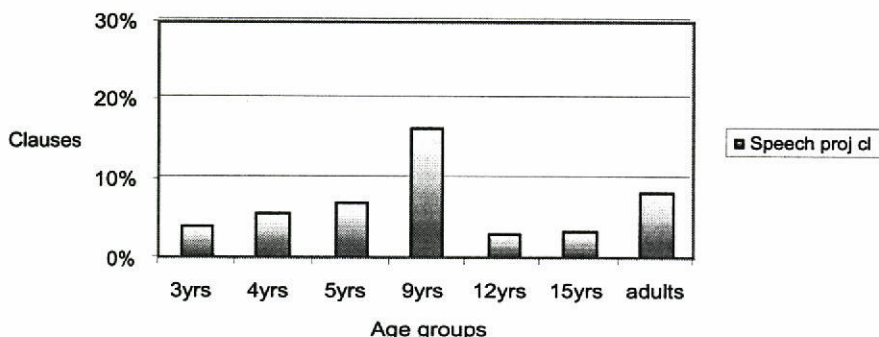
Figure 4-8: Percentage of speech projection clauses in the spoken narratives; mean values for each age group.

The mean values for the 3-, 4-, and 5-year-olds, are four, six, and seven percent, respectively, (see further Table 4:9 below). The corresponding figure for the 9-year-old group is sixteen percent and in the 12-year-old group as well as in the 15-year-old group, the mean values are three percent, whereas the adult group has a mean value of eight percent. Table 4:9 shows descriptive statistics (percentages) of the data by means of number of clauses involved in speech projections compared to the total number of clauses.

Age group	No. of narr.	Mean	Std. Dev.	Min.	Max.	Median
3	14	3.9	4.2	0	13.4	2.8
4	14	5.6	7.2	0	24.6	2.7
5	14	7.2	8.2	0	26.0	3.7
9	14	16.4	29.4	0	86.4	4.5
12	14	3.2	2.3	0	8.0	3.0
15	14	3.3	3.8	0	14.1	3.2
Adults	14	8.2	7.3	0	23.6	7.1

Table 4:9: Descriptive statistics for all age groups; number of clauses involved in speech projections compared to total number of clauses (percentages); spoken narratives.

It can be seen in the table above, that in all age groups there is at least one narrative that does not include a speech projection clause (the Min. value). Moreover, it is also evident from the table that the narrative containing the greatest proportion of speech projection clauses in each age group (the Max. value) ranges from 8.0% in the 12-year-old group, to 86.4% in the 9-year-old group. The mean value in the 9-year-old group is higher than any of the other groups (16.4%). In addition, the standard deviation value is extremely high (29.4%) in this group. This fact is illustrated in Figure 4-9 below (the error bars show the standard deviation for each age group).
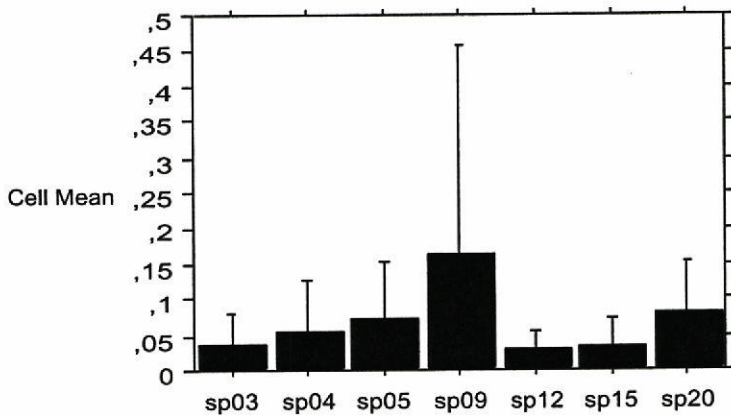

Figure 4-9: Interaction bar plot for proportion of speech projection clauses (mean values), spoken narratives, all age groups. Error bars: Standard deviations.

The explanation for the exceptional figures is that two narratives deviate greatly from the rest of the narratives in the 9-year-old group. ANT09SP and LIE09SP produce as many as 86.4% and 83.7% speech projection clauses in their narratives, respectively, whereas the remaining 9-year-olds produce a much smaller proportion (cf. the median value that is 4.5 in this age group). Yet, although the narratives of ANT09SP and LIE09SP are not representative of 9-year-old spoken narratives, it is of interest that they exist (i.e., it is an

indication that there are different ways of narrating in this age group). I will return to this question later on.

If the 9-year-old group with its high standard deviation value is excluded, ANOVA shows no significant values among the remaining age groups ($F(5, 78)$ = 0.176, p = 0.132). However, Post Hoc tests show a weak tendency that the adult narrators as a group make more use of speech projections than the teenage narrators (the 12- and 15-year-olds).[5] In addition, the median value in the adult group of narrators is higher (7.1%) than in any of the other age groups.

In conclusion, speech projection clauses are relatively few in the frog story narratives, although there is a slight tendency that adult narrators as a group make more use of them than the younger narrators. In addition, there are two 9-year-olds using speech projection clauses to a much greater extent than their peers. Let us now turn to the question of what the speech projection clauses consist of, in terms of forms of direct and indirect speech.

4.3.3 Development of forms

In total, 284 speech projections were coded in the spoken Frog story data. Twenty-five of these were coded as either UnclearIf (i.e., when difficult to decide if it was a speech projection at all) or UnclearType (i.e., when difficult to decide what type of form). Thus, 259 speech projections were coded into the categories defined above (Section 3.3.3, page 142), i.e. framed direct, free direct, and indirect speech. Figure 4-10 shows how the forms are distributed in each age group. Category *DirX* includes all framed direct speech tokens where a direct quotation is either preceded, followed, or interrupted by a clause with a speech act verb.

---

[5] 12-year-olds versus adults, p = 0.030, and 15-year-olds versus adults, p = 0.034, (Fisher's PLSD).
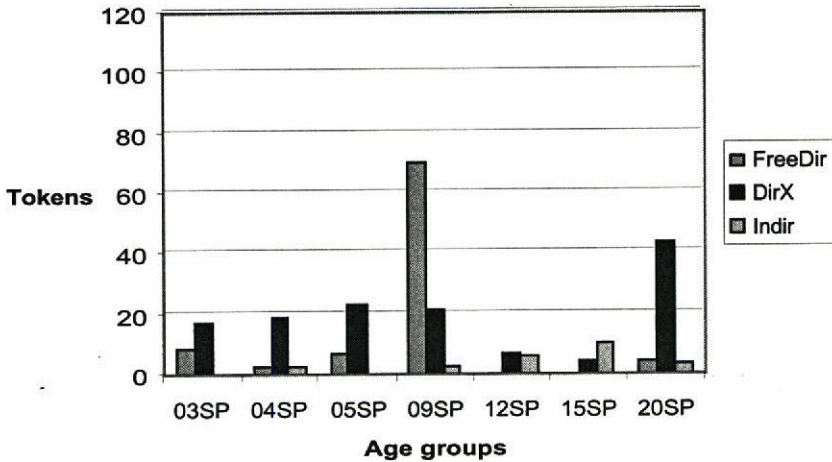
Figure 4-10: The distribution of forms of speech (raw numbers) in the spoken narratives; all age groups (number of narratives in each age group = 14); forms coded as UnclearIf or UnclearType are not included.

What can be seen in Figure 4-10, from a general point of view, is that the form of framed direct speech seems to be a favored means to project speech. Note also that, contrary to what may be expected (cf. Section 2.4.4.1), the youngest narrators often frame their direct quotes verbally. As a matter of fact, in the 3-year-old group framed direct speech scores nearly twice as high as free direct speech (17 tokens vs 9 tokens), and while only five out of the fourteen narrators produce at least one case of free direct speech, nine narrators project speech at least once with the form of framed direct speech.

An analysis of changes in tone of voice reveals that in the cases of direct speech, the quotes are sometimes marked off by prosody. Consequently, in the cases where the clause with the speech act verb precedes the quote, there is a falling intonation pattern in the final syllable(s) of this clause. In addition, at the clause boundary, as a way of marking this same boundary and to "take off", there is a short pause before expressing the quote. The quote is then uttered with a pitch distinct from the one used in uttering the clause with the speech act verb. This seems to be a relatively early achievement, in that most 3-year-olds, and all 4-year-olds, using framed direct speech display this prosodic pattern (cf.

also Harry at 4 years and 1 month in Example 4.12 on page 162). Figure 4-11 shows a F_0-analysis of a case of framed direct speech by a 3-year-old (CUL03SP).
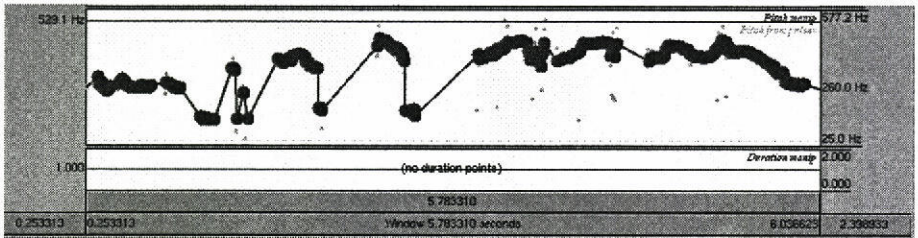
å pojken skriker ei	*hei*	*kom hit kom hit kom hiiiiiiiit*
'and the boy shouts ei	hei	come here come here come heeeere'

Figure 4-11: F_0 (fundamental frequency) correlate of an utterance by a 3-year-old in the Frog story corpus (CUL03SP).

The framing clause (*å pojken skriker*, 'and the boy shouts') has a mean pitch of 243 Hz, and the final syllable in this clause has a pitch of 129 Hz. The syllables *ei hei* are hesitation sounds taking place in the boundary between the framing clause and the quote and are followed by the mere quote, *kom hit kom hit kom hiiiiiiiit*, 'come here come here come heeeere', with a mean pitch of 395 Hz.

It should be kept in mind that Figure 4-10 above does not include cases where it has been problematic to decide if an utterance is a speech projection or not (i.e., UnclearIf), and these often have the structural (syntactic and deictic) characteristics of free direct speech. Hence, including these cases would have resulted in a higher proportion of free direct speech, especially in the youngest age groups. (I will return to the issue of cases of UnclearIf in relation to speaker perspectivizing below.) Whereas the youngest narrators (the 3-, 4-, and 5-year-olds) sometimes use free direct speech, the older school children (12-, and 15-year-olds) never do. One explanation for this avoidance of free direct speech (and preference for framed direct and indirect speech), by older children is that they seem to want to be as explicit as possible when displaying

perspective shifts by including a clause with a speech act verb and a speaker identity. As will be shown below, the youngest narrators only partly manage to make these shifts clear, for instance by changing the pitch, when using free direct speech. Indirect speech is generally more often used by older narrators than by the pre-school children, and in the groups of 12-, and 15-year-olds sixteen out of the twenty-eight speech projections are of the indirect type.

In the 9-year-old group, there is a noticeable peak of free direct speech use (73 tokens). However, as many as 68 of these tokens are produced by narrators ANT09SP and LIE09SP, who were described in Section 4.2.3 above as using a large number of speech projection clauses. Their narrations include long sequences of dialogue exchanges of direct speech and the recountings have a strong comic-strips character. The story is moved forward in these two narrations to a great extent by the subjects having the story characters say things and thereby describing what happens from inside the story, rather than having a narrator's voice provide the plot.[6] These examples illustrate the point I want to make that distinguishing narrative from evaluative functions and referring to direct quotes as being only (or primarily) evaluative, is not a plausible approach to take (cf. my critical comments in Section 2.3.1). Setting the speech projections of these two 9-year-olds aside (i.e., basing the results on the remaining twelve 9-year-olds instead of all fourteen subjects), results in a distribution resembling the one in, for instance, the 5-year-old group. Doing so illustrates that framed direct speech is preferred over free direct speech. This is true also of the adult group of narrators.

In a study of the participation of tense in narrative construction (Nordqvist 1999), I found that for all age groups except the three-year-olds, speech projections were typically used in narratives that were anchored in the past tense.[7] As stated earlier, past tense is typical of Swedish fairy tales, and past tense then becomes a marker of fiction. In Section 2.4.4.2 it was reported that

---

[6] This way of making use of speech projections is similar to the use of some of the children and mothers playing with the doll house (cf. section 4.2.3).
[7] The 12-year-olds were not included in this study.

children's first reading books contain a large amount of framed direct speech (Baker & Freebody 1989). The use of past tense in the framing clauses marks the presence of a narrator and a narrator's voice, and when direct quotes are used the deictic (perspective) shifts to story characters are made clear. Free direct speech used in narratives anchored in past tense signals deictic shifts by the use of present tense in the quote (although the information about *to whom* the perspective is shifted may not be made clear). Narratives that are anchored in the present tense cause more problems in this regard. When using free direct speech, the speaker perspective shift may appear less transparent since the same tense is used for narrator's voice and character's voice. To avoid these problems with deictic shifts in narratives told in the present tense, the use of speech projections can be minimized or indirect speech be employed (the perspective then remains with the narrator). Interestingly, from the review in Section 4.3.1 we know that ten out of the fourteen 15-year-olds anchor their narratives in present tense, and in these narratives we find examples of indirect speech (see Figure 4-10). It may thus be the case that indirect speech is a strategic choice by these narrators. The 3-year-olds also primarily narrate in the present tense, and as will be evident below, their use of free direct speech involve certain problems from the point of view of conveying shifts in perspective.

4.3.4   Use of the forms

From the examination of the relationship between scenes, or events visible in the pictures, and production of speech projections, it is clear that many narrators ascribe speech to the protagonists when it is not triggered or evident from what is visible in the pictures. This means that it is not necessarily when the boy has his hands cupped around his opened mouth (as, for instance, in picture five), that speech is projected onto the boy, but also in picture seven where the boy is carrying his dog: *"när skall du skaffa glasögon?" muttrade Sebastian [pojken] förebrående,* '"when are you going to get yourself a pair of

214

glasses?" Sebastian [the boy] asked reproachfully' (JOM20SP). Cases like these have been coded as Other in the category *Pict. matching* (see Section 3.3.3), and they are frequent in most age groups. This indicates that telling the frog story is about much more than only describing what is seen in the pictures and that narrators add information and make inferences. More than half of all speech projections by the 9-year-olds and the adult narrators, (66% and 56%, respectively), are of this type and in the three youngest age groups and in the 12-year-old group, these cases constitute around one third of all speech projections. However, in the 15-year-old group, only one speech projection refers to an "invented" event.

Narrators in all age groups, except in the two youngest (the 3-, and 4-year-olds), often project speech in relation to picture number twenty, where you can see the boy with his finger in front of his mouth as if he is hushing his dog. This event is clearly depicted in the picture and the event is crucial to the plotline (since the silencing of the dog indirectly leads to the finding of the frog behind the log). In fact, it is difficult to ignore the fact that the boy is addressing his dog by (non) verbal means. The 12- and 15-year-olds, who generally make use of very few speech projections, often depict this scene with direct or indirect speech. Nevertheless, this event is harder for the 3- and 4-year-olds to interpret. In fact, in several of the dyads with these young children, the adults try hard, and fail, to make their children describe what they see. The picture thus seems to make little sense to the children. In the scenes where you can see the dog barking at the beehive (pictures nine and ten), in contrast, the 3- and 4-year-olds often project speech (or onomatopoetic sounds) onto the dog. These pictures, or the event depicted in them, thus seems to attract little children.

In the pictures in which the boy can be interpreted as calling (pictures five, eight, nine, and fourteen), the 5-year-old narrators and the adults, to a greater extent than narrators in other age groups, project speech (or rather, calling) onto the boy. What the pictures showing the boy calling have in common is that one can clearly see that a speech act is performed. A five-year-old narrator

is likely to encode this event linguistically. In addition to the fact that this act of calling is clearly portrayed in these pictures, the calling has an obvious function of constituting a main thread throughout in the story (i.e. the fact that the story is primarily built up around the boy's searching for his frog). Indeed, as seen in Section 4.3.1 above, the 5-year-olds in this study performed better than the 4-year-olds in constructing an overall plotline. The frequent encoding of the calling (by means of framed direct speech) by the 5-year-olds further substantiates the claim that this age is an important developmental phase for narrative awareness and the ability to construct a story.

It is clear from the analysis of who the projected speakers are, that the boy is the main protagonist and the main speaker. In all age groups, he is the most frequent projected speaker. In the 3-, 4-, and 9-year-old groups the narrator projects speech onto the dog to a greater extent than in the other age groups. The 12- and 15-year-olds almost exclusively project speech onto the boy, [8] whereas (some of) the 9-year-old and adult narrators include a much wider range of projected speakers. Thus, in addition to the boy and the dog, the frog (and/or the frog family), the bees, the gopher, the owl and the deer on some occasion "speak".

So, what do the protagonists actually say? Do animal characters produce words? The adult narrators would seem to think so, since 90% of their projected speakers that are animals use words in their utterances. Below is an example from an adult narrator (CAR20SP), that is very reminiscent of the genre of fables.

Example 4.24

Men ugglan blev lite stött av att pojken trodde att han skulle göra honom illa å sa:
- Du måste lära dig att skilja dom dumma från dom snälla, alla vet väl att mullvadar biter folk i näsan när dom täcker för deras dörr med huvudet, men jag ja bara övervakar och tar ansvar för vad som händer i skogen jag glömmer inget och kommer ihåg allt.

---

[8] This is the case with the exception of two instances in the 15-year-old group, where one utterance is ascribed the frog, and another one the gopher.

'But the owl took offence when the boy thought that he would hurt him and said, "you must learn to separate the evil ones from the kind ones, everybody knows that gophers bite people on their noses when somebody's face appears covering their door, but I, I just keep watch and take responsibility for what is happening in the woods, I don't forget anything and I remember everything".'

The younger narrators also let their animal characters speak with words sometimes, although it is more common that they are only ascribed cries (in the 3-year-old group, the animal characters are exclusively ascribed onomatopoeic sounds).[9] From the analysis of the length of the mere speech quoted or described, it is evident that the quotes become longer by age. Example 4.24 above is an example of an extraordinarily long quote by an adult narrator.

From this description, general patterns of use of the speech projections connected to the age of narrator can be discerned (see Table 4:10 below). As was noted above, the 12- and 15-year-olds (almost) never let the animal characters speak. This indicates that they primarily choose the perspective of the boy in their narrations. Several of the 9-year-olds and the adult subjects, in contrast, also include various animal characters as speakers and shift perspectives between them. In Section 3.3.3., a successful act of speaker perspectivizing was defined as the narrator managing to make clear who is the intended projected speaker and/or successfully conveying shifts in perspective (e.g., from narrator to story characters, or among characters), by means of speech projections. (For further details of the operationalizations of these aspects, see Section 3.3.3.) Figure 4-12 below displays the percentages of speech projections in the spoken narratives where the narrators manage to make clear who is the intended projected speaker and/or successfully convey shifts in perspective.

---

[9] To illustrate two cases in point: firstly, an example of a five-year-old projecting a verbal proposition to the gopher: *å så kom de upp en liten sork som sa: "va ä de, ja ä ingen mus"*, 'and then a little vole appeared and said, "what is it about, I'm not a mouse"', (MAL05SP); secondly, a three-year-old projecting an animal cry onto the dog: *å sen säjer han: "voff off"*, 'and then he says, "bow-wow"', (CUL03SP).

4. Results

Age group	Speech events (picture matching)	Projected speaker	Quote content	Quote length in clauses
3, 4	barking at the beehive; waving goodbye; other pictures	boy; dog; some other animals	boy is projected words; the animals words or onomatopoetic sounds	≈ 1
5	calling for the frog	boy	verbal content	≈ 1
9	hushing scene; other pictures	only boy; or a range of animals (incl. dog)	primarily verbal content	≈ 1.5
12, 15	hushing scene	boy	verbal content	≈ 1.5
adults	calling for the frog; other pictures	boy; a range of animals	verbal content	≈ 2

Table 4:10: Typical patterns of contents of speech projections in the spoken frog story narratives; by age group.
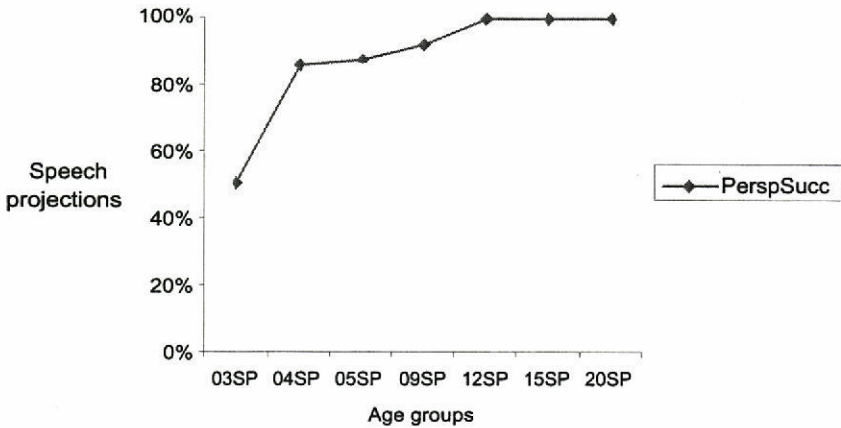
Figure 4-12: Speaker perspectivizing: Percentage of speech projections in the spoken narratives where the narrators manage to make clear who is the intended projected speaker and/or successfully convey shifts in perspective; by age groups.

The graph above (Figure 4-12) shows that about half (51%) of all speech projections produced by the 3-year-olds are successful from the point of view of speaker perspectivizing. Thus, nearly half (49%) of the speech projections are produced in such a way that it is difficult, or impossible, for the listener to decide who the intended speaker is. One plausible explanation for these results is that the 3-year-olds typically use present tense in their narration, and when using free direct speech the deictic shifts become unclear (cf. discussion in the preceding section). One year later (i.e., when these children have reached the age of four), these problematic cases are substantially fewer and the successful acts of speaker perspectivizing comprise 86% of all speech projections. In contrast to the narrations at the age of three, the speech projections are now used in past tense anchored narratives which clarifies the perspective shifts considerably. In the 5- and 9-year-old groups these proportions are 88% and 92%, respectively. In the 12-, 15-year-old, and adult groups all (100%) of the speech projections are successful in this respect.[10]

Typical of the youngest narrators (the 3-year-olds) is that they fail to make clear if an utterance is a speech projection at all (this has been coded as UnclearIf in the *Form* category). For instance, when CON03SP exclaims *jag ska leta!*, 'I will search [for the frog]!', it is hard to know if *jag*, 'I', refers to the narrator/subject himself, or if the narrator/subject is intending to project the speech act to, for instance, the boy character in the story. This supports the claim made by e.g. Hickmann (1993), that younger children tend to have problems marking explicit boundaries between narrative and narrated speech, and also, I would like to add, making explicit boundaries between reality and fiction.

The 9-year-olds, in contrast, typically demarcate the speech as projected although they in certain cases fail to make clear which of the story characters is

---

[10] I have been the listener judging the degree of success here, and since I know the story very well it is likely to be the case that I have managed to figure out the changes of speaker perspective to a greater extent than would a listener to which the story is unknown. If more coders had been used to judge the degree of success, the percentages would most likely have

the speaker. ANT09SP, for example, uses many speech projections in the form of free direct speech, and several animal characters are attributed speech (often verbal propositions) in his narration. In the swift dialogue exchanges it is sometimes difficult to know who is actually speaking. Nevertheless, as a listener to his narration you get certain clues that facilitate the tracking of the changes in speaker perspective. Firstly, he anchors the narrative in past tense in order to distinguish the narrator's voice from character speech; secondly, he changes his voice pitch in order to depict different speakers; and finally, he makes use of vocatives and animal cries in order to inform about projected speakers.

Use of speech projections as a narrative strategy have consequences for the liveliness of a narration, and choices of form have certain effects on the experienced distance between the narrator and story characters. As was explained in Section 3.3.3., each speech projection in the material has been coded for degree of distancing, where forms of free direct speech were coded as MinDist, cases of framed direct speech received the MidDist code, and indirect speech was coded as MaxDist. Since these functions are tightly connected to type of form, Figure 4-10 (page 213) above is informative. This figure shows that the two possible ways of decreasing the distance when projecting speech by using framed direct and free direct speech, are more frequently employed by the younger children and the adults than by the 12- and 15-year-olds. These older children only rarely use the option of projecting speech, and when they do, they tend to use indirect speech which contributes to increasing the distance to the protagonists and to upgrade the narrator's (or reporter's) role.

This review of the results shows that the adult narratives have certain aspects in common with the narratives of the older children, as well as with those of the younger narrators. Similarly to the 12- and 15-year-olds, the adults score high on successful speaker perspectivizing, that is, they are clear about

---

been lower. Still, I think the appearance of the curve and the *tendency* shown in Figure 4-12 would have been similar.

whose vantage point is taken. In other respects, the adult narratives are more similar to the narratives of the pre-school children and to some 9-year-olds. Framed direct and free direct speech, rather than indirect speech, are favored forms in connection with speech projections, a variety of animal characters has an active role and narrators project speech onto them, and speech projections occur frequently. In fact, the adult group of narrators has the highest mean and median value (7%) of all groups when it comes to proportions of speech projection clauses, and they use significantly more than the 12- and 15-year-olds. The pattern of use among the adult narrators increases the liveliness of the narrations and this way of narrating would probably attract little children. Importantly, the fact that some of the 9-year-olds and the 12- and 15-year-olds relatively rarely project speech and instead prefer forms that result in less lively narrations, must be put in relation to the fact that the frog story is primarily a children's story. Thus, the patterns of use among these (pre-) adolescent narrators may be explained by a lower degree of motivation than among other narrators included in the corpus. I return to this question in Section 5.2.

4.3.5   Summary of the results

A general conclusion from the analyses of speech projections in the spoken Frog story data, is that it is not so much the *degree of use* of speech projections that differs between the age groups, but *the way in which they are used* and *how they are distributed.*

The youngest narrators (the 3-year-olds) had some difficulties in narrating without conversational scaffolding by the adult conversation partner. This is in accordance with the findings presented on three-year-old Harry and Tea and the peers playing with the doll house in Section 4.2. The 3-year-olds telling the frog story rarely had an overall global structure (as evidenced by their low scores on the three core plot components, see Table 4:8 on page 203), and most of the children at this age kept moving into and out of the picture description mode and the narratives (or conversations) took place in the present tense.

Most 3-year-olds produced at least one speech projection but in half of the cases they failed to make clear to the listener (or at least the coder) who the speaker was. By way of comparison, the 4- and 5-year-olds failed in only about 15% of the cases to make clear who the intended projected speaker was.

The proportion of speech projection clauses was around 5% in all three groups of pre-school children (the 3-, 4-, and 5-year-olds), and common to these groups of narrators was the fact that framed direct speech was the most frequently produced form, to be followed by free direct speech, while indirect speech was used only occasionally. In the narratives of the pre-school children, the boy and the dog were typically the main protagonists and verbal content was projected onto them (although animal cries were also projected onto the dog in the 3-year-old's narratives).

Turning to the school-age children, the 9-year-old group of narrators turned out to be homogeneous in the sense that most narrators produced speech projections, they preferred past tense as their anchoring tense, and they often started out their storytelling with a fairy tale-like opening formula. In other respects the group was heterogeneous, and we can divide the types of narratives into two, where the first type of narration was represented by two narrators (ANT09SP and LIE09SP). Their narratives consisted of around 85% speech projection clauses and the overwhelming majority of the speech projections were of the free direct type delivered in a dialogue exchange fashion resulting in lively narrations. Speech was projected onto the whole range of animal characters (sometimes in the form of words, and sometimes only by onomatopoetic sounds), and the act of projecting speech was not at all tied to what was explicitly seen in the pictures, but many speech projections were "invented". Since speech act verbs and speaker identities were verbally expressed only rarely, some utterances caused problems in tracking changes of speaker perspective.

The second narrative type in the 9-year-old group was represented by those narratives (n=12) including few if any speech projections. Past tense was typically the dominant tense, they succeeded in making clear who the projected

speaker was, and the hushing-scene was the most frequent speech projection evoking scene. The narratives of this group resembled the ones of the 12- and 15-year-olds. Most narrators in these latter groups on some occasion made use of speech projections, but the mean value for proportion of speech projection clauses was low (approx. 3%) in both groups. In contrast to the younger groups of narrators, the 12- and 15-year-olds were likely to make use of indirect speech. In all but two cases, the boy was the speaker, which indicates that the story was mainly narrated through the eyes of the boy. The hushing scene was the one most often referred to in the context of speech projecting. Mainly due to the preference of indirect speech, and the non-existence of free direct speech, the narratives had a distanced quality.

In several respects, the adult narratives resembled the ones produced by the pre-school children. For example, the adult narrators clearly favored framed direct speech over indirect speech, and produced some forms of free direct speech as well. The use of these forms resulted in a decreased distance between narrator and characters, and similarly to the 5-year-olds, the adult narrators often projected speech to the boy in relation to those pictures where the boy is seen calling. Many of the adults projected speech onto a range of different story characters and these quotes were verbal rather than in the form of animal cries. The frequent changes in perspective were successfully employed in that the speaker identities were made clear in all cases of speech projecting, (including the cases of free direct speech). In addition, the quotes made by the adult narrators were longer (in terms of included clauses) than in any of the other age groups.

4.4 Cross-modal analyses

In the preceding sections I noted that the school children's oral frog story narratives in several respects differed from the pre-school children's narrations on the one hand, and from the adult's oral narratives on the other. I will now further explore the developmental issues from a cross-modal perspective by

comparing the written Frog story data by the 9-, 12- and 15-year-olds and the adults, to the spoken data of the same age groups.

The disposition in this section is similar to the preceding one. Thus, the first subsection (Section 4.4.1) deals with more general narrative features (tense anchoring and lengths of the narratives). In 4.4.2, the number of speech projections in the spoken and the written narratives are compared and in the following subsection (Section 4.4.3) the distributions of forms of direct and indirect speech are explored. In 4.4.4 I discuss how the forms are used by the narrators of different ages and examine if the forms are used differently in the spoken and the written narratives. The section is concluded with a summary of the results.

4.4.1 Tense use and narrative length

In Section 4.3.1 a general preference for the 9-, 12-year-olds and the adults to anchor their spoken narratives in past tense became evident, whereas ten of the 15-year-olds chose present tense as the dominant tense (cf. Figure 4-7 on page 206). The same type of analysis has been carried out for the written narratives and this yields the figures as are shown in Table 4:11 below (the numbers from the analysis of the spoken narratives are repeated for clarity and in order to facilitate comparisons).

	PRESENT	PAST	MIXED	total
09SP	4	10	0	14
09WR	1	12	1	14
12SP	3	10	1	14
12WR	1	7	6	14
15SP	10	4	0	14
15WR	8	5	1	14
20SP	5	8	1	14
20WR	4	9	1	14
total	36	65	11	

Table 4:11: Tense anchoring in the spoken narratives versus the written narratives; by age.

The table shows that the same pattern that was true for the spoken narratives applies to the written narratives as well. The only difference worth noticing is that six written narratives in the 12-year-old group are of the mixed type and this is the case in only one spoken narrative. [1] Yet, the difference in tense use between speaking and writing in this age group is not statistically significant (χ^2 (2) = 1.333; $p = 0.513$).

The length of the narratives has been calculated by means of included clauses. Table 4:12 provides a summary overview of the different age groups, the spoken and written narratives, range, mean length, and standard deviation.

Data	No of narr.	Total	Range	Mean	Std. Dev.
09SP	n=14	911	32—90	65.1	17.8
09WR	n=14	810	38—117	58.9	21.9
12SP	n=14	965	36—134	71.5	26.3
12WR	n=14	824	36—99	65.0	17.8
15SP	n=14	1064	43—151	76.0	33.6
15WR	n=14	967	39—126	68.4	24.8
20SP	n=14	1750	72—201	125.0	45.8
20WR	n=14	2012	56—259	143.7	63.0

Table 4:12: Range and mean number of clauses by age group; speech and writing.

This table shows that the groups of 9-, 12-, and 15-year-olds, as groups, do not deviate too much from each other, whereas the adult group produces longer narratives, and displays great individual differences (Std. Dev). When comparing the length of spoken versus length of written narratives in the four age groups, the paired *t*-test showed that there were no significant differences. Thus, the fact that the narrating of the story was carried out in two different

---

[1] Four of these written narratives were produced after the spoken narratives were told, whereas two of the written narratives being of the mixed type were produced before the spoken narrative.

modes of production, having different processing constraints tied to them (cf. discussion in Section 2.2.3), seems to have had little effect on the length of the story in terms of clauses.[2] Still, it is worth noting that for all three groups of school children the mean length of the spoken narratives is greater than for the written ones, whereas the contrary is true for the adult narratives (the mean length of the written narratives is greater than for the spoken ones). This is most likely a result of the adults being more skilled at writing than the younger children, as well as having more patience with the task. It is the case that the frog story consists of twenty-four pictures depicting events, and encoding the events in each picture by means of writing takes its time, due to the fact that writing is a slower mode of production that demands more of the producer than speaking does.

In order to examine developmental aspects, ANOVA was applied to see if there were significant differences in length between age groups, in speech and writing, respectively. Age showed to have an effect in speaking ($p <.0001$) as well as in writing ($p <.0001$). An Ad Hoc test (Fisher's PLSD) showed that the adult narratives were significantly longer than the narratives of the 9-, 12-, and 15-year-olds, whereas there were no significant differences among the three groups of school-age children. Thus, the number of clauses included in the narratives does not differ to any great extent among the children at different school-ages. Berman & Slobin (1994) in their study of frog story narratives found, among other things, that the adults produced longer narratives than the 9-year-olds. Berman & Slobin did not include adolescent narrators, but with these findings I can add to the developmental picture the assumption that constructions of long and elaborate frog story narratives is a late achievement. This seems to be particularly true of the written narratives.

One 12-year-old, ROL, evidenced mixed tense usage in both conditions (speaking and writing).
[2] One explanation for these results may be found in the experimental conditions of the elicitation of the narratives. Although the subjects were not instructed to do so, they may have interpreted the task as to "tell the same story twice". Since the elapsed time between the two occasions of

4.4.2 Proportion of speech projection clauses

The corpus of written narratives (N=56) comprises in total 4,613 clauses. 627 of these clauses have been coded as speech projection clauses. Table 4:13 shows descriptive statistics of the written narrative data:

Age group	N of narr.	Mean	Std. Dev.	Min.	Max.	Median
9	14	23.4	27.2	0	84.3	9.8
12	14	4.9	4.4	0	13.6	4.2
15	14	2.7	3.0	0	8.2	2.1
Adults	14	15.2	13.0	0	41.7	13.0

Table 4:13: Descriptive statistics for all age groups; number of clauses involved in speech projections compared to total number of clauses (percentages); written narratives.

As was also the case with the spoken data, the mean value in the 9-year-old group is high (23.4%), as is the maximal value (84.3%), and the individual differences are great (Std. Dev.: 27.2%). The 12- and 15-year-olds produce comparatively few speech projection clauses (Mean: 4.9% and 2.7%, respectively) and the individual differences are relatively small. In contrast, the adult narrators have a fairly high mean value (15.2%) and the individual differences are great. The mean and standard deviation values in the written frog story narratives are illustrated in Figure 4-13 on the next page. Figure 4-14 exhibits the mean values in the written data contrasted to the corresponding mean values for the spoken data.

elicitation (speaking versus writing) did not generally exceed half an hour, the narrators may also have had memory traces from the first occasion affecting the second narration.
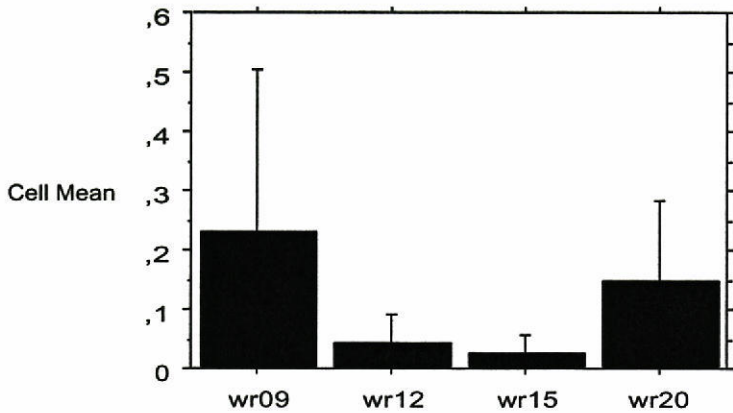
Figure 4-13: Interaction bar plot for proportion of speech projection clauses (mean values), written narratives, 9-, 12-, 15-year-olds and adults. Error bars: Standard deviations.
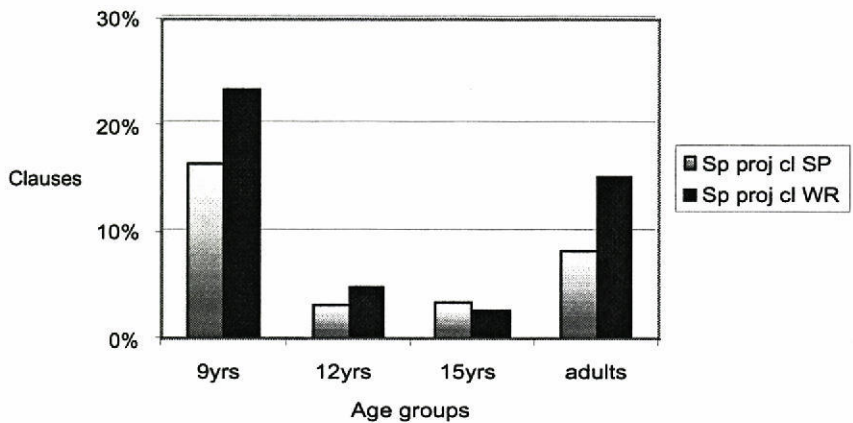

Figure 4-14: Percentage speech projection clauses in the spoken versus the written narratives; mean values for each age group.

Comparing the spoken and the written narratives within the age groups, the paired *t*-test showed no significant differences, except for the adult group ($p =$

0.05). Hence, the adult narrators use proportionally more speech projections in writing than in speaking

The 9-year-olds constitute a fairly heterogeneous group regarding the spoken narratives, in that two narrators rely far more on speech projections than the remaining twelve narrators do (in this latter group of narrators no narrative consists of more than 13% speech projection clauses). This fact is reflected in a median value as low as 4.5%. In the written data of the 9-year-olds, in contrast, the median value is nearly 10%. As can be seen in Figure 4-15 below, which indicates percentages of speech projection clauses for all the 9-year-olds and their spoken and their written narratives, as many as five 9-year-old writers exceed the 25%-level. This means that in these particular narratives, at least 1/4 of all clauses are speech projection clauses. Moreover, nine out of the fourteen narrators produce a greater proportion of speech projections when writing than when speaking.[3]
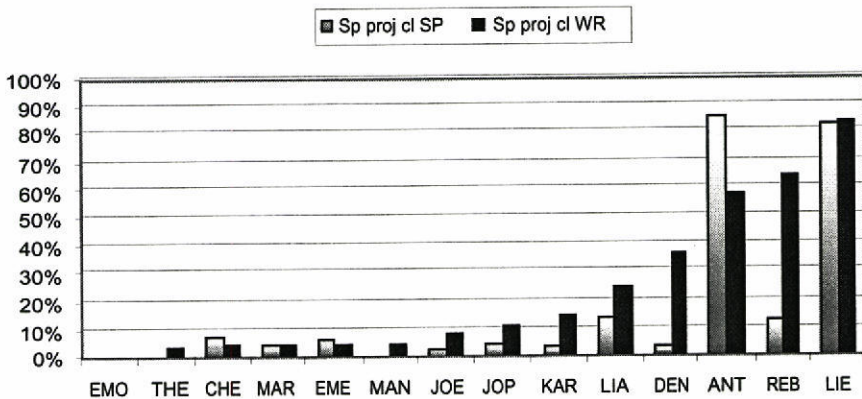

Figure 4-15: The distribution of proportional use of speech projection clauses, by 9-year-old subjects (N=14); in speaking and in writing.

---

[3] A statistical analysis (the Wilcoxon Signed Rank test) was carried out to see if the order of production had any effect on amount of use, and the outcomes were non-significant (speaking to writing (N=7): $z = -1.15$; $p = 0.25$; and writing to speaking (N=7): $z = -1.52$; $p = 0.13$). Hence, the order of production does not explain why the majority of the subjects produce a greater proportion of speech projections when writing than when speaking.

4. Results

The difference shown between speaking and writing is indeed noteworthy, and a similar effect is discovered when we consider the total corpus. The size of the corpus of written narratives (covering the age groups of 9-, 12-, 15-year-olds and adults), is about 4,600 clauses and the size of the corpus of spoken narratives produced by the same subjects, (i.e., the 3-, 4-, and 5-year-olds excluded), is almost the same: 4,690. Yet, in the spoken corpus the number of speech projection clauses is 385 whereas it is 627 in the written corpus, (cf. also the indications in Figure 4-14 above). Comparing all the written narratives to the spoken by applying the paired t-test, gives significant results on the 0.05 level ($t(55)$ = -2.55; p = 0.014). Hence, leaving individual differences and matters of development aside, the contents of the data indicate that speech projection clauses are used to a greater extent in written narratives than spoken narratives. Why is this the case? One explanation for this is that the consideration of different perspectives that needs to be taken into account when projecting speech is a cognitively demanding process (cf. discussion in 2.1.3). Therefore, speech projections are easier to produce in writing when the subjects have more time to plan and reflect (see Section 2.2.3). Moreover, as has been discussed previously (e.g. in Section 2.4), direct and indirect speech frequently occur in genres like written literary texts, fairy tales and novels. It is likely that the subjects included in this study have experience of these kinds of genres and make use of this knowledge when writing the frog story (cf. Tannen's notion of literary discourse and the discussion on "relative focus of involvement" in Section 2.2.3).

4.4.3   Development of forms

In total, 297 speech projections were coded in the written Frog story data. Out of these, 26 cases were coded as either UnclearType or UnclearIf. This means that 271 speech projections were coded according to the categories of form: Indir; DirX (DirBef, DirAft, DirEnfr); and FreeDir. Of these 271 speech projections, 106 were found in the 9-year-old group, 25 in the 12-year-old

group, 12 in the 15-year-old group, and 134 in the adult group of narrators. Below is a figure illustrating the distribution of forms in the four age groups.
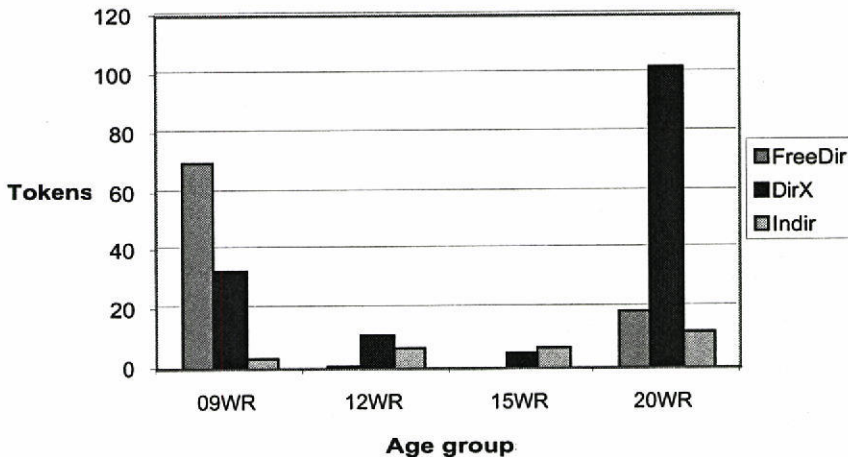

Figure 4-16: The distribution of forms of speech (raw numbers) in the written narratives; all age groups; forms coded as UnclearIf or UnclearType are not included.

Comparing this figure to the one showing the distribution in the spoken narratives (Figure 4-10 on pagepå sidan 213), it is evident that the distributions are similar in both modes of production. For instance, the relationship between the three bars in 09WR look similar to the relations between the bars in 09SP (see page 213), i.e., there are more tokens of free direct speech than of framed direct speech, which, in turn, occurs more frequently than indirect speech. However, in contrast to the spoken data, (where two of the 9-year-old narrators account for the greater part of the forms of free direct speech in this age group), several of the narrators use free direct speech in writing. This has consequences for the success of speaker perspectivizing as will become evident below. As far as the adult narrators are concerned, they produce more than twice the number of speech projections in their written narratives in comparison to their spoken ones (134 versus 53 instances, this fact is reflected in Figure 4-14 above as

well). The relationships among the three types of forms are however similar in the spoken and written narratives of this age group, suggesting that the narrators make similar choices of forms in the two modalities.

So far, the different types of framed direct speech possible for use have not been taken into consideration. In Section 2.4.3.3 it was mentioned that the clause with the speech act verb may precede, interrupt, or follow the quote, and I said that it is open for discussion whether any of the types commonly occur in spoken or written discourse. The results from the examination of children's early reading books that were reviewed in 2.4.4.2, suggest that the framing clauses most often follow the quote in writing. Indeed, this is confirmed in this written production data. In all age groups the quotes more commonly precede the clause with the speech act verb (DirAft) in writing. Interestingly, in speaking the conditions are quite the reverse, i.e., the quotes are preceded by that clause (DirBef). Below is a table summarizing the numbers:

	DirBef		DirAft		DirEnfr		TotalDir	
SP	43	**56%**	24	31%	10	13%	77	100%
WR	28	18%	92	**60%**	34	22%	154	100%

Table 4:14: Distributions of types of direct speech in all spoken narratives (N=56) and in all written narratives (N=56): Direct speech where the quote is preceded by the clause with the speech act verb (DirBef); is followed by that clause (DirAft); or interrupted by it (DirEnfr).

As I have indicated with bold face in the table, DirBef make up 56% of all instances of framed direct speech in the spoken data, whereas the corresponding figure in the written data is only 18%. However, in the written data, as much as 60% of the framed direct speech are of the DirAft type, and it is even the case that DirEnfr is more frequently employed in writing than DirBef is (22% vs 13%). In her spoken narrative, the 15-year-old JEL expresses it as follows, *Tim tittar på hunden å säjer dumma hund varför gör du så?*, 'Tim looks at his dog and says, silly dog why do you do that?', whereas in her written version of the same episode she expresses it in the following way,

dumma hund, säger Tim, och tar upp honom i sin famn, 'silly dog, says Tim, and picks him up'. While the adult narrators are the ones producing most of the instances of DirEnfr, the 9- and 12-year-olds also use forms of DirEnfr in their writing. For instance, JOP09WR writes *Tyst, sa pojken till hunden du skrämmer bort groda om du skäller så,* 'Quiet, the boy said to the dog, you frighten the frog if you bark like that'.

A psycholinguistic explanation for the general preference of DirBef in speaking, is that the narrator gains some planning time regarding the contents of the quote. This means that the framing clause functions as a signal to the listener *that* a quote is to appear, and at the same time, the narrator gets some extra time to plan *what* to include in the quote.[4] The speech is produced and perceived on-line (cf. Section 2.2.3) and introducing a quote with a framing clause thus alerts the listener that speech is to be quoted. In writing, quotation marks or a hyphen serve this same function, and therefore, the clause with the speech act verb can be placed after the quote. Additionally, in writing the production rate is slower, there is more time to plan, and the writer is then more free to choose form of expression. These conditions of production may explain why DirEnfr is employed in written discourse to a greater extent than in spoken discourse, since this form has a more complex structure than the other forms and requires some cognitive effort to produce. Moreover, the reader of a narrative also has more time to process the discourse and the shifts in perspective than the listener of a spoken narrative.

4.4.4 Use of the forms

When studying the projected speakers in the written narratives, the same tendencies emerge as in the spoken narratives. This means that the 12- and 15-year-olds exclusively project speech onto the boy, whereas the 9-year-olds and

---

[4] The fact that CUL03SP hesitates before he expresses the quote in the example of framed direct speech (DirBef) provided in Figure 4-11 on page 214, may work as an illustration of this. CUL03SP, by means of the hesitation markers, signals that he has not finished his planning of what exactly he wants to express in the quote and/or how to express it.

the adult narrators project speech onto a variety of story characters. Thus, besides the boy, the dog says things (fifteen cases are found in the 9-year-olds' narratives, and twenty-one in the adults'), as do the frog, the gopher, the deer, the owl, and the bees. About one third of the speech projections by 9-year-olds are of this kind, and in the adult group nearly half (43%) are speech projections onto animal characters. This latter percentage is higher than it was in the corresponding spoken narratives, that is, the adult narrators project speech more often onto animal protagonists (and less often to the boy) in the written narratives. Moreover, the speech projected onto animal figures, always consists of words in the adult group (cf. Example 4.24 on page 218), whereas in the younger age group, the speech projected is often only animal cries.

Turning to scenes that are conducive to speech projections, the 9-year-olds and the adults show the same pattern. Around 75% of the speech projections are made in relation to scenes other than those included in the coding scheme. This percentage is clearly higher than it was in the spoken narratives (around 50%). Approximately 10% of the speech projections are produced in relation to any of the calling scenes, and nearly 10% to the hushing scene. The 12- and 15-year-olds typically project speech to the boy when the boy asks his dog to be quiet (the hushing scene). These results are similar to what was found in the analyses of the spoken data.

The function of distancing yields the same pattern in the written narratives as in the spoken data, i.e., the 9-year-olds' and adults' preference for using framed and free direct speech contributes to a decrease of the distance whereas the older school children's avoidance of speech projections and more frequent use of indirect speech, increases the sense of distance.

In Section 2.3.3 I contended that direct and indirect speech have been neglected to a great extent in previous analyses of frog story narratives. This is unfortunate since speech projections can play a prominent role in the narration

234

of the story. I would like to illustrate this point with an extract from an adult narrator (RAG20WR) writing to pictures 14 to 21.[5]

Example 4.25

Pojken klättrade ända upp på stenens topp och tog där tag i en gren för att hålla balansen.
- Finns det någon som har sett grodan här?, ropade pojken
Plötsligt reste sig grenen som pojken höll i. Det var ingen gren, det var en hjort.
- Vill du verkligen veta var grodan är?, sa hjorten till pojken.
- Ja , det vore roligt, sa pojken lite förskräckt över situationen. Hjorten bar pojken fram till ett stup och skrattade,
- Nu ska ni få träffa grodan.
Hunden var i full galopp och följde hjorten och pojken. Men hunden upptäckte inte att hjorten plötsligt stannade till vid stupet. Både pojken och hunden föll framåtstupa ner för branten. Dom skrek och skrek.
- Hjälp, hjälp!!!
Hjorten skrattade snällt när han såg dem båda falla i den lilla dammen.
- Såja, om ni nu är riktigt tysta och snälla så lär ni nog träffa den lilla grodan igen, sa hjorten.
Våta, men glada satt sig pojken och hunden för att lyssna. Där bakom det gamla trädet kunde dom båda höra ett litet kvackande ljud. Hunden blev så glad att han började skälla.
- Tyst! Annars skrämmer du iväg den lilla grodan!
Hunden tystnade och smög sig nära pojken för att fortsätta jakten på den lilla grodan.

'The boy climbed all the way up to the top of the rock and grabbed a branch to keep his balance.
"Is there anyone here who has seen my frog?", the boy shouted.
Suddenly the branch that he had held onto rose. It wasn't a branch, it was a deer.
"Do you really want to know where the frog is?", the deer said to the boy.
"Yes, that would be fun", the boy said a little frightened about the situation.
The deer carried the boy to a cliff and laughed,
"Now you'll meet the frog."

---

[5] These pictures show: the boy has climbed up the rock and is holding onto what he believes are branches; he is calling for the frog; a deer hiding behind the rock picks up the boy on its antlers and runs off with the boy; the dog in hot pursuit; the deer throws the boy and the dog down a cliff; they land at the bottom in a pond below the cliff and the dog climbs on the boy's head; the boy puts his hand to his ears and tells the dog to be quiet; they climb over an old log. (For picture series, see Appendix.)

The dog ran along at a gallop and followed the deer and the boy. But the dog didn't notice that the deer suddenly stopped at the edge of the cliff. Both the boy and the dog fell headfirst down the cliff. They yelled and yelled.
"Help, help!!"
The deer laughed kindly when he saw them fall into the little pond.
"Well, if you are really quiet and good now you probably will meet your little frog again", the deer said.
Wet, but happy, the boy and the dog sat down to listen. There, behind the old log, they could to hear a little "ribbeting" sound. The dog became so happy that he started to bark.
"Quiet! Otherwise you'll frighten our little frog away!"
The dog got quiet and sneaked close to the boy in order to continue hunting for the little frog.'

This is a fairly elaborate narrative, and it contains quite a few instances of direct speech. Direct speech has been argued to have a major function in making a story expressive and "a brief 'time out' from the telling of the story proper" (Toolan 1998b: 625, see Section 2.3.1); what, then, will remain if we take away all speech projections from this narrative passage? In the extract below, all forms of framed direct and free direct speech have been removed.

Example 4.26

Pojken klättrade ända upp på stenens topp och tog där tag i en gren för att hålla balansen. Plötsligt reste sig grenen som pojken höll i. Det var ingen gren, det var en hjort. Hjorten bar pojken fram till ett stup. Hunden var i full galopp och följde hjorten och pojken. Men hunden upptäckte inte att hjorten plötsligt stannade till vid stupet. Både pojken och hunden föll framåtstupa ner för branten. Dom skrek och skrek. Hjorten skrattade snällt när han såg dem båda falla i den lilla dammen. Våta, men glada satt sig pojken och hunden för att lyssna. Där bakom det gamla trädet kunde dom båda höra ett litet kvackande ljud. Hunden blev så glad att han började skälla. Hunden tystnade och smög sig nära pojken för att fortsätta jakten på den lilla grodan.

'The boy climbed all the way up to the top of the rock and grabbed a branch to keep his balance. Suddenly the branch that he had held onto rose. It wasn't a branch, it was a deer. The deer carried the boy to a cliff. The dog ran along at a gallop and followed the deer and the boy. But the dog didn't notice that the deer suddenly at the edge of the cliff. Both the boy and the dog fell headfirst down the cliff. They yelled and yelled. The deer laughed kindly when he saw them fall into the little pond. Wet, but happy, the boy and the dog sat down to listen. There, behind the old log, they could to hear a little "ribbeting" sound.

The dog became so happy that he started to bark. The dog got quiet and sneaked close to the boy in order to continue hunting for the little frog.'

The variant in Example 4.26 is far from as expressive as the original one, but the storyline is fairly intact and intelligible to us. Still, it lacks certain crucial components. First, it is not evident that the boy is calling for the frog when standing on the rock holding onto the stick. Second, it is not clear that the reason why the deer starts running away with the boy on his antlers and then throwing him over the cliff, is actually that the deer wants to help the boy and the dog find their frog (which he believes can be found below the cliff). Third, the information that it is because the boy exhorts the dog to be quiet (since the barking would otherwise frighten off the frog) that the dog gets quiet, is not at all provided. Thus, the matter is not as simple as regarding forms of direct speech only as some kind of expressive marker that can easily be removed, but in addition, speech projections have crucial narrative and plot-advancing functions.

Turning to the issue of speaker perspectivizing, Figure 4-17 below shows the degrees (percentages) to which speaker identities are made clear in the speech projections in the written narratives.
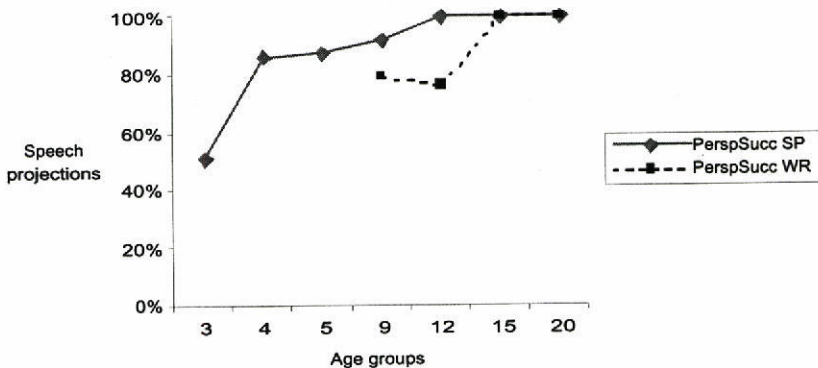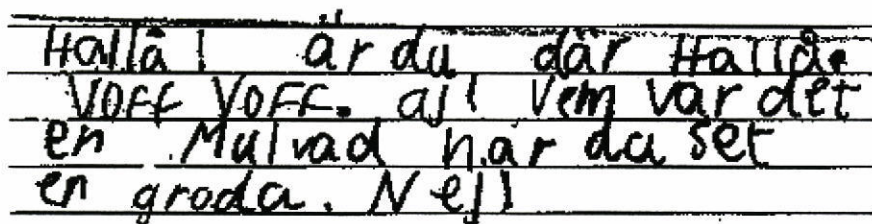

Figure 4-17: Speaker perspectivizing; success rate (percentage of speech projections where it is made clear who is the projected speaker), all age groups; speaking vs writing.

237

(The rates for the spoken narratives, already presented and discussed in Section 4.3.4, are included for comparison.) What is evident from this figure, is that the 15-year-olds and the adults always succeed (100%) in making clear who the projected speaker is when they include speech projections in their writing. 9- and 12-year-olds, however, fail in around 20% of the cases, and fail more often than when narrating orally.

There is one important explanation as to why the younger school children seem to have greater problems speaker perspectivizing in writing than in speaking. When writing you cannot use changes in the tone of voice in order to convey shifts in perspective. When speaking, prosodic modifications can be a very powerful device to accomplish this, especially in the cases where direct quotes are not verbally framed (free direct speech). In writing, the conventional way of marking quoted speech is to use quotation marks. It should then be observed that the 9-year-old writers who show a preference for free direct speech, rarely use these kinds of markers. Observe the example below, written by 9-year-old LIE in relation to pictures nine and ten (the boy searches in a gopher hole and the dog is sniffing at a beehive; a gopher comes out of the hole and bites the boy on the nose):

Example 4.27

'Hello! are you there Hello.
Bow-Wow. ouch! Who was that
a Gopher have you seen
a frog. No!'

LIE uses punctuation marks such as exclamation marks and dots. Moreover, she occasionally capitalizes letters. However, despite the dialogical character of

the writing she does not use any quotation marks to show where a speech line starts and ends, and she does not even start on a new line for a new speaker. In this way, she does not make clear who the intended speakers are.

In one case, a 9-year-old marks off quoted speech with a colon (LIE09WR: *sen sa pojken: Nää Nu måste vi gå och lägga oss*, 'then the boy said: No Now we have to go to bed'), and in a couple of other cases a comma is used to separate the mere speech projected and the clause with the speech act verb (e.g., LIA09WR: *Jag hör nåt, sa pojken*, 'I hear something, the boy said'). However, *none* of the 9-year-olds use quotation marks or dashes. For the reader of the narrative, this fact results in difficulties when keeping track of changes in perspective.

The 12-year-olds do better on punctuation, although the conventional ways to mark speech lines (as illustrated in Example 3.3 and Example 3.4 on page 147) in Swedish are not fully employed.[6] What causes trouble for the reader, is that in some cases it is difficult to determine whether the writer has intended to project speech or not (this has been coded as UnclearIf in the Form category). The fact that many of the 12-year-old writers anchor their narratives in present tense further contributes to the deictic changes becoming less transparent. The 15-year-olds, in contrast, succeed in making speaker identities clear. This can partly be explained with reference to the fact that these writers prefer indirect speech. The adult writers make use of many direct quotations, and mark them clearly and conventionally, so speaker perspectivizing is not a problem for this age group. Moreover, when the adult writers use free direct speech, they embed the utterances in the larger discourse context, which helps the reader to infer who the speaker is. Consider the examples below extracted from JOT20WR and KAP20WR, respectively.

---

[6] For example, KES12WR writes: *schyy! sa Villy till Sammy*, 'quiet! Villy said to Sammy'. Here the exclamation mark indicates what Villy expresses to Sammy. Note that neither quotation marks, nor hyphens are used.

Example 4.28

Pelle måste ideligen påminna Plutt att inte vifta så förfärligt på svansen.
- Tyst nu!

'Peter must remind Plutt over and over again not to wag his tail so furiously.
"Be quiet now!"'

Example 4.29

- Vad är det? frågade Bert förvirrat.
- Tror du inte att han kan ha gömt sig någonstans?
- Varför skulle han gömma sig?
- Tja, vi känner ju inte varandra så väl. Han kanske blev rädd.
- Knappast. Kom nu, vi går ut.
Bert började bli otålig.

'"What is it?" Bert asked confused.
"Don't you think that he might have hidden somewhere?"
"Why should he hide?"
"Well, we don't know each other that well. Maybe he got frightened.
"Hardly. Come on, let's go out."
Bert started to get impatient.'

In Example 4.28 there is no verb of saying present and although there is no traditional framing clause, it can easily be inferred from the preceding context that Peter is telling Plutt to be quiet rather than vice versa. In Example 4.29 the speakers' identities can be inferred from the fact that the free direct speech utterances are part of an exchange of utterances, starting with a question by Bert. In contrast to the younger narrators, all cases of free direct speech in the written narratives of the adults are embedded in the narrative discourse in a sophisticated fashion as illustrated in Example 4.28 and Example 4.29. The adult narrators are well aware that prosodic cues are not present in writing, and they adjust their speech projection strategies accordingly.

Projecting speech in writing means rendering speech in a different modality. In order to represent speech, you can use orthographical conventions such as punctuation marks. Using quotation marks, for instance, is a powerful way of signaling that speech is projected, and LIE09WR's use of exclamation marks in Example 4.27, indeed creates certain effects. Besides punctuation

marks, there are other ways of depicting the speech of others such as tone of voice, pitch and aspects of delivery in writing. Such options were discussed in 2.4.3.3, but similar attempts were also found in the written data.

Reduplicating letters in order to depict a lasting cry and/or emphasis, is found in all age groups: *Hååååå sa hunden*, 'Hooooo, the dog said', (9-year-old); *Guuud så gulliga*, 'Goooosh they are so cute', (12-year-old); *Schyyy sa Peter*, 'Hushhhh, Peter said', (15-year-old); "Groooodan... var é du?", 'Froooog... where are you?', (adult). Capitalizing letters to mark emphasis or pitch, however, is only found twice in the adult data, and in one of the 12-year-olds' narratives (SAB12WR):

Example 4.30

'Suddenly Sniffy finds a
biscuit and wants to jump from
the window! He does!
- NO SNIFFY!!!

He jumped!'

Note that the subject also uses multiple exclamation marks.

In the adult narratives, there are attempts at marking aspects such as speech register. JOM20WR projects the speech line "*Va' e' re om?*", 'What is it all about?', to the gopher, where the accents and the reduced forms bring about a sense of spoken language and a particular register that deviates from the one attributed to the boy in the same narrative. The same writer depicts stuttering on the part of the boy at a later point: "*F-förlåt*" stammade Sebastian

fram, '"S-sorry", Sebastian stammered'. Here the hyphen is used to depict a disturbance in the flow of speech. Another writer, CAR20WR, uses the hyphen in order to mark lengthening of a syllable in the calling for the frog instead: *"Kva-ack, var är du?"*, 'Rib-bet, where are you?'.

Although there are examples like those described above to depict different aspects of delivery in the written narrative data, they are comparatively rare. The adults use depict aspects of delivery more than the children, but not even the adults use it to a great extent. Indeed, it may be problematic to depict such aspects in writing, and in certain cases it is not even desirable to do that. For stylistic reasons, for instance, you may avoid using reduplicated letters or capitalized words. When you cannot, (or when you do not want to), depict aspects of delivery in the ways as exemplified above, you can describe them. You can do it in several ways, for instance, by using speech act verbs.

The 3-, 4-, and 5-year-olds use the speech act verbs *säga*, 'say'; *ropa*, 'call'; *skrika*, 'shout'; *fråga*, 'ask'; and *skälla*, 'bark'. The 9-year-olds use *säga*, *ropa*, and *skrika* in both speaking and writing, as do the 12-year-olds. One instance of *viska*, 'whisper', is found in a written narrative in this age group as well. The 15-year-olds only use *säga*, in their spoken narratives, whereas they add *be*, 'ask'; *fråga*, 'ask'; and *vråla*, 'howl' in their written narratives. Thus, there is a core of speech act verbs used by the children and adolescents in the study, namely, *säga*, *ropa*, and *skrika*. These are fairly neutral verbs of saying, where *säga* is the most neutral and frequent speech act verb in Swedish. The frequent occurrence of *ropa* and *skrika* is most certainly because the story is about a search for a frog and the pictures show the boy with his hand next to his mouth. In the 15-year-old group, some other types of speech act verbs that are not so neutral (e.g., *howl*) can be discerned, and these subjects also use more types in writing than in speaking. The weak tendency of more varied forms in writing than in speaking in the 15-year-old group, is, on the other hand, very strong in the adult group. In the spoken narratives, thirteen different types of speech act verbs are used. In the written narratives, as many as thirty-three different types are found:

be	'ask'	*skratta*	'laugh'
fråga	'ask'	*skrika*	'yell'
fräsa	'hiss'	*skrocka*	'chuckle'
förebrå	'reproach'	*skälla*	'bark'
föreslå	'suggest'	*stamma (fram)*	'stammer (out)'
genmäla	'reply'	*svara*	'answer'
hoa	'hoot'	*säga*	'say'
hojta	'shout'	*undra*	'wonder'
hyssja	'hush'	*utbrista*	'exclaim'
hånskratta	'jeer'	*utropa*	'exclaim'
lova	'promise'	*viska*	'whisper'
morra	'growl'	*voffsa*	'bark'
muttra	'mutter'		
pipa	'squeak'	*börja*	'start'
påminna	'remind'	*försöka*	'try'
ropa	'call'	*höras*	'be heard'
råda	'advise'	*låta*	'sound'

Table 4:15: Types of speech act verbs in the written narratives of the adult controls.

(The four final verbs do not have an inherent meaning of a speech act, but these forms have been used in the particular contexts to refer to a speech act.) In addition to these numerous forms, that indeed contribute to more detailed descriptions of the speech than the word *säga*, 'say', the adult narrators often use adverbial modifiers. For example, phrases such as *utbrast pojken förtjust*, 'the boy exclaimed delightedly'; *frågade hunden surt*, 'the dog asked sourly'; *morrade hunden förundrat*, 'the dog growled with puzzlement'; *skrek pojken av skräck*, 'the boy screamed in fear'; *'skällde hunden oroligt*, 'the dog barked anxiously'; and *hoade ugglan elakt*, 'the owl hooted in a mean fashion', are found in the material.

These examples of framing verbs serves to illustrate to that the adult narrators use different ways of encoding and depicting events in their writing than they do in their oral recountals of the story, and I argue that they make these distinctions to another extent than the younger narrators do. The development of awareness of modality-specific properties and conditions will be further discussed in Chapter 5 below.

4.4.5 Summary of the results

The analyses of the spoken data indicated a U-shaped-like development in that, on the one hand, the narrations by the 9-year-olds had certain aspects in common with the narrations by the adults, and on the other hand the narratives by the 12- and 15-year-olds were in many respects similar to each other. This pattern was even more pronounced in the written data. Framed direct and free direct speech were used to a greater extent than indirect speech in the 9-year-old and adult groups, whereas the 12- and 15-year-olds generally produced very few speech projections. However, indirect speech occurred more frequently in these age groups than in the other age groups. The 9-year-old and adult narrators projected speech onto several of the story characters, in all different kinds of scenes, and were thus not tied to what was explicitly displayed in the pictures when projecting speech. This fact was particularly obvious in the adults' written narrations where the animal characters as actors were given a considerable role. When projecting speech onto animal characters, the younger children often projected animal cries, whereas the adults projected words and propositional content. The 12- and 15-year-olds all chose the perspective of the boy, and the most frequent speech projection was onto the boy when he is seen hushing his dog in picture twenty.

Similarly to what was found in the spoken data, the written speech projections made by the 9-year-old narrators and the adults contributed to a decreased distance between the narrator and the protagonists. In the 12- and 15-year-old group, in contrast, this distance was increased and the narrator (or reporter) role upgraded. The fact that few speech projections were used in general, and indirect speech was preferred, strengthens the impression of these narratives as reported rather than enacted.

In one respect the written narratives by the 12-year olds were more similar to the 9-year-olds, and that concerned the degree to which it was made clear who the intended projected speaker was (speaker perspectivizing). Both groups scored lower in writing than in speaking, and this was mostly due to the fact

that quotation marks were rarely used. Thus, whereas the 15-year-olds and the adults had no problems with this issue (i.e., conveying shifts in perspective), younger school children had more difficulties indicating that the marking of quotes takes some time to learn.

It was shown above that attempts at depicting aspects of delivery were found in the written corpus. The strategy of reduplicating letters to depict prosodic aspects in writing was found in all age groups, whereas capitalizing letters was a less frequently employed strategy. In the adult narratives attempts at depicting speech register were also found. An analysis of types of speech act verbs used, revealed that Swedish equivalents to *say* and *call* were used by narrators of all ages both in speaking and in writing, and besides them, only a few other types of verbs occurred. The adult narrators, however, differed from the younger ones in that a broad range of different types of verbs of saying was used. This was particularly true of the written narratives, where thirty-three verb types were distinguished, and these were often combined with, and semantically modified by, adverbial clauses. Thus, besides resulting in stylistically captivating renditions, rich information (in this case, information concerning the speech quoted) was provided and packaged into these linguistic constructions.

If we turn to the question of whether we can point out aspects of use that are connected to type of mode of production (speech versus writing), we found in the adult group that speech projections were used significantly more frequently in writing than in speaking. Comparing *all* written narratives (of all age groups) with *all* spoken narratives also showed significant differences, in that speech projections were more frequently employed in writing than in speaking. However, there were no differences among the types of forms chosen by the subjects in the two modalities. Thus, what Tannen (1986) and Chafe (1982) have suggested concerning use of direct speech (i.e., that free direct speech is more common in spoken discourse than in written texts) is not confirmed in these data. There was a slight tendency for framed direct speech to be interrupted by a clause with a speech act verb (DirEnfr) more often in

writing than in speaking (e.g., Tannen 1986), and in line with what Perera (1986) found in children's reading books, the framing clause more often followed the quote in the written narratives. On the contrary, the framing clause more often preceded the quote in the spoken narratives. These facts held for all age groups.

4.5   Free indirect speech

Up until this point I have been concerned with the reporting and projecting of speech in terms of free and framed direct speech and indirect speech. However, there are other ways to convey speech and one is to use free indirect speech. This section examines the occurrence of free indirect speech in the entire corpus. To distinguish these forms, and in order for an utterance to be coded as a candidate of free indirect speech, the following two conditions apply: *a) utterance Y is likely to refer to a speech event*; and *b) if the utterance Y is transformed into a direct quote construction, then the word-order of Y should remain intact whereas the deictic elements need to be shifted.* (See 3.3.5 for more details.) The section is divided into four subsections. In the first one (Section 4.5.1) the use of free indirect speech in the Longitudinal case studies, the three-year-olds and their mothers playing with the doll house, and the 3-, 4- and 5-year olds telling the frog story, is examined. The second section (Section 4.5.2) accounts for free indirect speech use among the 9-, 12- and 15-year-olds and the adults. The third section (Section 4.5.3) presents the results of the analysis of use of mental verbs in the frog story narratives. In the fourth section (Section 4.5.4), finally, the results are summarized.

4.5.1   Pre-school children

The coding process yield extremely few and unclear examples of free indirect speech in the Longitudinal case study material, in the dyads of the doll house play, and in the frog story narratives by the 3-, 4- and 5-year-olds and their

interaction partners. In the interactions of Harry and Tea with their mother and grandmother, I find only a handful of candidates of free indirect speech in total. Interestingly, these cases have certain aspects in common in that they are *speech reports* rather than speech projections. This means that the forms are used in a discourse context where earlier experiences are discussed and reported on. However, the relationship between the free indirect speech report and the original speech uttered is vague, so that the forms have the *character of a paraphrase*. I cannot be sure of what the original speech actually consisted of by interpreting only the free indirect speech form. Additionally, the examples all turn out to include *modals*. More specifically, the modals are *(inte) få*, 'must (not)', and *(inte) vilja*, '(not) wanting to'. In the examples below, all produced by Harry's and Tea's mother, the candidates for free indirect speech are typed in bold and followed by an example where the form has been turned into a direct construction (in squared brackets) and the deictic effects of this transformation are specified.

Example 4.31

Harry 24_16.cha

*MOT:	å så ritade du de h ritade du på armarna här kommer du ihåg de?
%eng:	and then you drew that on your arms here, do you remember
*HAR:	that?
	hmm
%eng:	uhum
*MOT:	**de fick inte du göra**
	[jfr. dir: (då sa jag) "det/så *får* du inte göra"]
%eng:	[lit.] it/that you *could*n't do
	[cf. dir: (then I said) "it/that you *can*'t do"]

Example 4.32

Harry 28_02.cha

*MOT:	sa du att dom skulle ta upp sakerna ?
%eng:	did you tell them that they should pick up the things?
*HAR:	ha
%eng:	yeah

*MOT: **dom fick inte slänga sakerna på golvet**
          [jfr. dir: (du sa) "*ni får* inte slänga sakerna på golvet"]
%eng: they could not throw the things on the floor
          [cf. dir: (you said) "*you can*'t throw the things on the floor"]

Example 4.33

Tea46_22.cha

*MOT: men familjen ville gärna ge bort den hunden dom fråga oss om
          vi ville ha den
%eng: but the family wanted to give away that dog they asked us if we
          wanted to have it
*TEA: ö^ö
%eng: uhm
*MOT: **men de ville vi nte** [jfr. dir: men (vi sa) "(nej) det *vill* vi inte"]
%eng: but we didn't want that [cf. dir: but (we said) "(no) we *do*n't
          want that"]

It is fairly easy to imagine the situation the mother is referring to in Example 4.31 and "hear" how she tells Harry not to draw on his arms, and in Example 4.32 to imagine how Harry told his friends not to throw the things around. Still, it should be noted that common to these examples is the fact that they are not self-evident forms of free indirect speech. The utterances marked in bold may well be only comments by the mother on an earlier event that did not even necessarily include speech. The loose bonds between these occurrences and possible earlier speech events may be particularly obvious in Example 4.33. It is unlikely, not the least because of courtesy reasons, that the mother (or someone) literally replied *we don't want that [your dog]*, when they were offered the dog. Rather, the utterances given in Example 4.33 probably summarize an event that included a great deal more than these utterances express.

In Example 4.34 below, Harry makes a paraphrastic report similar to those described above including modals. In this example, Harry asks a lamb that is lying in his lap if it would like to have a swim. On his mother's request, Harry reports what the lamb replied. By Harry's report, we may imagine the little lamb originally having "said": "Yes I'd love to do that!".

Example 4.34

Harry35_29.cha

*MOT:	tror du de? att lammet vill bada i havet? du får fråga honom
%eng:	do you think so? that the lamb would like to bathe in the sea? you should ask him
*HAR:	du vill bada i havet ha?
%eng:	you'd like to bathe in the sea huh?
*MOT:	va sa han?
%eng:	what did he say?
***HAR:**	**han ville de** [jfr. dir: *jag vill* det/ja det *vill jag*]
%eng:	he wanted that [cf. dir: *I want* that/yes *I'd* like that]

Yet, this example may also be interpreted as a case of indirect speech without a framing clause (*(han sa att) han ville det*, '(he said that) he wanted that'). Other candidates of free indirect speech by the children are rare, although they sometimes report or project speech by means of direct speech where the deictic elements are "incorrect" (see e.g. Harry's report from when he rode a roller coaster in Example 4.11 on page 162). I regard Example 4.35 below as perhaps the only clear case of free indirect speech in the Longitudinal case studies.

Example 4.35

Harry43_01.cha

***MOT:**	först **fick vi nte gå ut i köket** när vi ville **sa pappa** för han hade väl en hemlis te oss där [jfr. dir: "*ni får* inte gå ut i köket" sa pappa]
%eng:	first we couldn't go out in the kitchen when we wanted to, Daddy said, probably because he had a secret for us out there [cf. dir: "*you can*'t go out to the kitchen", Daddy said]

One explanation for this is the fact that it is followed by a framing clause with a verb of saying.

   Thus, we may conclude that free indirect speech is extremely rare in the Longitudinal case studies, and it typically occurs in contexts of speech reporting (rather than speech projection). The few candidates found refer to Harry's and Tea's mother (rather than their grandmother), and only one utterance by Harry is (possibly) a case of free indirect speech. Taking into

account the age of the children by the time their mother use the forms, we find that the earliest case occurs in the interaction with Harry (as shown in Example 4.31) at 24 months. Harry himself reports what his lamb said (see Example 4.34) when he is just about to turn three. So, what about the three-year-olds playing with the doll house? Are any instances of free indirect speech evident there?

As a matter of fact, cases of free indirect speech are rare in these data as well. I found two cases by two different mothers that met the criteria of free indirect speech. Interestingly, the types of use are very similar to each other. The examples are shown below.

Example 4.36

DOLL.042-03

MOT:	så GÅ han hit åsså koka han koka koka koka
eng:	then he GOES here and then he cooked cook cook cook
MOT:	**NU ä han färdi** [=! glatt, röstförändring]
	[jfr. dir: "nu är *jag* färdig"]
eng:	NOW he's ready [=! in a gay voice]
	[cf. dir: "now *I*'m ready"]

Example 4.37

DOLL.025-03

MOT:	ja // då lagar mamma MATEN här // hon laga å rör i PANNORNA så DÄR å kokade
eng:	yeah // then Mummy cooks the FOOD here // she cooks and stirs in the PANS like THAT and cooked
MOT:	**nu var de klart** [=! höjer rösten, "ropar"]
	[jfr. dir: "nu *är* det klart"]
eng:	now it was finished [=! raises the voice pitch, calling]
	[cf. dir: "now it *is* finished"]

These two examples describe how one or two dolls enter the kitchen in the doll house, and how the father or the mother doll is cooking and announces when the cooking is done. The utterances marked with bold face qualify as speech projections (i.e., indirect voicing, the dolls are made to speak) by means of the

changes in pitch. By changing the tone of voice, the utterance and the "voice" is distinguished from the preceding narrative clauses and the "voice" of the narrator (the real mother). Thus there is a speaker perspective shift. Yet the deictic elements are not adjusted in the way that would be expected if the perspective was simply shifted from the narrator's deictic center to the dolls' deictic centra, and this is also what qualifies them as examples of free indirect speech. If the utterance in Example 4.37 was a direct quote, *var* ('was') should be *är* ('is'), and in Example 4.36 *han* ('he') should be *jag* ('I'). Additionally, there is a certain conflict in temporal deixis *within* the utterance in Example 4.36, in that *nu* ('now') conflicts with *var* ('was'). One side of the coin is to see this as an aspect of "play pragmatics", that is, the *now* of the play is something different from the *now* of the real world (cf. Strömqvist 1984). Another side of the coin refers to what is seen as typical of free indirect speech, that is, the speech is meant to be enacted and uttered within the play (the perspective of the doll) but the perspective of the player (narrator), in a sense, still remains (see e.g. 2.4.1.2 and 2.4.2.3).

The children in the Doll house data do not use the type as described above. I have found only one candidate of free indirect speech in total in the speech of the children. This potential case of free indirect speech occurs in a play dialogue between a mother and her three-year-old daughter. They discuss a telephone conversation that has taken place between the child doll and someone else:

Example 4.38
DOLL.040-03

MOT:	de kanske va MORMOR som ringde å fråga om HON fick komma å grilla KORV
eng:	maybe it was GRANNY who called and asked if SHE could come and grill HOT DOGS
CHI:	**mä de va så LÅNGT hit** [jfr dir: men (mormor sa), "de *är* så långt *dit*"
eng:	but it was so FAR to get here [cf. dir: but (Granny said), it *is* so far to get *there*"]

251

MOT:	VA?
eng:	what?
CHI:	de va så LÅNGT hit
eng:	it was so FAR to get here
MOT:	va re så LÅNGT hit?
eng:	was it so FAR to get here?

It is difficult to tell what status the utterance marked with bold in this example actually has, but one interpretation is that the child reports that the grandmother has expressed that she considered the distance to the house of the dolls too far away to walk. The fact that the mother asks the child to repeat and clarify himself may indicate that it is not obvious what the child intends to express.

Turning to the 3-, 4- and 5-year-olds telling the frog story, this corpus shows the same tendency as in the data presented above, that is, free indirect speech is rarely used. In fact, there is no case found in the adult speech. This is not surprising in the light of the fact that the analyses of the three-year-olds and their parents (see Section 4.2) showed that speech projections were used only twice (two instances of framed direct speech) in the data as a whole. Hence, in this type of context the adults did not typically project speech (although they prompted their children to do so, see Section 4.2). One unclear case of free indirect speech is found in the narrative of a 4-year-old child, CYS04SP. She expresses *åhhh wow dom fick ta en [grodunge]*, 'oh wow they could take one [baby frog home]', in relation to the final picture where one can see the boy holding a frog in his hand. What leads me to believe it might be a speech projection are the initial expressive markers (*oh wow*) and the fact that turning this utterance into a direct quote would result in something like *åhhh wow vi får ta en [grodunge]*, 'oh wow we can have one [baby frog]'. Yet, the utterance can also be regarded as only a spontaneous expression and comment from the child's (the narrator's) own point of view, rather than as a speech projection onto a story character (cf. discussion on speaker perspectivizing in Section 4.3.4).

In the section below, I will explore whether the infrequent use of free indirect speech in the frog story narratives is consistent over development and age, and over the mode of production (speaking versus writing).

4.5.2 School children and adults

When coding for and analyzing free indirect speech in the frog story narratives by school-age children and adults, we encounter similar problems as we did in the coding work of the data of the pre-school children. This means that there is a certain cloudiness involved in determining the status of the utterances and deciding whether they are cases of free indirect speech or not. In the utterances by Harry's mother in Example 4.31 and Example 4.32, the problems of determining the status had to do with the relation to the supposed earlier event and *if* it was actually (and only) speech that was passed on. In the cases of the mothers playing with the doll house in Example 4.36 and Example 4.37, the ambiguity consisted of determining whether the utterances were actually speech *projections* onto the dolls or only comments from the point of view of the play and story organizer (i.e. the mother/narrator). Besides these types of ambiguity, i.e., *speech report or not?*, and *speech projection or narrator's comment?*, the analyses of the frog story result in one further uncertainty: *speech or thoughts projected?*.[1] This third type refers to a vagueness with respect to whether a story character (or someone similar) is intended to be speaking or to be thinking a certain content. It is noteworthy that this type of ambiguity is not salient in the free indirect forms of the pre-school children and the adults interacting with them. This may indicate that this aspect is too complicated for children at this age to produce and process and that the adults avoid applying it, but also that free indirect speech with this "inherent" ambiguity may not be likely to occur and be produced in the types of activities engaged in.

---

[1] Note that these types do not exclude each other (cf. the discussion on double ambiguity in section 2.4.1.2).

I would like to illustrate the ambiguities and how free indirect speech may be interpreted and analyzed in different ways. Figure 4-18 below includes an authentic passage from a 9-year-old telling the frog story orally. The text in italics is from the transcript of her narration, and the speech and thought balloons are added here as an attempt to illustrate how the events may be conceptualized by a reader of/listener to the story. The text in bold is the candidate for free indirect speech.
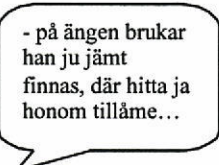
'he [the boy] was thinking...'	'and suddenly he realized!'	'in the meadow, there he [the frog] always was, he even found him there, that was his favorite spot!'
han [pojken] tänkte...	*å plötsligt kom han på det!*	**på ängen där fanns han [grodan] jämt ju, där hitta han honom tillåme, det var hans favoritplats!** (a)

(b) på ängen brukar han ju jämt finnas...

(c) - på ängen brukar han ju jämt finnas, där hitta ja honom tillåme...

Figure 4-18: Ways of interpreting a depicted event: (a) as narrator's voice, (b) as a projected mental act, or (c) as projected speech. The text written in italics is a transcription of a 9-year-old telling about the boy's attempts to find his lost frog.

The idea behind Figure 4-18 above is that this 9-year-old first (as represented by the two first squares) describes and narrates what the boy is doing, i.e., that the boy is thinking and trying to find a solution to his problem, and that an idea then suddenly pops up. Then (as represented by third square) the idea is expressed – the meadow is a place worth exploring since the frog is likely to be there. As a listener of this final utterance, we may imagine the scene in different ways. We may "hear" a narrator's voice, i.e., a narrator providing us with information on what is happening (cf. the texts in square 1 and 2). This is

represented by "(a)" in the figure above. The perspective of the narrator remains. Otherwise, as indicated in "(b)", we can imagine how the boy has an inner monologue and expresses "silent speech". Here the narrator is projecting explicit thoughts onto the boy and the perspective is shifted to him. A third alternative, "(c)", is that the boy is intended by the narrator to burst out in speech. In this interpretation, the perspective is shifted to the boy in the same way as in (b), however, the narrator projects speech to the boy rather than quiet words and inner monologue. The object of this thesis is to investigate the reporting and projection of *speech* rather than thinking. However, as this analysis shows it is, in many cases of free indirect speech, impossible to separate thoughts from speech. Therefore, some of the examples given below are equivocal.

For convenience and for clarity of presentation, I distinguish four partly overlapping sub-types of free indirect speech below: *interrogative clauses*, *stylistic quotes*, *parts of dialogue*, and *framed quotes*. Each of these types has certain characteristics, but the differences should be regarded as gradual rather than representing mutually exclusive categories.

In contrast to the parts of dialogue and the framed quotes, the interrogative clauses and the stylistic quotes are unframed (i.e., they are not preceded or followed by a clause with a speech act verb). They are both examples of the type that was discussed in Section 2.4.1.2 and which has been thoroughly examined by literary researchers for a long period of time. Tokens of free indirect speech are then seen as a stylistic device, used primarily by professional writers to represent the narrator's or the story character's thinking or speaking in a typically ambiguous way. It is then left to the reader to interpret what is expressed and from whose point of view it is expressing (cf. Figure 4-18 above). What distinguishes utterances of the interrogative type from those I have called stylistic, is, of course, the fact that the former are questions:

CAR20WR:	***Kanske hade Kvack hoppat ut genom fönstret?***
	'Maybe Ribbet had jumped out the window?'
ANN20SP:	***oj oj oj va ska dom göra?***
	'oh oh oh what shall they do?'
EMI15SP:	***va kunde re vara?***
	'what could that be?'

Table 4:16: Free indirect speech in the Frog story data, school children and adults, speaking and writing; examples of *interrogative utterances*.

Although the examples provided in this table may indeed be thought of as speech projections, they most likely reflect mental activities by the boy (inner monologues). The narrator can use this type of utterance to move the plot forward. For example, the phrase *Kanske hade Kvack hoppat ut genom fönstret?*, 'Maybe Ribbet had jumped out the window?' [CAR20WR], serves among other things to inform the audience that the boy now is about to look (or actually *is* looking) out the window to search for his frog. In addition, the interrogative utterances can be used by the narrator to comment on the events: *oj oj oj va ska dom göra?*, 'oh oh oh what shall they do?' [ANN20SP]. These utterances seem to have a clearer function of providing the narrator's comments than actually projecting speech to the story characters.

The stylistic quotes (see table below) are also ambiguous, yet, the stylistic quotes in my opinion are more likely to actually represent a character's speech than the interrogative utterances above. Similarly to what I found in the analyses of free indirect speech use by Harry's and Tea's mother (see the preceding section), several of the stylistic quotes in the frog story narratives contain modals such as *skulle* and *kunde* ('should' and 'could'). Common to many of the stylistic quotes is also the inclusion of expressive words as *åh*, ('oh'), and *Gud*, ('God'); exclamation marks (in the written examples), and the discourse particle *ju* (there is no equivalent word for it in English). These contribute to a "speech-like" feeling. In fact, the examples of the 12-year-old [SAB12WR] and the 15-year-old [ANJ15WR] narrators in the table below may

even be regarded as direct speech projections and thus instances of free direct speech. It is true that the fact that past tense is used instead of present tense indicates that we are dealing with free indirect speech. Still, in the contexts where these utterances are expressed, past tense is also straightforward when considering them as examples of direct speech.[2] A reason not to see them as free direct speech, however, is the fact that both narrators make use of direct speech elsewhere in their narratives and then mark them in terms of speech hyphens (see the example of SAB12WR on page 243 above), whereas the utterances given in the table below are not marked.

KAP20WR:	*Inte kunde han och Frog redan skiljas, de hade ju precis träffats!* [lit.] 'Not could he and Froggie separate already, they had just met!'
JOM20WR:	*De hade själva suttit i samma båt och de förstod sin vän så väl. En dag skulle de säkert själva trilla dit.* 'They had been in the same boat themselves and they understood their friend so well. One day they would probably get there themselves.'
ANJ15WR:	*Gud, vad dom var söta.* 'God, they were pretty.'
SAB12WR:	*Nu blev det ju glasskärvor över hela gräsmattan!* 'Now there were bits of glass all over the lawn!'

Table 4:17: Free indirect speech in the Frog story data, school children and adults, speaking and writing; examples of *stylistic quotes*.

Stylistic quotes (as well as interrogative quotes) are primarily used by the adult narrators (for numbers, see Table 4:20 below), and especially in their written narratives. This gives support to the claim made by e.g. Polanyi (1982) and Banfield (1993) that the form typically appears in written discourse (rather than

---

[2] *Var* ('were') in *Gud, vad dom var söta*, ('God, they were pretty'), then refers back to the experience of the moment when the protagonist discovered them (the frogs); and when the protagonist scolds his dog for having broken the glass jar by saying *Nu blev det ju glasskärvor*

spoken). Moreover, the fact that it is more frequent in the adult group than in the younger age groups, indicates that this form is more accessible to trained writers than to writers with less practice.

The third type, part of dialogue, fairly clearly refers to speech since the utterances are included in an exchange of utterances and are replies to (or a continuation of) a preceding speech projection.

JOT20WR:	*den tidigare rymlingen [grodan] frågar sina föräldrar om han inte kan få följa med Pelle och Plutt hem.* **Visst får han det.** 'the previous fugitive [the frog] asks his parents if he couldn't follow Peter and Plutt to their house. Certainly he can.'
EMI15WR:	*Barnen frågade om de ville leka med* **och det ville de.** 'The children asked if they wanted to play too and they wanted that ([lit.] and it wanted they)'

Table 4:18: Free indirect speech in the Frog story data, school children and adults, speaking and writing; examples of *parts of dialogue.*

A common denominator to the five instances of this type that I found in the data is the fact that the utterances are preceded by an indirect speech form. I do not regard them as only a continuation of the indirect speech form due to the word-order and the fact that the free indirect form represents the speech of a new speaker (dialogue exchange).[3] Note also the resemblance between the example of [EMI15WR] and the form produced by Harry at age three when he reports what his lamb said: *han ville det,* 'he wanted it/that' (see Example 4.34 on page 251).

över hela gräsmattan!, ('Now there were bits of glass all over the lawn!'), the boy contends that the jar broke at an earlier point in time and that the result was bits of glass all over the lawn.

[3] Turning these examples into indirect speech would affect the word-order. Cf. the example of [EMI15WR] *och det ville de* ('and it they wanted') that would become *och [de svarade att] de ville det* ('and [they replied that] they wanted it') as indirect speech.

Only two cases in total, produced by two different 15-year-olds, refer to
the final category, *framed quotes*. These are very similar to framed direct
speech, but in order to count as direct speech, the deictic elements in the quote
should be adjusted from the narrator's perspective to the quoted speaker's
perspective. This is not the case in the example shown in Table 4:19 below
(*kunde* ('could') should be in present tense), and therefore it has been regarded
as a case of free indirect speech.

JAN15SP:	*var kunde grodan vara, sa dom* 'where could the frog be, they said'

**Table 4:19: Free indirect speech in the Frog story data, school children
and adults, speaking and writing; example of *framed quote*.**

It is impossible to know whether this was a conscious choice by the narrator, or
if he failed to make the proper deictic adjustment. Since this occurred in the
subject's oral narration, it could also be explained as a "mistake" caused by the
fact that the speech planning is made more or less simultaneously to the speech
production (cf. discussion in 2.2.3).

Summarizing the numbers of the candidates of free indirect speech forms
in the spoken and written frog story narratives by the 9-, 12- and 15-year-olds
and adults, yields the figures as in Table 4:20 below. The table shows that free
indirect speech is relatively rare in the frog story narratives and it is more
frequent in written (a total of 25 cases) than in spoken narratives (10 cases).
Adult narrators use it to a greater extent than younger narrators do. Nine of the
fourteen adults make use of free indirect speech and they use it in the form of
interrogative utterances (that typically reflect mental activities of the story
characters but they also allow the narrator to provide comments on the events
and to move the plot), as a stylistic means, and they integrate the form in
dialogues.

	Inter-rogative utterances	Stylistic quotes	Parts of dialogue	Framed quotes	Total
20WR	9	7	2	-	18
20SP	3	1	2	-	6
15WR	2	1	1	-	4
15SP	1	-	-	-	1
12WR	2	1	-	-	3
12SP	-	-	-	2	2
09WR	-	-	-	-	0
09SP	-	1	-	-	1
Total	17	11	5	2	

Table 4:20: Free indirect speech in the Frog story data, school children and adults, speaking and writing; total number of instances of *interrogative utterances*, *parts of dialogue*, *framed quotes*, and *stylistic quotes*.

The school children rarely make use of free indirect speech. From the analyses in Sections 4.3 and 4.4 we also know that (framed and free) direct speech and indirect speech are infrequent in the spoken and written narratives by many of the 9-year-olds and nearly all of the 12- and 15-year-olds. All these forms can be used in order to make evaluations of the events in the story. A resulting question is if these children and adolescents make use of other means than speech projection to make evaluations?

4.5.3 Mental verbs

One way to include evaluations in the narratives is to depict and describe the story characters' cognitive acts and mental processes. According to theories of understanding of minds, and the development of cognition and narrative abilities (cf. discussion in Sections 2.1 and 2.3), we can expect adolescents to choose and to manage to depict mental acts to a greater extent than their younger counterparts. To examine if this is the case, I coded the data for framing clauses where the verb included was a mental verb rather than a speech act verb and the types of verbs coded for were the following:

*grodan kanske finns där, **tänkte** han*
'maybe the frog is there, he **thought**'

*han **undrar**, var kan grodan vara nånstans?*
'he **wonders**, where can the frog be?'

The verb *undra*, 'wonder', is ambiguous in that it can be used to describe both a speech act and a cognitive act (cf. the inclusion of this verb in Table 4:15 on page 245). In the coding process I relied on contextual clues to determine the status of the individual instances of *undra*. I also coded for four types of verbs denoting mental activities (not having the framing function):

*renen **tänker** (har för avsikt att) rädda pojken*
'the deer **intends** to save the boy'

*han **tänker** på (reflekterar över) sin groda*
'he is **thinking** (reflecting) about his frog'

*dom **undrade** (funderade på) var grodan kunde tagit vägen*
'the **wondered** where the frog could have left'

*han **märkte** (upptäckte) att grodan var borta*
'he **noticed** (discovered) that the frog was gone'

*pojken **tycker** (upplever det som) att han hör någonting*
'the boy **thinks** (experiences) he hears something'

The verbs *märka*, 'notice', and *tycka*, 'think, experience', are perceptual as well as cognitive in the meanings and contexts above. I have not included instances of *tycka* when it refers to 'having an opinion'.

The table on the next page shows the instances of these mental verbs found in the Frog story data, and it is striking to see how few instances of cognitive verbs (two in total) there are in the 3-, 4- and 5-year-old groups.

261

	Mental verbs used as framing verbs		Mental verbs (not framing verbs)				
	tänka 'think'	*undra* 'wonder'	*tänka* 'think; intend'	*undra* 'wonder'	*märka* 'become aware of; notice'	*tycka* 'think; exper-ience'	
20WR	11	1	8	1	12	2	35
20SP	9	3	3	-	5	2	22
15WR	5	-	4	3	10	-	22
15SP	1	5	-	2	7	2	17
12WR	3	-	-	-	8	1	12
12SP	2	-	-	1	1	-	4
09WR	3	-	1	2	3	1	10
09SP	4	-	3	2	1	-	10
05SP	-	-	-	-	1	-	1
04SP	-	-	1	-	-	-	1
03SP	-	-	-	-	-	-	0

Table 4:21: Verbs denoting mental activities in the Frog story data: pre-school children, school children and adults; speaking and writing; absolute numbers.

Studies have shown that children's understanding of the presuppositional and implicational properties of mental verbs are learned around 4-5 years of age (see Aksu-Koç 1994). The semantics of the verbs as well as their place in syntactic structures are learned at this age, along with the emerging awareness about the minds of others. Kavanaugh *et al* (1997) report that at age 4 children ascribe emotions and cognitions to inanimate objects in make-believe play, and at the same age children start to regard the events and actions of a story from the perspective of its protagonists, understanding their beliefs, mental states and intentions (Astington 1990). Therefore, it may be surprising that mental verbs are generally absent in the 5-year-olds here. This probably reflects the fact that the frog story is a complex narrative task where two levels are to be interwoven; the child is expected to construct a coherent text relating a long sequence of events, and in addition make it meaningful from the point of view of the characters experiencing those events. Aksu-Koç (1994), in her analyses

of Turkish children's' frog story narratives, found that the 5-year-olds had difficulties combining these two levels and that the children seemed to construct their stories as an objective series of events. The 9-year-olds, in contrast, managed to include both levels and to attribute subjective experiences to the characters. If considering the mental verbs in the above analysis as a way of attributing cognition to story characters and making them into experiencing subjects, I find the same type of development in my data. Whereas there are hardly any mental verbs in the narratives of the pre-school children, there are between four and twelve verbs in the narratives of the 9- and 12-year-olds and around twenty in the narratives of the 15-year-olds and in the spoken narratives of the adults. The most cases, a total of thirty-five, are found in the written narratives by the adults.

In all groups except the 3-, 4- and 5-year-olds, *tänka* ('think') is used as a framing verb, that is, thoughts are projected. Nevertheless, thoughts (or silent, inner speech) are not projected instead of, or more often than, speech. As a matter of fact, it is true for all age groups that speech is projected to a greater extent than thoughts are (cf. the analyses presented in Sections 4.3 and 4.4). Hence, not only do narrators of the frog story project speech from an earlier point in time (around three years of age), but they also generally project speech more *frequently* than they project thoughts. This is indeed an interesting phenomenon since both types of projections involve ascription of intentions, and a question to ask is why speech seems to be easier or more likely to be ascribed to the protagonists in the story than thinking. Moreover, in accordance with the findings of free indirect speech in the preceding section, I observe that mental verbs are more frequent in writing than in speaking across the age groups (79 verbs in the written narratives and 53 in the spoken). In fact, what is evident is a general trend that cognitive and communicative activities (manifested by the use of direct and indirect forms of speech and cognitive predicates) are expressed to a greater extent in writing than in speaking.

4.5.4   Summary of the results

The analyses of the Longitudinal case studies and the Doll house data yielded meagre results from the point of view of instances of free indirect speech. The few candidates found in the input data to Harry and Tea are found in the speech of the children's mother, and these cases had certain characteristics in common. They were speech reports rather than speech projections, the relationship between the free indirect speech report and the original speech uttered was vague and the reports conveyed a sense of being summaries of an earlier event, and the forms typically contained modal expressions. In the Doll house data, two candidates of free indirect speech by two different mothers were found. The mothers signaled that the dolls were the speakers by holding onto them and by changing their tone of voice. Although these utterances had the word-order characteristics of direct speech, not all deictic elements within them were adjusted to the here and now of the projected speakers. This kind of blending of temporal and personal reference, which can be regarded as a type of play discourse, was not found in the children's own language.

Child uses of free indirect speech among the pre-school children were rare and the status of them unclear (one case by Harry at 3 years, one by a 3-year-old playing with the doll house, and one by a 4-year-old relating the frog story, respectively).

The analyses of the spoken and written frog story narratives by the school children and the adults showed that free indirect speech occurred only occasionally in the 9-, 12-, and 15-year-olds' narratives, whereas the adults sometimes used it, especially in their written narratives. These forms were typically ambiguous in nature and could serve to reflect mental activities by the characters, and/or to represent a narrator's voice, and/or to express speech by story characters. The analysis of occurrences of the mental verbs *tänka*, *undra*, *tycka* and *märka* in the frog story narratives yielded results similar to the analysis of free indirect speech. In other words, mental verbs were extremely rare in the pre-school children's narratives, they were used to a greater extent in

264

the school children's narrations than in their younger peers', but they were most frequent in the adults' written narratives.

5 General discussion

This chapter is divided into two major sections. The first section discusses the findings presented in the previous chapter and attempts to answer the research questions that were stated in the introductory chapter and in Section 2.5. The second section is devoted to a discussion of the methodological framework.

5.1 Answering the research questions

There are four main types of research questions of this thesis that concern the development of direct and indirect speech, the use of direct and indirect speech, input characteristics, and later development in speech and writing (see Sections 1.1.3 and 2.5). The answers to these questions are discussed in three subsections below.

The first sub-section (Section 5.1.1) examines the development of direct and indirect speech in the pre-school children, whereas the subsequent section (Section 5.1.2) deals with the development of these forms in speech and writing by the 9-, 12- 15-year-olds and the adults. The third section (Section 5.1.3) is especially devoted to the issue of input, primarily the adults' use of the forms when interacting with the pre-school children. Questions concerning how the forms are used are considered in all three sub-sections.

5.1.1 The development of the forms in early years

One of the main tasks of this thesis has been to establish the course of development of the forms of direct and indirect speech. Harry and Tea represent the youngest language users in the data of the thesis and their use of direct and indirect speech provides important information on what types of forms are the first to be acquired and which forms that are most accessible to young language learners. The analyses of the Longitudinal case study data showed that mimicking of animal cries in relation to "reading" of books was frequent in a first phase of development (before age 2). Although this is not

267

about speech projection proper, the animals are ascribed sounds by the children and the adults are likely to put the mimicking into a frame of direct speech (cf. *ja så säjer lammen, bä:: säjer lammen*, 'yes that's what the lambs say, baa:: the lambs say' (Harry's mother in Example 4.1 on page 154). Around the children's second birthday, Harry and Tea start to produce their first forms of direct speech. These are typically prompted by the adults and the quotes are free rather than framed.[1] The use of free direct speech is explained by two facts: first, the preceding adult speech projection or reporting prompt itself consists of a framing clause, thus, there is no need for the child to frame the subsequent quote (cf. Example 4.3 on page 155); second, the children's ability to coordinate two clauses at this stage of development is restricted (cf. the MLU-value is between 1-2). Hence, it seems more functional to the children at this stage to express the mere quotes rather than to express clauses that frame quotes. Simply speaking, at this age the children quote and enact rather than frame and narrate.

The first forms by Harry and Tea that are not prompted by an adult, are found around 26 months. In Harry's case this first form appears in a conversation about personal experiences (i.e. speech reporting), and Tea makes use of free direct speech in a context of make-believe play (speech projection). Nevertheless, it is not until the children approach their third birthday that the forms are more often spontaneously used by the children (i.e., the forms are not prompted or contingent on previous adult utterances), and indirect speech appears (although the deictic elements are not adjusted appropriately in all cases) in addition to full-blown forms of framed direct speech.

We can thus summarize the development of forms as free direct speech to be the first form to appear (around 26 months), to be followed by framed direct speech (28-30 months of age), and slightly later (and not as common in use) indirect speech (at 35-36 months). These results are in accordance with Ely & McCabe (1993). They found in a similar study of three English-speaking

---

[1] Note, however, Example 4.4 on page 155 where Harry at age 24 months says: *sa fy*, 'said shame' (i.e., a speech act verb is included).

children that the first forms of direct speech (no distinction was made between free and framed direct speech) appeared between 24 and 30 months of age (see review in Section 2.4.4.1).

When considering the linguistic structures of these forms, this order of development is to be expected. Thus, the simplest form from a structural point of view, free direct speech, may consist of around only one word, whereas framed direct speech requires a combination of at least two clauses. This is true for indirect speech too, but in addition, this form involves subordination of clauses. The ability to subordinate clauses is indeed a later emerging skill in children's langauge development (e.g. Plunkett & Strömqvist 1992). Also from a pragmatic and deictic point of view, this course of development is to be expected. The use of free direct speech requires no shift of deixis and in the case of speech reporting the utterance is more or less "repeated". When direct speech is framed, the perspective is shifted from the reporting speaker to the reported speaker, and the deictic elements need to be shifted accordingly. In indirect speech, the speech of another speaker is reported (or projected), but the perspective remains with, and the deictic elements are adjusted to, the reporting (or projecting) speaker. The form of free indirect speech is even more complex deictically, since the utterance reported has the form of direct speech and claims to convey the words as they were originally uttered whereas the deictic elements are adjusted to the speech situation of the reporting speaker. The structural and pragmatic complexities of this form are reflected by the fact that free indirect speech is extremely uncommon in the langauge use by, not only Harry and Tea, but also in the 3-year-olds playing with the doll house and the 3-, 4- and 5-year-olds relating the frog story.

One achievement in the child's pragmatic development that has consequences for the encoding of perspectives and the use of direct and indirect speech, is the child's emerging awareness that the reported or projected speakers have other perspectives and independent minds. Previous research has shown that children start to attribute intentions to others between 30 and 36 months of age (Poulin-Dubois & Shultz 1988; see Section 2.1.3), Kavanaugh *et*

al (1997) point out that children at this age have toy figures carrying out make-believe actions in play, and Wolf *et al* (1984) found that children began to ascribe intentions to the figures in play with small replicas around the age of 31 months. I claim that ascription of speech to others is a part of the attribution of intentions to others, and the fact that both Harry and Tea increase their use and control over speech reporting and speech projection during the months immediately before they turn 3 supports this. For instance, Tea and her caretakers are often and from an early point in time engaged in make-believe play, but it is not until Tea is 31 months that she herself occasionally projects speech in terms of free direct speech to the dolls (still, in comparison to the adults, she is not doing it frequently at this point). From 35 months and onwards, we find longer episodes of projected dialogue exchanges between the dolls with whom Tea is playing.

In order to report and project speech successfully, the narrator also has to take her interlocutors into consideration. For the speech reporting/projecting child this means to take the perspective of the interlocutor into account and to make use of shared knowledge. Thus, in order for Tea's mother to be able to keep track of the changes of speaker perspective in Tea's make-believe play, and for Harry's grandmother to understand who the reported speakers are in Harry's vivid speech reports, the children need to provide their interlocutors with sufficient information. By age 3 both children do this to some extent. Tea, for instance, changes her tone of voice in order to depict the speech of a certain doll, Harry frames the quotes with framing clauses, and both children weave in direct and indirect speech in narrative frames.

Turning to the three-year-old children in the doll house play, all fourteen children make use of speech projections in their play to a greater or lesser extent. In fact, the children project speech significantly more than their mothers, indicating that the option of projecting speech to dolls is appealing to these young children. The overwhelming majority of the speech projections have the form of free direct speech and they are parts of dialogue exchanges. Typically the child plays the part of either of the dolls and the mother takes the

part of the other, but some children create entire dialogues on their own. There are functional explanations for the fact that free direct speech is used to a greater extent than framed direct and indirect speech in this type of activity. Since the children have concrete objects (dolls) to hold onto, it is clear who the projected speaker is (e.g. by holding onto the dolls, wiggling them, directing the gaze towards them). Thus, there is no need to add a framing clause. Moreover, the projection of speech onto the dolls has consequences for, and may be employed as a strategy for, the advancement of the plot in the play (cf. a child accompanying the doll's walking in the house by projecting *ja ska gå å lägga mig*, 'I'm going to bed' (DOLL010-03)). In the doll house interactions, we also find evidence of packaging of information that depicts speech and leads to successful speaker perspectivizing. Several of the children modify their voices, use prosody and make lexical choices in order to depict speakers and speech registers. Indeed, playing house seems to be a good forum to show and practice not only linguistic skills, but, in addition, pragmatic and sociolinguistic skills.

The finding that free direct speech was the most common form of speech projection to be used by the children playing with the doll house is in accordance with the results of the Longitudinal study of Harry and Tea. Free direct speech thus seems to be very functional in play contexts. When it comes to situations of book reading in the Longitudinal case studies, in contrast, direct quotes were typically found to be framed. Interestingly, this tendency is true for the three-year-olds telling the frog story as well. All but two three-year-olds projected speech, and they often used framed direct speech. In my opinion, there are at least two explanations for this pattern of use. First, framed direct speech has been shown to occur frequently in children's books and fairy tales (see Section 2.4.4.2). In those cases where the children have experience of this type of language use (for instance if they have had stories read to them), the use of framed direct speech indicates an emerging sensitivity of genre by the three-year-olds. Second, the frog story is a picture story, and in contrast to the doll house there are no concrete objects to hold onto. Consequently, in order to

271

adapt to these conditions it is functional to refer to the characters by verbal means (like adding a framing clause). In those cases where the children use free direct speech, they fail to make clear who the projected speaker is to a greater extent than their counterparts playing with the doll house.

The problems of conveying the shifts in perspective evidenced in the 3-year-old frog story narrators, are mainly due to an inability to distinguish the narrator's and/or a character's voice from the child's own perspective. The borders between reality and fiction then become fuzzy, at least for the listener. In the doll house activity, we find a related type of ambiguity. In Section 2.3.4, I made the distinction between direct and indirect voicing (Sawyer 1996). Direct voicing refers to cases when a child "becomes" a character (i.e. role play) whereas indirect voicing refers to when a child enacts a play role through the medium of a toy figure and, for instance, projects speech onto it. The mothers in this study typically stick to indirect voicing whereas some of the children go back and forth between the two play strategies. To an outside observer (and perhaps also to the mothers) it is difficult to follow these changes in role identity. The dolls available in relation to the doll house were meant to reflect the structure of a family and the mothers and the children were instructed to "show what happens on an ordinary day in your family" (see Section 3.2.3). Taking these conditions into account, it is likely that a child sometimes "plays herself", rather than acting via a doll that is only a symbol of her. The fact that these types of ambiguities are found within the children's play and only rarely in the play of the mothers is consistent with Corsaro's (1986) claim that indirect voicing is the most advanced form of make-believe play.

The 4- and 5-year-old narrators of the frog story use about as many speech projections as the 3-year-olds do and they too prefer framed direct speech to other forms (cf. Figure 4-8 on page 210 and Figure 4-10 on page 213). However, the 4- and 5-year-old narrators evidence fewer problems conveying shifts in perspective than the youngest narrators (cf. Figure 4-12 on page 220). This difference between the 3- and 4-year-olds in managing the shifts of

perspective should be related to other research results on the development of children's understanding of minds. As was reviewed in Section 2.1.3, this type of research shows that 4-year-olds (but not 3-year-olds) understand false belief, the distinction between appearance and reality, and that something seen may present different visual experiences if the observer views it from different positions in space (e.g. Forguson & Gopnik 1988, Flavell 1988). Hence, the results I present in this thesis fit nicely into this body of research.

Previous research on children's cognitive development in relation to make-believe contexts also shows that three-year-olds ascribe intentions to make-believe figures, whereas four-year-olds, in addition, ascribe cognition to figures (e.g. Wolf et al 1984, Kavanaugh *et al* 1997). However, the analyses of a selection of mental verbs in the frog story narratives (Section 4.5.3) yielded in total only two occurrences of mental verbs in the narratives by the 4- and 5-year-olds. Speech was thus projected to a greater extent than thinking (cf. *"what can I do", the boy said* and *"what can I do", the boy thought*). One explanation for this fact is probably that the attribution of a mental activity such as thinking to a toy figure or story character requires more of the narrator and her theory of mind, than ascription of speech does (cf. also the discussion above that children ascribe intentions to make-believe figures before they ascribe cognition). In addition, when sticking to what is explicitly displayed in the pictures, we can see (or infer) the boy performing speech acts (e.g. calling for the frog), however, we cannot *see* the boy thinking. Hence, speech projections can then be expected to occur more often than mental projections.

To conclude, the potential ambiguity between projected mental content and speech content is not really a problem in the data of these age groups, since the figures (the characters in the frog story and the dolls in the make-believe play) are neither ascribed cognitions nor projected free indirect speech. Rather, as described above, the problems have to do with marking the borders between fiction and reality clearly, and providing the interlocutor with sufficient background information.

5.1.2 Later development in speech and writing

The pre-school children relating the frog story showed a preference for direct speech over indirect and free indirect speech. Moreover, the pre-school children sometimes failed in the speaker perspectivizing process. The mean percentage of speech projection clauses was around 5% in all three age groups (3-, 4- and 5-year-olds). However, the groups differed from each other in certain respects. Among other things, the youngest children were found to project onomatopoetic sounds to the animals, and in relation to a wide range of depicted events. The 5-year-olds, in contrast, mainly projected verbal content onto the boy and typically in relation to the pictures showing the boy calling for the frog. In this sense, the 5-year-old group was more homogeneous than the younger groups of narrators. Yet, in another sense, the oldest group is heterogeneous since there are greater individual differences here than among the younger children when it comes to how much speech is projected. That is to say, some of the 5-year-olds use the option of projecting speech, whereas others choose alternative narrative strategies that do not include speech projections.

The type of narration that is performed by the majority of the school children (i.e. in the 9-, 12- and 15-year-old groups), has certain characteristics. The narratives include relatively few cases of speech projections and the projections typically have the form of framed direct and indirect speech. Moreover, in most cases the boy is the projected speaker and a majority of the projections are made in relation to the picture depicting the boy asking his dog to be quiet. A few cases of free indirect speech are used and the boy is occasionally ascribed mental activities (more frequently in the 15-year-old group than in the 9- and 12-year-old group). Due to these characteristics, the narratives have a distanced feel. As was discussed in Section 2.2.3, Chafe (1982) introduces the notion of *detachment* as a property typical of written langauge, whereas *involvement* is characteristic of spoken langauge. Yet, despite the fact that the narratives here are oral, they express more of a feeling

of detachment than of involvement. Indeed, elsewhere (Nordqvist 1998b) I have described these narrators as following a speak-as-you-write strategy.

Why is it the case that these narrators produce such "written-like" discourse? I think that one part of the explanation is the fact that written langauge is often associated with books, formal situations and expository prose (cf. Tannen 1985, see Section 2.2.3). This kind of language use is to a great extent learned and practiced in school. It is likely that (some of) the children participating in this study find this narrative task similar to tasks normally carried out in school, and that the situation is experienced as formal. (I will return to methodology and the issue of experimental conditions in Section 5.2 below.) School children's increased experience with written expository texts probably also contributes to the children using more conventional and "written-like" ways of reporting and projecting speech than younger children do (cf. Pontecorvo & Orsolino 1996).[2] This is indicated by the fact that these narrators use indirect speech (which contributes to a distanced quality), but avoid free direct speech (which may cause problems from a speaker perspectivizing point of view, especially in writing).

In contrast to this type of "detached" narration, there are two narratives in the 9-year-old group that are characterized by involvement (to use Chafe's terminology) and where the distance between the narrator and the story characters is minimized. These narratives consist to a great extent of dialogues between the story characters, and instead of using verbal framing clauses to mark the identity of the projected speakers, the narrators modify their voices and use vocatives and onomatopoetic sounds. Typically, they manage to convey who the projected speaker is, although not in all cases. In these narratives, the speech projections serve to move the story ahead. This style of narration has clear affinities to what is found in comic strips. Indeed, the frog story booklet consists of a series of pictures depicting events and one way of

---

[2] Cf. also the discussion in section 2.4.4.2 about the girl Malka who grew up in an environment where the emphasis was placed upon literacy-related activities (Wolf & Hicks 1989). She was found to tell a less lively oral story than her peer Rene, who came from a home where oral story-telling was encouraged.

narrating the story is to simply add "balloons with speech". Importantly, however, the subjects provide a narrative frame by introducing their stories with *det var en gång...* ('once upon a time') and anchoring in past tense (i.e., besides free direct speech utterances we find narrative clauses in past tense). It should also to be taken into account that the frog story is a children's book, and a vivid narration including dialogue exchanges and voice modifications probably attracts little children. This style of narration thus may be explained as a way of adapting to the fact that the story is meant for children younger than the subjects themselves.

Elsewhere in this thesis (Section 2.4.4.1), I reviewed studies of Swedish pre-adolescents and teenagers, and their use of speech reporting in naturalistic peer conversations (e.g. Nordberg 1986, Eriksson 1993). These subjects were found to use a considerable amount of onomatopoeia, pseudoquotations, and free and framed direct speech in their interactions. So, to conclude what has been said so far about school children's use of speech reporting and projections in speaking, we distinguish at least four types of use that school-age children are likely to master. First, in formal renditions (as represented by the majority of the frog story narrators in the 9-, 12- and 15-year-old groups), speech projections are rarely used, and when they are, framed direct and indirect speech dominate. In a second type of use, speech is projected in a comic strip dialogue fashion and free direct speech is common. This type is characterized by less formality, as is the third type of narration, in which speech is projected in a vivid manner in order to appeal to younger children. Reilly's study of 10-11-year-olds telling the frog story to pre-school children supports this point, since the school-age children in her study were found to include affective devices in their narrations in order to adapt it to their pre-school-age listeners (Reilly 1992, reviewed in Section 2.4.4.1). Fourth and finally, in informal peer dialogues, the speech of others is reported by means of direct quotes, onomatopoeia, pseudoquotations and unconventional speech act verbs (e.g.

ba(ra), 'just').[3] Although these four types of registers are not evidenced in the same group of subjects here, the findings suggest that most pre-adolescent and teenage children can switch between the different modes of reporting and projection of speech, and adapt to the situational context, aim, and degree of formality.

Turning to the written data, the analyses show that many of the school-age children producing spoken narratives with a detached quality, use the same type of strategy in their written stories. Hence, speech projections are relatively rare, direct quotes are framed and indirect speech is used. This is true irrespective of the order of production (i.e. speaking before writing, or vice versa). However, this pattern applies more to the 12- and 15-year-olds than it does to the 9-year-olds, since it turns out that speech projections are more frequent in the 9-year-olds' writing than they are in their spoken counterparts. Thus, more 9-year-old writers use the type of narration that was found in two of the 9-year-old oral narrators (cf. Figure 4-15 on page 231). These written narratives include passages of dialogue including framed and free direct speech utterances, which are attributed to a variety of story characters. As a consequence of this strategy, the narratives are characterized by involvement rather than by detachment.

So why is speech depicted to a greater extent in the written accounts than in the spoken ones? Why do these narrators prefer *writing about speech* to *speaking about speech*? One explanation is that the writer has more time to plan and edit her discourse than the speaker has, and this may favor the use of more elaborate langauge (like the inclusion of a dialogue within a narrative). Another explanation is that the oral narrator can make use of voice modifications and gestures along with the verbal message in order to make the story interesting to listen to, whereas the writer needs to find other ways to serve a similar purpose. One way to do that is to "get into" the story characters

---

[3] This style of conversation probably serves social functions such as strengthening peer relationships, achieving group solidarity, and influencing individuals' perception of people and events.

and attribute speech to them. I think the point Tannen (1985) makes that markers of involvement and interaction are not reserved for spoken discourse but are important features of genres as written narratives and creative writing, is important (see Section 2.2.3). Indeed, dialogues of speech frequently occur in novels, and children at this age are likely to encounter this type of language use in their own reading. In addition, Swedish children are encouraged and trained to write their own informal stories early on in school. It is likely that reporting of dialogues is a relatively recently gained skill for these young writers, and also that some writers integrate this type of language use to a greater extent than others do.

Elsewhere (Nordqvist 1998b) I have referred to the 9-year-old writers as following a write-as-you-speak strategy. This strategy has consequences for the intelligibility of changes in perspective within the narrative. Hence, paralinguistic features, such as voice modifications and gestures that function as indicators of speaker identity in speaking but not in writing, are not always taken into account by the 9-year-old writers (this is in accordance with the findings by Michaels & Collins (1984, see Section 2.2.3). Figure 4-17 (page 239) shows that the 9-year-olds, and the 12-year-olds, do not do as well in writing as they do in speaking when it comes to conveying shifts in speaker perspectivizing. The main reason for these shortcomings is the fact that conventional markers of speech, like hyphens and quotation marks, are rarely applied by these writers. These results are in accordance with the studies of Ferreiro & Zucchermaglio (1996) and Ledin (1998) that show that children of these ages often use direct speech in their writing and know that it is distinguished from the surrounding text, yet they do not control how the graphic conventions should be used (see Section 2.4.4.1). In this sense, the art of *speaking within writing* is not yet fully developed.

The adult subjects' frog story narratives show a variety of characteristics that on the one hand distinguish them from the younger subjects' narratives, and on the other hand evidence modality specific features that distinguish the spoken narratives from the written ones.

The adult subjects produce significantly longer narratives, and a greater number of speech projections, than the younger narrators. The quotes are longer than in any other group, the adult narrators project speech onto a variety of animal story characters and these projections consist of verbal content rather than of animal cries, they project thinking onto the protagonists, and they use free indirect speech. It is not surprising to find these aspects in the adult narratives since "mature narrators ... manipulate perspectives, introduce narrator's comments, ... [and] speculate about cognitive or emotional states of protagonists" (Aksu-Koç 1994: 381). Interestingly, the written narratives contain more of these aspects than the spoken narratives do. Consequently, the written narratives are longer than the spoken ones, they include significantly more speech projections than the spoken stories, more story characters get to say something (hence, the speaker perspective is shifted more frequently), and speech is projected in relation to a greater range of depicted events in the pictures. Moreover, in the cases of framed direct speech, the framing clauses either follow or interrupt the quote (whereas they precede the quote in speaking), and a considerable number of different types of verbs of saying are used (33 types in writing in comparison to 13 in speaking). In addition to this, free indirect speech and mental verbs are used to a greater extent in the adults' writing than in their speaking.

To summarize the findings of the adults' narratives, the written narratives are more elaborate than the spoken counterparts, and the adult narrators adjust to the fact that different conditions hold for the use of speech in contrast to writing (i.e. the adults neither use a speak-as-you-write strategy, nor do they write-as-they-speak). The number of varied speech act verbs illustrates this latter aspect. Hence, since the voice cannot be used in writing to depict a certain speaker and the quoted speech, the speech act verb can serve a similar purpose and describe the speech. In fact, in the 15-year-old group we also evidence a weak tendency towards more types of speech act verbs, and slightly more free indirect speech and mental predicates in writing than in speaking.

279

Speaker perspectivizing in relation to the use of direct and indirect speech is not a problem for the adult narrators, since they frame the quotes and embed the quotes in larger contexts of discourse, they mark speech quotes conventionally in writing, and they modify their voices in speaking. Nevertheless, the use of free indirect speech sometimes causes problems in this respect. In comparison, then, to the conclusion of the preceding section (5.1.1), the oldest frog story narrators have little difficulty making the borders between fiction and reality clear, and the interlocutor is provided with sufficient background information in order to make the story comprehensible. However, they (more or less consciously) leave the listener or reader of the story in a state of uncertainty about whether an utterance is meant to refer to a narrator's voice or to a character's, or to thinking or speaking. The mature narrators indeed know the arts of *speaking within speaking* and *speaking within writing* and they manipulate the options they have at hand.

5.1.3 Using the forms with children

The narrative strategies of the adults relating the frog story – lively narrations with dialogues taking place between animal characters speaking like humans – are probably to a great extent chosen to appeal to little children. The fact that the adults' narrations are, in some respects, more similar to the pre-school children's than the older children's narrations, supports this claim. Another type of narrative situation (e.g. telling the story to an adult or using a different kind of elicitation material) might have yielded another style of narration by the adult subjects. Thus, with these data we can only speculate about the adults' adaptation to a potential child listener (this issue will be discussed further in the next section). In relation to the Longitudinal case study material, the Doll house data, and the pre-school children telling the frog story, in contrast, we can discuss the adult's interaction styles.

In the analyses of the forms of direct and indirect speech, I find certain patterns that could be explained with reference to the type of activity engaged

in. Harry's and Tea's mother and grandmother show a general preference for framed direct speech. It is the most frequently employed form and it occurs in several types of activities. Free direct speech, in contrast, is used almost exclusively in make-believe contexts (especially in the interactions with Tea), whereas indirect speech is only used in relation to book reading and discussions related to previous experiences and/or hypothetical events. From the review of the results of the frog story above, we know that framed direct speech is the most commonly used form in the narratives, and to this we add the fact that the most frequently employed form by the mothers in the doll house play, is free direct speech. Figure 5-1 below summarizes these findings. The figure combines the adult data from the Longitudinal case studies, with the adult data of the Doll house corpus and the Frog story corpus, and shows the distributions of the forms in the three major types of activities as described above.
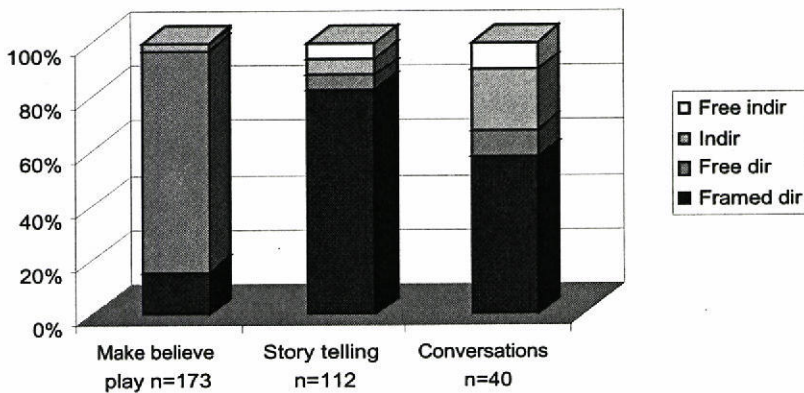
Figure 5-1: Percentages of adults' use of framed direct speech, free direct speech, indirect speech and free indirect speech in three types of activities. *Make believe play* **(play with toy figures) includes adult data from the Longitudinal case studies and the Doll house corpus, (N=16);** *Story-telling* **(reading and/or discussion of books, telling of stories) includes Longitudinal case study data and data from the Frog story corpus (N=58); and** *Conversations* **(discussions about past events, habits of the child and hypothetical (future) events) refer to the adults' use in the Longitudinal case studies (N=2).**

This figure also shows that free indirect speech is infrequent in all activities studied here, especially in make-believe contexts (0,1% of the forms). The cases found in story-telling (5,4%), refer only to adult subjects' monological narrations of the frog story, consequently, free indirect speech is never used by the adults interacting with the pre-school children narrating the frog story. Four out of forty speech reports (10%) in the category of conversations are of the free indirect type.

One of few studies correlating parents' and children's use of direct and indirect speech (Ely *et al* 1995, see Section 2.4.4.2), showed a positive correlation between the extent to which the mothers (but not the fathers) used speech reporting and the extent to which the children reported speech, but no correlation was found between the types of forms used (the mothers typically used indirect speech whereas the children produced direct forms). This suggests that the types of forms heard in the input do not necessarily predict the types of forms used by the children themselves. However, in the Longitudinal case studies presented here, it is evident that Tea uses free direct speech in her make-believe play (exactly as her mother and her grandmother do), the three-year-olds playing with the doll house prefer free direct speech (as do their mothers), Harry and Tea – as soon as their linguistic ability allows it – produce framed direct speech in book reading situations and when discussing personal experiences (exactly as their interlocutors do), and the three-year-olds relating the frog story often frame their quotes (as we have noted that adult narrators also do). Similarly, the extremely infrequent use of free indirect speech by the children can be explained in terms of cognitive and deictic complexity (cf. discussion in Section 5.1.1 above) but also by the fact that it is almost non-existent in the input. Moreover, it is true that the occurrences of indirect speech in the children's and the adults' speech in the Longitudinal case study data are too few from which to draw any conclusions, but the fact that Harry hears a little more indirect speech in his input may explain why he himself makes use of the form more often than his sister. These facts taken together indicate that, in contrast to the results of Ely *et al*, there *is* a relationship between types of

forms used by the adult interlocutors and the children's own use, and in addition, that children make inferences from the activities in which they are engaged and become socialized into activity-related uses.

This reasoning is in line with the claim made elsewhere in this thesis that the models of language (in this case, forms of direct and indirect speech) that the environment provides have consequences for the langauge acquiring child (cf. Section 2.1.2). I have also argued that the caretakers' willingness to interact with the child contributes to the child's linguistic, communicative, and pragmatic development. One way to engage the child in communication is to prompt the child to give a reply, for instance by prompting the child to report or project speech. The analyses of prompts in the Longitudinal case studies showed that these types of prompts were not as frequent as were models of direct and indirect speech, and Harry was found to be prompted to quote speech in relation to discussions of personal, habitual or hypothetical nature, whereas Tea typically received prompts in her input in make-believe play contexts and book reading. Moreover, the adults telling the frog story with their 3-year-old children made more speech projection prompts than did the mothers playing with their 3-year-olds in the doll house play. Considering these results, we need ask what functions these prompts actually have and why the parents make them.

The frog story task demands quite a bit of the child since it consists of a long series of pictures (24 pictures) showing events that are to be interpreted and related to each other. A 3-year-old's attention and memory span is limited, and in order to carrying out the frog story task some of the narrating children need to be supported and be put on the right track by the parents. One way to focus the children's attention on the story and its components, is to make speech projection prompts. The prompt is a question and a request for information, and the child is supposed to provide this information. In most cases, the child attempts to do this. Thus, by this procedure the parent ensures interaction (i.e. a communicative function), and simultaneously, the attention of the child can be directed to components that are relevant to the plotline (i.e. a

pedagogical function). The analysis shows that the most frequent type of prompt in the material is *what does the boy say there?* (or similarly) in connection to the pictures depicting the boy calling for his lost frog. The linguistic encodings of the boy's calling for the frog, contribute to maintaining the plotline (which is the search for the frog). The fact that the 4-year-olds do not need as much adult scaffolding as the 3-year-olds, and the 5-year-olds need even less (cf. figures given in Table 3:2 on page 122, and discussion in Section 4.3.1) points to a developing narrative awareness and a greater independence.

There are several differences between the frog story task and the doll house play which explain why explicit speech projection prompting is not equally common in both activities. First, the doll house play is probably an easier (and maybe a more appealing) task for the child to accomplish than the frog story, and therefore the child does not need as much adult scaffolding. Second, the frog story has a clear beginning, a series of events representing the middle, and an end or a conclusion to the story. This is more or less pre-defined by the pictures and the child (and the adult) should follow this plot. Certainly, a play narrative can also have a beginning, a middle and an end, and there is of course a play language and rules of play (play pragmatics) to be acquired. Nevertheless, the events in (or the plot of) the doll house play, in principle can take any turn and there is no specific pre-defined plot to stick to (thus the mothers do not need to put an effort into directing the child's attention to certain plot components). Third, prompts can indeed be of other kinds and be more subtle than explicit questions like *what does the boy say there?*. The doll house data include many projected dialogues (exchanges of utterances) between dolls, and a contribution by one participant/speaker in a dialogue is likely to trigger a contribution of another participant/speaker (cf. question-answer sequences). In this sense, make-believe play with dolls is a good arena for practicing pragmatic skills such as turntaking.

Speech projection prompts (and related strategies) can thus be seen as having both communicative and pedagogical functions. This discussion of prompts as a strategy by the adults implies that the adult is the one directing the

conversation and the child is only to follow her interaction partner. This is, of course, not always the case. Rather, adults very much adapt to the children. The interests of the children and what they want to talk about often govern the conversations and the activities engaged in, and adults adjust their langauge to a level that is processable by the children (Section 2.1.2). In the Longitudinal case study material, the caretakers were even instructed to let the children take the lead as much as possible during play and other activities in order to avoid triggering a certain behavior (Richthoff 2000: 7). Still, this did not result in passive behavior by the adults. Indeed, adults taking no initiatives at all would neither be natural in these types of interactions with little children, nor would they be particularly desirable. As was described in Section 4.1 both children (Harry and Tea) were prompted to produce direct and indirect speech, and this *is* actually an example of "triggering of a certain behavior". It was also shown that Harry and Tea were prompted to quote speech in different contexts, and this may partly explain why they become skilful in these different domains as they grow older. I think the figures showing the first instances of an adult produced form, a prompt, and the children's first unprompted use of a form in four types of conversational contexts (Figure 4-2 on page 176) quite nicely show the dynamics between input and production and the adults' adjustments to the development of the child. This is particularly clear in the case of Tea, where adult models of direct and indirect speech appear first, to be followed later by prompts, and finally, Tea herself makes use of the forms.

In this section I have primarily been concerned with the interaction between the pre-school children and their adult interlocutors in this particular corpus, but of course "input" is about much more than parent's speech. In the previous section I discussed examples of other types of input that may influence the langauge use and the development of the children; these examples included practice in school and the increasing experience of written discourse of various types. Indeed, the children included in this study are part of a society that, to a great extent, relies on and emphasize literacy, and above I showed examples of how written-like features could "spill over" and be integrated into

285

the school children's oral narratives. However, although the analyses showed that the two modalities were not always well-defined, I also found examples where this distinction was clearly made. Hence, children at all ages, in those cases where they used framed direct speech, tended to place the framing clause before the quote in speaking, whereas they typically placed the framing clause within or after the quote when they were writing. This pattern of the written narratives should be related to the findings of Perera (1996) who showed that the framing clauses rarely precede the quotes in children's early reading books.[4] It thus seems to be a feature typical of written language that is taken up by relatively young language users.

As far as the topic of this thesis is concerned – direct and indirect speech – the child thus obtains knowledge about what types of forms are available in order to project or report speech in her interaction with the environment, and in addition, the child is socialized into different types of uses. At the same time, the child is also driven by what is actually functional in a certain situation. For example, in playing with toys it is typically superfluous to add framing clauses to the speech projections since it is obvious to the child herself who is speaking. In this sense, the child's langauge use is not determined by what she is *told* to use, but what she herself finds *practical* to use.

5.2 Methodological considerations

In the introduction to Chapter 3, I stressed the strengths of the approach of this thesis in that the analyses build on several types of data, broad age ranges are covered and different activities and modalities occur. One main objective of this thesis has been to create a methodological framework that contributes to an

---

[4] For an attempt to explain from a functional and psycholinguistic point of view, why a quote is likely to be preceded by the framing clause in spoken discourse, and followed by it in writing, see section 4.4.3.

extension of the body of knowledge of long-term language development, and of the uses of direct and indirect speech in narratives in particular, in the ways I have described in the preceding sections.

The design mainly relies on cross-sectional data, i.e., the subjects are not followed longitudinally (apart from the Longitudinal case studies of Harry and Tea). However, considering that the purpose of this thesis was to study long-term development from infancy through adolescence, the application of a longitudinal approach would have been far too time-consuming. In the analyses, I have complemented quantitative with qualitative analyses. The size of the entire corpus is large (approximately 275,000 words), and the size allows for the possibility of quantitative calculations. Still, the number of subjects in each age group is relatively small, and the qualitative analyses have added important information on individual strategies and described different types of uses.

The Frog story data serve as the basis of the claims I make about long-term and later language development in this thesis. In Section 3.1, I argued that the use of the frog story in order to elicit narratives has proven to be a successful method with which to study narrative development (e.g. Berman & Slobin 1994). In addition, I stated that the examination of long-term development with experimentally elicited data requires a type of task that works for all age groups. Below I will discuss the possibility of using the frog story in that kind of elicitation task. Importantly, previous approaches using this material have not included adolescents (nor have they included written elicitations).[1]

As far as the youngest children included in the study are concerned, I pointed out already in Section 5.1.3 above that narrating the frog story is a demanding task for such a young a child to accomplish, and this also explains why the parents are likely to intervene and support the child. This is in

---

[1] For other studies considering the Swedish spoken and written Frog story data, see Strömqvist & Hellstrand (1994); Strömqvist (1996); Strömqvist & Ahlsén (1998); Nordqvist (1998b, 1999); Wengelin (*forthcoming*); Strömqvist, Nordqvist & Wengelin (*forthcoming*).

accordance with Berman & Slobin (1994) who observed that the 3-year-olds included in their study found the story too long and complicated and did not manage to tell a story on their own. Their aim was to collect accounts that were produced by the children without support from the interlocutor and they therefore had to reject several three-year-olds' accounts: "In fact it took some 60 sessions with 60 different 3-year-old children to obtain 12 usable stories" (Tanya Renner, quoted in Berman & Slobin 1994: 59). The stories left out were either too short, unintelligible, or partly prompted by the adults. I question how representative the "usable stories" actually are for three-year-olds, and consequently, no such procedure of elimination was applied in this thesis.

The frog story narrations of the youngest children were tape-recorded by their parents at home. What this procedure gains in naturalness (i.e., the story is told collaboratively in a home-setting), it may lose in comparability. For example, we have no control over what may have preceded the actual recording (e.g. whether the story was discussed before turning on the tape-recorder), and there are different parental interaction styles that may have an impact on the children's output. It is true that this makes it difficult to compare the children, but, on the other hand, it then becomes possible to examine the behavior of the parents (e.g. prompting for speech projections). I claim that the parental input is most valuable to explore in order to get a picture of how children learn to use language in interaction with the environment.

Aksu-Koç & von Stutterheim (1994) point out that the frog story is ambiguous in that it yields accounts that are a hybrid of description and narration.[2] Typically, the pre-school children of this thesis described what they saw in the pictures (cf. Section 4.3), but I found description-like accounts in all age groups, not least in the 12- and 15-year-old groups. Many of these narrators used present tense and started out their narration with phrases such as *det här handlar om...*, 'this is about...'. Moreover, they used indirect speech which

---

[2] They refer to two main strategies, the first in which the static pictures are transformed into a continuous story and the pictures are no more than clues to an advancing plotline, whereas the other strategy consists of producing a number of descriptions simply following the picture frames.

Clark & Gerrig (1990) point out is about describing speech rather than depicting it (see 2.4.3.1), and they projected speech mainly in relation to events where the boy is clearly depicted as speaking (or vocalizing).

In the Berman & Slobin (1994) study, all subjects were to tell the story to an adult listener who already knew the story and was able to see the pictures. This was true for the pre-school children in my study as well. However, the 9-, 12-, 15-year-olds and the adult subjects did not receive any specific instructions about the intended receiver and told the story monologically in front of a video-camera (for details about the procedure, see Section 3.2.2). The question is how these conditions might have influenced the subjects' way of relating the frog story. Interestingly, Bamberg (1994) found that the adult narrators who were to tell the frog story to another adult sitting next to them, did not necessarily see this listener as the real receiver. Interviewing these narrators after having told the frog story, Bamberg found that the narrators could be placed into two distinct groups according to the imagined audience; one group of narrators imagined a child-like audience for the narrative (despite the presence of an adult), and the other presented the narrative to the actual adult interviewer. Bamberg further found that this difference in "audience construction" correlated strongly with increased story length and increased use of evaluative devices for the narrators who imagined a child audience group, and a decrease in both measures in the group that simply told the story to the adult interviewer. Unfortunately, we conducted no such interviews with the subjects after the recordings in our study in order to find out what type of audience the subjects imagined. However, we can suspect that many of the adults told the story "to a child" since these narratives include many evaluative devices.

In order to control for the factor of the identity of the listener, I conducted a study (Nordqvist 1998d) where I gave fifteen 15-year-olds the same task as the 15-year-olds of this thesis, but the conditions were changed so that the subjects told the story to a well-known peer sitting beside them. In this procedure, the narrators had a clear receiver, they knew that this person had never seen or heard the story before, and the listener could give feedback to the

289

narrator. The results of this study were strikingly similar to the ones of this thesis, i.e., speech projections were extremely rare, and they were typically of the indirect type. The fact that the accounts did not differ between the two types of condition – the first being a monological setting and the second involving a physically present interlocutor – indicates that narrating is in some sense "monological" and contributions from another speaker are not required to the same extent as, for instance, in a conversation (Nippold 1998). We do not know, however, whether the adolescents would have told the story differently if they were instructed to tell the story to a younger child. Reilly (1992, see also Section 2.4.4.1) examined how 7-8-year-olds told the frog story to an adult and to a three-year-old listener. The different receivers turned out to have little impact on the narrations since the 7-8-year-olds behaved similarly in the two conditions (linguistically as well as paralinguistically). In the same study, Reilly found that the 10-11-year-olds telling the story to a 3-year-old used a considerable number of prosodic affective devices and did so more than the 7-8-year-old narrators. However, there were no differences in how much speech quoting was included in the different conditions and ages. This suggests that the number of direct and indirect speech utterances does not necessarily increase when telling the frog story to a pre-school child.

The frog story is a typical children's book, and motivating adolescents to give a captivating rendition of this story is quite a challenge. However, despite the fact that it may be problematic to find the "perfect" picture series that appeals to all, I consider the method of using a picture story as an elicitation instrument (such as the frog story) to be fruitful in examining aspects of long-term narrative development. This applies to the development in writing as well as in speaking.

6   Conclusion

Children start to produce forms of direct and indirect speech at an early age. The uses of the forms, however, continue to develop, and this is a truly long-term process.

6.1   The development of direct and indirect speech as a long-term process

The reporting and projection of speech are about more than only selecting and producing an appropriate linguistic form (e.g. direct or indirect speech). As Berman & Slobin (1994, see also Section 2.1) discuss, language and narrative development include cognition (an ability to conceive a full range of encodable perspectives), communicative abilities (i.e., assessing the interlocutor's viewpoint), and an ability to express the linguistic forms. Thus, language development is as much about pragmatic development as it is about linguistic development in a narrow sense. As regards the development of uses of direct and indirect speech, the child needs to have an understanding of the mind and perspective of the animate or inanimate object whose speech is reported, or onto whom the speech is projected. In addition, the child needs to consider the perspective of, and the shared knowledge with the interlocutor and provide sufficient information about the speech reported or projected. This means that the child should relate her own perspective, on the one hand, to the perspective of the reported/projected speaker; and on the other hand, to the point of view of the interlocutor. These perspectives are then to be linguistically encoded and further adjusted to the situation in which the child is participating.

The analyses of this thesis have shown that direct and indirect speech can be produced by children at an early age. To be more specific, the analyses of the two Longitudinal case studies (i.e., Harry and Tea) showed that free and framed direct speech appeared before age three and indirect speech only slightly later. By age three Harry and Tea used direct and indirect speech

without being prompted by adults, they enframed the speech to some extent (e.g., by framing clauses or as a part of a narrative sequence), and they used non-verbal strategies to signal whose speech was quoted (e.g., Tea modifying her voice to depict the speech of a certain doll). Typically, the children used framed direct speech in story-telling activities and in the narrating of previous events, whereas they used free direct speech in make-believe play with toy dolls. Exactly the same tendency of choices of forms was true for the 3-year-olds narrating the frog story (i.e., preference for framed direct speech), and the children playing with the doll house (i.e., greater use of free direct speech). Hence, we notice different types of use in relation to different types of activities which indicate pragmatic awareness already at this age. Still, these children sometimes failed to assess the interlocutor's point of view and the shared knowledge, resulting in ambiguity about the reported/projected speech event, and in who the reported/projected speakers were. The 4- and 5-year-olds narrating the frog story clearly managed to convey shifts of perspective better than the 3-year-old narrators did. However, the 4- and 5-year-olds did not manage as well as the 12-, 15-year-olds and the adult narrators did (cf. Figure 4-12 on page 220). These findings suggest that a complete consideration of the listener's viewpoint takes time to develop.

Children as young as 3 years show awareness about activity-related use and school children continue to develop such skills, for instance, by distinguishing formal and informal ways of using language. We find this type of evidence in that many of the school children in this study chose a formal and detached way of narrating, while other researchers (e.g. Nordberg 1986) have shown that children at the same age use langauge in a very lively manner in more informal peer dialogues and personal narratives. As regards the production in speech versus writing, the school children in this thesis showed similar strategies in both modes of production. This resulted, among other things, in speech projections not always being adjusted to the written medium by the 9- and 12-year-olds, and as a consequence, the speech projections were difficult to code. In other words, children at these ages did better when

speaking-about-speech than they did when *writing-about-speech*. The group that most clearly differentiated between speaking and writing and applied modality-specific strategies, was the adult narrators. Thus, the skills connected to spoken and written performance develop in later adolescence.

As concerns narrative development in general, Nippold (1998: 178) concludes that younger children primarily concentrate on story structure (production, embedding and cohesion of episodes), older children get inside the characters and express evaluation (i.e., emotions, thoughts and plans of the characters), whereas the most mature narrators additionally put an effort into the performance (i.e., to entertain and engage the listener). By and large, this description applies to the narrative data of this thesis as well. From the analyses of the overall plotline in the frog story narratives (see Section 4.3.1), we know that most 5-year-olds and school children created a plotline. The analyses of use of free indirect speech and of occurrences of a selection of mental verbs in the frog stories (Section 4.5.3) showed that these devices were nearly non-existent in the narrations by the pre-school children, rare in the 9- and 12-year-olds' narratives, but more frequent in the 15-year-olds' and especially frequent in the adults' narratives. Moreover, the fact that the adults produced longer, more elaborate and vivid narratives (cf. Example 4.25 on page 237), than did most of the 9-, 12-, and 15-year-olds, supports the claim that efforts to engage the listener (or reader) is a skill acquired at a later stage.

This thesis has shown that language development, (in this case the development of direct and indirect speech), consists of much more than only an acquisition of lexicon and knowledge about how the words are combined. Importantly, in addition to these skills, language development is about making *use* of these linguistic components for different purposes and in order to create certain effects. More specifically, I have argued that various kinds of information can be packaged into a linguistic construction such as direct or indirect speech. For example, on a lexical level, the speaker needs to consider and adjust the deictical elements (i.e., words) according to the situation of language use. In addition, different types of speech act verbs may be selected

(cf. the finding that the adult writers used a great variety of different types of verbs to describe speech), as well as particular words depicting a certain speaker (cf. the 3-year-old using *verkligen*, 'indeed', to depict the speech of an older woman in Section 4.2.3). Syntactically, the clauses need to be related to each other and combined. In the case of indirect speech, for example, we are dealing with subordination. Regarding framed direct speech, I found that the positioning of the framing clause varied to some extent between speech and writing (see 4.4.3). Information may also be distributed over even larger fragments of discourse, as in the example of an adult narrator in Example 4.29 where the identity of the quoted speaker and his speech can be inferred from the exchange of quotes in the dialogue. A speaker additionally can make use of aspects such as prosody, voice quality, pitch, phonological stress, and gestures; and the writer may use orthography to represent aspects of delivery. Hence, there are many ways of making use of direct and indirect speech in order to express, e.g., evaluation, plot advancement, speaker perspective, and degree of involvement.

In conclusion, the analyses of this thesis have been aimed at describing children's and adolescents' development in learning how information can be packaged into direct and indirect speech and how these forms can be used for different purposes. Besides cognitive and linguistic maturity, children's interaction with adults and the adults' language behavior, the children's amount of practice and experience of langauge use in different types of activities, their exposure to storytelling in different contexts and modalities, the functionality, and the conditions of production (speech and writing), are examples of factors I have suggested as having an impact on the development of forms and functions of direct and indirect speech.

6.2 Future research

The review of earlier studies on the development of direct and indirect speech in Section 2.4.4, showed that there are certain research gaps within this area. In

the present study I have filled out some of these gaps. Consequently, I have conducted an investigation of the development of these forms by Swedish-speaking children, adolescents, and adults. In the review of previous studies I established that there was a predilection for examinations of children's reports of actual past speech events. In this thesis, the majority of the data have been concerned with speech quoting within frames of fiction, i.e., when speech is projected onto toy figures and/or story characters. In addition, my analyses cover the development and use of direct and indirect speech in writing. The results of this study can be used as a foundation for further research and there are several types of studies that could be conducted in order to provide a more complete picture of later stages of language development and the use of speech reporting and projections.

One major result of these studies was the finding that already at the age of three, children used different forms in different contexts (cf. the use of free direct speech in make-believe play and framed direct speech in story-telling). Further research needs to be conducted in order to reveal the relationships between the uses of the forms and the type of activity, and in connection with investigations of children's acquisition of direct and indirect speech, these relationships should be taken into account. [1] My analyses additionally included free indirect speech, and this form turned out to be less frequent than framed and free direct speech and indirect speech. The form was extremely rare among the pre-school children, and the analyses of this thesis suggest that children do not encounter this form in their spoken language input to the same extent as the other forms. Taking into account the fact that free indirect speech is often found in adult novels, and that the analyses of adult narratives showed that free indirect speech occurred more frequently in writing than in speech, it would be

---

[1] Previous research suggests that it is indeed the case that the use of certain types of forms is related to the type of activity in which the individual is engaged. For example, Ely *et al* (1995) found indirect speech to be the most commonly used form by mothers in dinner time conversations, Perrin (1999) showed that direct speech occurred at six times the rate of indirect speech when adults reported on their life and interests, Hickmann (1993) found adults choosing between framed direct or indirect speech when reporting a puppet dialogue, and Tannen (1986)

interesting to examine whether it occurs in books for children.[2] Moreover, further research and additional types of data are required to decide whether this form *is* primarily a stylistic literary device to be used by trained narrators, or if it occurs in spontaneous casual spoken language as well (cf. discussion in Section 2.4.1.2).

A finding of this thesis that suggests that certain kinds of pragmatic adjustments to the interlocutor continue to develop during the school-years, was the fact that pre-school children and 9-year-olds did not, in all cases, succeed in linguistically encoding and making clear shifts in speaker perspective whereas the 12-, and 15-year-olds succeeded in this respect. In addition to this, I found that the children had greater difficulties marking off quoting in their writing than they had in their speaking. It is likely to be the case that a child who includes dialogue in her writing, makes use of her experience of spoken language. Indeed, the analyses showed that the school children's spoken and written productions had certain features in common and that they affected each other. These findings indicate a later development of pragmatic and modality-specific awareness that should be further investigated. This could be done by examining the use of direct and indirect speech by children and adolescents in other types of data and contexts than those used in this thesis, and in addition, in other types of language use than speech reporting/projection.

The frog story has been used to elicit narratives from a wide range of different languages (see e.g. Berman & Slobin 1994; Strömqvist & Verhoeven *forthcoming*), and consequently, there are great opportunities for replicating the analyses I have done. In what ways do, for instance, typological characteristics of a certain language influence the use of direct and indirect speech in the frog stories (cf. the finding by Özyürek (1996) that Turkish speaking children acquire indirect speech later than English-speaking children)? To what extent

---

showed that 26% of the speech quotations in personal narratives were unframed (i.e. free direct speech).

[2] This area of research would benefit from a cross-disciplinary approach, including, e.g., linguistics, literary science, psychology, and education.

do narrative and cultural traditions, and the role of literacy and oral traditions, influence the inclusion of dialogues in the narrations (cf. discussion in Section 2.4.4.2)? Gayraud *et al* (2000) point out that spoken and written French differ in many respects, and they even describe the situation as one of diglossia. For such a language, it would be interesting to examine and compare the use of speech projection and reporting in speech and writing. Most frog stories collected up to this point are orally produced. However, my experiences of using ScriptLog (Strömqvist & Malmsten 1998) in order to elicit the written frog stories are positive, and I encourage elicitations of written narratives in languages other than Swedish as well. Elsewhere in this thesis, I found that direct and indirect speech were generally used to a greater extent in writing than in speaking. To what extent is this true for other languages, and for genres others than the narrative analyzed in this thesis? Moreover, from a psycholinguistic point of view, and taking into account the rich possibilities of analysis that ScriptLog provides, it would be interesting to investigate pause and editing patterns in relation to production of speech reports/projections, and compare these to production in speaking.

Finally, this study has dealt with children with assumedly normal development in writing skills and further research is needed in order to establish how functional disabilities affect speech reporting and projection. A study of a group of adults with severe reading and writing difficulties and their use of speech projections in written frog stories, showed that they, similarly to the 15-year-olds in this thesis, rarely projected speech and used a small number of different speech act verbs (Nordqvist 1998a). In addition, these writers produced forms that were mixtures of direct and indirect speech of a kind that did not occur in the adult control group. In the same study, I found that nine, out of a group of ten congenitally deaf writers, projected speech, and they more often used direct speech than indirect speech. This group of subjects had (written) Swedish as their second language, and despite the fact that they did not hear themselves or use spoken language, they quoted speech in their writing. Systematic comparisons between the ways of quoting in sign

language,[3] in speaking, and in writing would provide insights into the possibilities and the limitations of a particular modality of expression.

---

[3] See Emmorey & Reilly 1998.

SUMMARY OF THE THESIS

The purpose of the present study is to examine the emergence and development of forms of direct and indirect speech, and functions connected to them, from a long-term perspective. More specifically, the research questions explored concern at what point in time children start to use the forms of direct and indirect speech and what the course of development looks like, how the forms are used by children and adults in different activities and in speech in comparison to writing, and what types of direct and indirect speech forms are used by adults interacting with young children (for a more detailed description of the research questions, see Section 1.1.3, page 5, and Section 2.5, page 97). The empirical data that are investigated include Swedish-speaking monolingual children who have just started to produce their first words, pre-school children telling a narrative and playing with a doll house, and school children and adults narrating a story in speech as well as in writing.

Chapter 1, *Introduction*, presents the aim and scope of the study. In Chapter 2, *Theory*, the theoretical framework of the thesis is introduced and the chapter includes examinations of relevant areas of linguistic research. In Section 2.1, I establish that language development is a long-term, open-ended and a non-linear process. The fact that the language-acquiring child is a part of a social context is emphasized, as is the fact that the child's language development is intimately connected with her cognitive development. In this chapter a discussion about functional similarities and differences between two modes of language production (speaking versus writing) is also included. A brief review of research on children's literacy development and emerging awareness about the differences connected to the different modalities is also conducted. Section 2.3 discusses how direct and indirect speech can be integrated into narrative frames, and a distinction between *personal* and *fictional narratives* is made. The first type of narrative refers to reports of personal experiences, and the latter type refers to narratives produced within

contexts of fantasy, for instance, within make-believe play (Ninio & Snow 1996).

Section 2.4 is devoted to a detailed presentation and critical examination of previous research of direct and indirect speech. This review leads to the development of a theoretical model that serves as a base for the analyses of the data included in the study. This model is shown below.
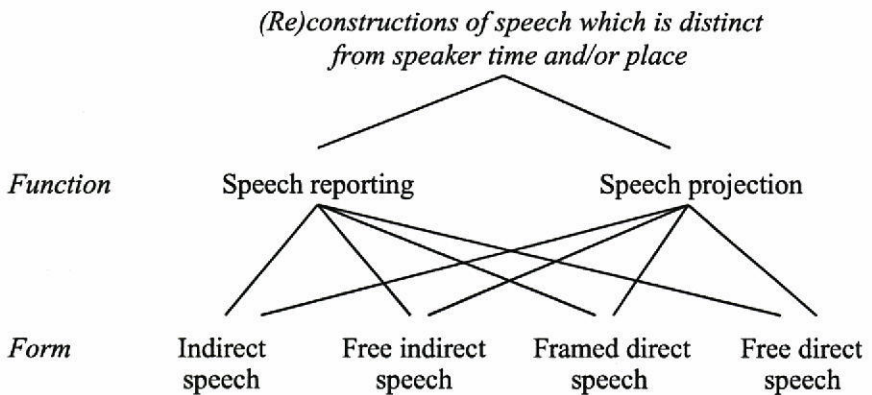

(Re)constructions of speech which is distinct from speaker time and/or place

Function   Speech reporting   Speech projection

Form   Indirect speech   Free indirect speech   Framed direct speech   Free direct speech

Figure 2-1: (Re)constructions of speech which is distinct from speaker time and/or place; types of speech (functions), types of speech structures (forms), and their possible combinations.

The model thus covers *(Re)constructions of speech which is distinct from speaker time and/or place*, and two major types of functions are distinguished – *speech reporting* and *speech projection*. The former of these two functions concerns speech that passes on information contained in earlier actual utterances. This refers typically to the use of either of the four forms of direct and indirect speech in personal narratives. The latter type, speech projection, refers to uses of direct and indirect speech in fictional narratives (i.e., where no original utterance exists or when the original utterance could only be seen as generalized script). In Section 2.2, *form* is defined as a unit with certain grammatical-structural properties, and in relation to this model, it is argued that at least four types of linguistic forms can be used to express (re)constructions of speech: *indirect speech, framed direct speech, free direct speech,* and *free*

indirect speech (see Section 2.4.2.1 (pages 55-59), and Table 2:1 (page 66) and Table 2:2 (page 68), for illustrations of the forms). The analyses of the children's language use, thus explore to what extent these forms are used. In addition, as illustrated in Figure 2-2 (page 84), different types of information can be *packaged* (Berman and Slobin 1994, Strömqvist 1996) into these forms, and these forms typically express several functions simultaneously. The aspects of packaging of information and multifunctionality, are also examined from a developmental point of view. Moreover, the review of previous research of children's use of direct and indirect speech in Section 2.4.4, reveals that the concentration has primarily been on speech reporting, whereas this study concentrates on speech projection (i.e. projections onto dolls and story characters).

Chapter 3, *Methodology*, then presents corpora consisting of altogether close to 275,000 words. The data are of three major types. The first type refers to two naturalistic Longitudinal case studies of a boy and a girl, between 1 1/2 – 4 years of age, interacting with their close family members (Strömqvist, Richthoff & Andersson 1993, Richthoff 2000). The second corpus consists of fourteen dyads of 3-year-olds and their mothers playing make-believe with a doll house for about half an hour (Söderbergh 1982; Wiberg, Humble & de Château 1989). The third and final corpus is comprised of spoken narratives elicited by means of the picture story book *Frog, where are you?* (Mayer 1969, see Appendix). The data collected include oral narratives by 3-, 4-, 5-, 9-, 12-, 15-year-olds and adults (Berglund & Eriksson 2000; Nauclér & Boyd 1996, 1997; Strömqvist *et al* 1994, 1998). The older subjects (the 9-, 12-, 15-year-olds and the adults) additionally produced a written version of the story. Section 3.3, describes the operationalizations used, the coding procedures used, and the types of analysis carried out, in order to answer the research questions. The types of analysis include quantitative as well as qualitative measurements. These procedures allow quantitative comparisons of use among subjects, activities and modes of production, as well as qualitative analyses of, e.g., individual strategies.

Chapter 4, *Results*, presents the results of the analyses. The first section (Section 4.1) reveals the findings of the analyses of the Longitudinal case study data (for a summary overview of the development of the forms in these children, see Table 4:1, page 165). The subsequent section (Section 4.2) contains the results of the analyses of the 3-year-olds narrating the frog story, and those children playing with the doll house. Sections 4.3 and 4.4 present the outcomes of the analyses of the narrative data elicited by means of the picture story book, whereas the final section (Section 4.5) focuses on the analyses of occurrences of free indirect speech in the data. Each of these sections is concluded by a summary of the results.

In Chapter 5, *General discussion*, the research questions are answered and the findings discussed. The most important findings include the following ones:

• Based on the analyses of the Longitudinal case study data, the development of forms occurs as follows: free direct speech is the first form to appear (around 26 months), this is followed by framed direct speech (28-30 months of age), and slightly later (and not as common in use) indirect speech (at 35-36 months). Free indirect speech was extremely infrequent in the data of the pre-school children.

• The emergence and the development of the use of the forms can be related to the development of understanding of minds. The ascription of speech to others by the two children starts during the same time period as previous research has indicated (e.g. Poulin-Dubois & Shultz 1988).

• The adults interacting with the two children made adjustments to the level of development of the child. Typically, the adults in the early phases produced direct and indirect speech themselves. As the children grew older, the adults prompted them to report speech, and little by little the children started to produce the forms independently.

• The results of the analyses indicate that children at age 3 can integrate direct and indirect speech into personal as well as fictional narrative frames, use changes in voice quality, make use of gestures, and select and modify the pronunciation of a particular word in order to depict the speech of a speaker. Speech quotations were in addition to evaluations employed by the children in order to move the plot forward. Moreover, at this age the children distinguished between different types of activities by using free direct speech in relation to

make-believe play (such as playing with dolls), and framed direct speech in story-telling activities (e.g., when relating the frog story).

• The 3-year-olds narrating the frog story failed to convey shifts in speaker perspective in around half of the cases, whereas the 4- and 5-year-olds clearly managed better in this respect. This is in accordance with previous research showing that 4-year-olds (but not 3-year-olds) understand false belief (e.g. Flavell 1988). The school children (the 9-, 12-, 15-year-olds) had little difficulty marking shifts in speaker perspective in their oral narrations. However, the changes of perspective were not conveyed clearly in some of the 9- and 12-year-olds' written narratives, since the quotations were not framed and marked properly. Thus, the skill of *speaking within writing* does not seem to be fully developed at these ages.

• The type of narration performed by the majority of the school children had a *detached* (Chafe 1982) feel, due to the relatively infrequent use of speech projections, and the preference for indirect speech. However, in the 9-year-old group some narratives included a great deal of free as well as framed direct speech and were thus characterized by *involvement*. This was particularly true of the written narratives.

• The adult subjects produced significantly longer narratives, and they included a greater number of speech projections than the younger narrators. Moreover, the quotes were longer than in any other age group, and the adult narrators projected both speech and thinking onto a variety of story characters.

• The adult group of narrators was found to be the only one differentiating clearly between the spoken and the written productions. Consequently, the written narratives are longer than the spoken ones, they include significantly more speech projections in the written narratives, a greater number of different types of verbs of saying are used, and free indirect speech and mental verbs are used to a greater extent in the adults' writing than in their speaking. The adult narrators thus had command over *speaking within writing* as well as *speaking within speaking*.

In the final chapter, *Conclusion*, the fact that the development in learning how information can be packaged into direct and indirect speech and how these forms can be used for different purposes is a truly long-term process is emphasized. Several factors are suggested as having an impact on this development. Besides cognitive and linguistic maturity, children's interaction with adults and the adults' language behavior, the children's amount of practice

and experience of language use in different types of activities, their exposure to story-telling in different contexts and modalities, the functionality, and the conditions of production (speech and writing), are examples of such factors. However, further research needs to be conducted in order to obtain more insight into these issues, and to gain further knowledge about the forms and functions of direct and indirect speech (for suggestions of future research, see Section 6.2).

REFERENCES

Aisenman, R. (ed.) (1999). *Working Papers in Developing literacy across genres, modalities and languages. Volume 1*. Tel Aviv University, Dept of Linguistics.

Aksu-Koç, A. (1994). Development of linguistic forms: Turkish. In R. A. Berman & D. I. Slobin (eds.), *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Aksu-Koç, A. & von Stutterheim, Ch. (1994). Temporal relations in the development of narrative. In R. A. Berman & D. I. Slobin (eds.), *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Allwood, J. (1976) *Linguistic Communication as Action and Cooperation*. Gothenburg Monographs in Linguistics 2. Göteborg University, Dept of Linguistics.

Allwood, J. (1980) An analysis of communicative action. In M. Brenner (ed.), *The Structure of Action*. Oxford: Blackwell.

Allwood, J. (ed.) (1999). Talspråksfrekvenser - ny och utvidgad upplaga. Frekvenser för ord och kollokationer i svenskt tal- och skriftspråk. *Gothenburg Papers in Theoretical Linguistics* S21. Göteborg University, Dept of Linguistics.

Allwood, J., Nivre, J. & Ahlsén E. (1990). Speech-management - on the non-written life of speech. *Nordic Journal of Linguistics*, 13: 3-48.

Allwood, J., Nivre, J. & Ahlsén E. (1992). On the semantics and pragmatics of linguistic feedback. *Journal of Semantics*, 9: 1-26.

Andersen, E. (1990). *Speaking with style: The sociolinguistic skills of children*. London: Routledge.

Astington, J. (1990). Narrative and the child's theory of mind. In B. K. Britton & A. D. Pellegrini (eds.), *Narrative thought and narrative language*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Austin, J. L. (1962). *How to do things with words*. Oxford: Oxford University Press.

Baker, C. D. & Freebody, P. (1989). *Children's first school books*. Oxford: Blackwell.

Bakhtin, M. M. (1981). *The dialogic imagination*. Austin, Texas: The University of Texas Press.

REFERENCES

Bakhtin, M. M. (1986). *Speech genres and other late essays*. Austin, Texas: The University of Texas Press.

Bal, M. (1985). *Narratology: Introduction to the theory of narrative*. (Translated by C. van Boheemen.) Toronto: University of Toronto Press.

Bally, C. (1912). Le style indirect libre en français moderne. *Germanisch-Romanische Monatsschrift* IV.

Bamberg, M. (1985). *Form and function in the construction of narratives: Developmental perspectives*. Unpublished doctoral dissertation, University of California, Berkeley.

Bamberg, M. (1991). Narrative activity as perspective taking: The role of emotionals, negations, and voice in the construction of the story realm. *Journal of Cognitive Psychotherapy*, 5: 275-290.

Bamberg, M. (1994). Development of linguistic forms: German. In R. A. Berman & D. I. Slobin (eds.), *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bamberg, M. & Marchman, V. (1990). What holds a narrative together? The linguistic encoding of episode boundaries. *Papers in Pragmatics*, 4: 58-121.

Bamberg, M. & Reilly, J. (1996). Emotion, narrative, and affect: How children discover the relationship between what to say and how to say it. In D. I. Slobin, J. Gerhardt, A. Kyratzis, & J. Guo (eds.), *Social interaction, social context and language: Essays in honor of Susan Ervin-Tripp*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bamgbose, A. (1986). Reported speech in Yoruba. In F. Coulmas (ed.), *Direct and indirect speech*. Berlin: Mouton de Gruyter.

Banfield, A. (1982). *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Boston and London: Routledge & Kegan Paul.

Banfield, A. (1993). Where epistemology, style, and grammar meet literary history: the development of represented speech and thought. In J. A. Lucy (eds.), *Reflexive language: reported speech and metapragmatics*. Cambridge: Cambridge University Press.

Barnes, S., Gutfreund, M., Satterly, D., & Wells, G. (1983). Characteristics of adult speech which predict children's langauge development. *Journal of Child Language*, 10: 65-84.

Bates, E., Bretherton, I. & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge: Cambridge University Press.

Bates, E. & MacWhinney, B. (1979). A functionalist approach to the acquisition of grammar. In E. Ochs & B. Schieffelin (eds.), *Developmental Pragmatics*. New York: Academic Press.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, .P., Reznick, J., Reilly, J. & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, 21: 85-123.

Beal, C. R. (1988). Children's knowledge about representations of intended meaning. In J. W. Astington, P. L. Harris & D. R. Olson (eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.

Berglund, E. (1998). Language assessment of Swedish three-year-olds using a protocol of natural speech: PLUS-3. *Reports from the Department of Psychology*, Stockholm University. No. 846.

Berglund, E. (1999). *Early communicative and language development in Swedish children: Methods, results, clinical implications, and prospects for the future*. Doctoral Dissertation. Stockholm University, Dept of Psychology.

Berglund, E. & Eriksson, M. (2000). Communicative development in Swedish children 16-28 months old. The Swedish early communicative development inventory - words and sentences. *Scandinavian Journal of Psychology*, 41, 2.

Berman, R. A. (1996). Form and function in developing narrative abilities. In D. I. Slobin, J. Gerhardt, A. Kyratzis, & J. Guo (eds.), *Social interaction, social context and language: Essays in honor of Susan Ervin-Tripp*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Berman, R. A. & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Caldas-Coulthard, C. R. (1987). Reported speech in written narrative texts. In M. Coulthard (ed.), *Discussing Discourse*. Birmingham: University of Birmingham.

Caldas-Coulthard, C. R. (1994). On reporting reporting: The representation of speech in factual and factional narratives. In M. Coulthard (ed.), *Advances in Written Text Analysis*. London: Routledge.

Chafe, W. L. (1982). Integration and involvement in speaking, writing and oral literature. In D. Tannen (ed.), *Spoken and written language: Exploring orality and literacy*. Norwood, NJ: Ablex Publishing Corporation.

Chafe, W. L. (1985). Linguistic differences produced by differences between speaking and writing. In D. R. Olson, N. Torrance & A. Hildyard (eds.),

Literacy, langauge, and learning: The nature and consequences of reading and writing. Cambridge: Cambridge University Press.

de Château, P. & Wiberg, B. (1977a). Long-term effect on mother-infant behavior of extra contact during the first hour post partum: First observations at 36 hours. *Acta Pædiatrica Scandinavica*, 66: 137-144.

de Château, P. & Wiberg, B. (1977b). Long-term effect on mother-infant behavior of extra contact during the first hour post partum: A follow-up at three months. *Acta Pædiatrica Scandinavica*, 66: 145-151.

de Château, P. & Wiberg, B. (1984). Long-term effect on mother-infant behavior of extra contact during the first hour post partum: Follow-up at one year. *Scandinavian Journal of Social Medicine*, 12: 91.

Clark, H. H. & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, 66: 764-805.

Cooley, C. H. (1902). *Human nature and the social order.* New York: Scribner.

Corsaro, W. A. (1986). Discourse processes within peer culture: From a constructivist to an interpretative approach to childhood socialization. *Sociological Studies of Child Development*, 1: 81-101.

Coulmas, F. (1985). Direct and indirect speech. General problems and problems of Japanese. *Journal of Pragmatics*, 9: 41-63.

Coulmas, F. (ed.) (1986). *Direct and indirect speech.* Berlin: Mouton de Gruyter.

Coulmas, F. (1986a). Reported speech: Some general issues. In F. Coulmas (ed.), *Direct and indirect speech.* Berlin: Mouton de Gruyter.

Coulmas, F. (1986b). Direct and indirect speech in Japanese. In F. Coulmas (ed.), *Direct and indirect speech.* Berlin: Mouton de Gruyter.

Coulmas, F. (1998). Reported speech. In J. L. Mey (ed.), *Concise encyclopedia of pragmatics.* Amsterdam: Elsevier.

Dore, J. (1989). Monologue as reenvoicement of dialogue. In K. Nelson (ed.), *Narratives from the crib.* Cambridge: Harvard University Press.

Dubois, B. L. (1989). Pseudoquotation in current English communication: "Hey, she didn't really say it". *Language and Society*, 18: 343-359.

Eisner, J. (1975). *A grammar of oral narrative.* Unpublished Ph.D. dissertation, University of Michigan.

Ely, R. & McCabe, A. (1993). Remembered voices. *Journal of Child Language*, 20: 671-696.

Ely, R. & Gleason J. B. (1995). Socialization across contexts. In P. Fletcher & B. MacWhinney (eds.), *The Handbook of Child Language.* Blackwell.

Ely, R., Gleason, J. B., Narasimhan, B., & McCabe, A. (1995). Family talk about talk: Mothers lead the way. *Discourse Processes*, 19, 201 - 218.

Ely, R., & McCabe, A. (1996a). Gender differences in memories for speech. In P. Thompson (ed.), *International yearbook of oral history and life stories: Memory and gender*. New York: Oxford University Press.

Ely, R., Gleason, J. B., & McCabe, A. (1996b). "Why didn't you talk to your mommy, honey?": Parents' and children's talk about talk. *Research on Language and Social Interaction*, 29 (1): 7-25.

Emmorey, K. & Reilly, J. (1998). The development of quotation and reported action: Conveying perspective in ASL. In E. V. Clark (ed.), *The proceedings of the twenty-ninth annual Child Language Research Forum*. Stanford: CSLI Publications.

Engberg-Pedersen, E. (1993). Point of view expressed through shifters. In K. Emmorey & J. Reilly (eds.), *Language, Gesture, and Space*. Hillsdale, NJ: LEA.

Eriksson, Ma. (1994). Anföringens funktion i ungdomars berättelser. In A. Holmberg & K. Larsson (eds.), *Svenskans beskrivning 20*. Lund: Lund University Press.

Eriksson, Ma. (1995). A case of grammaticalization in modern Swedish: The use of 'ba' in adolescent speech. *Language Sciences*, 17: 19-48.

Eriksson, Må. (*submitted*). Narratives Validates Communicative Development Inventories.

Eriksson, Må. & Berglund, E. (1999). Swedish early communicative development inventories: words and gestures. *First Language*, Vol. 19, 55: 55-90.

Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.

Farrar, M. J. (1990). Discourse and the acquisition of grammatical morphemes. *Journal of Child Language*, 17: 607-624.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethnick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D. J., Bates, E., Hartung, J. P., Pethnick, S. J. & Reilly, J. (1993). *The MacArthur Communicative Development Inventories. User's guide and technical manual*. San Diego: Singular Publishings Group.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B. & Ikuko F. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16: 477-501.

REFERENCES

Ferreiro, E. & Pontecorvo, C. (1999). Managing the written text: the beginning of punctuation in children's writing. *Learning and Instruction*, 9: 543-564.

Ferreiro, E. & Zucchermaglio, C. (1996). Children's use of punctuation marks: the case of 'quoted speech'. In C. Pontecorvo, M. Orsolini & B. Burge (eds.), *Children's early text construction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Fielding, H. (1998). *Bridget Jones's diary: A novel*. Picador.

Flavell, J. H. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. W. Astington, P. L. Harris & D. R. Olson (eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.

Flavell, J. H., Speer, J. R., Green, F. L. & August, D. L. (1981). The development of comprehension monitoring and knowledge about communication. *Monographs of the Society for Research in Child Development*, 46.

Fónagy, I. (1986). Reported speech in French and Hungarian. In F. Coulmas (ed.), *Direct and indirect speech*. Berlin: Mouton de Gruyter.

Forguson, L. & Gopnik, A. (1988). The ontogeny of common sense. In J. W. Astington, P. L. Harris & D. R. Olson (eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.

Gaarder, J. (1996). *Sophie's world*. (Translated by P. Møller.) London: Phoenix.

Garvey, C. (1990). *Play*. Cambridge, MA: Harvard University Press. (Enlarged edition; original ed. 1977.)

Genette, G. (1988). *Narrative discourse revisited*. Ithaca, NY: Cornell University Press.

Gombert, J. E. (1990). *Metalinguistic development*. Chicago: University of Chicago Press.

Goodell, E. W. & Sachs, J. (1992). Direct and indirect speech in children's retold narratives. *Discourse Processes*, 15: 395-422.

Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. L. Morgan (eds.), *Syntax and Semantics 3: Speech Acts*. Academic Press: New York.

Gayraud, F., Jisa, H., and Viguié, A. (2000) The development of syntactic packaging in French children's written and spoken texts. In E. Espéret (ed.), *Writing and learning to write at the dawn of the 21st century*. Proceedings from the 1998 European writing conference, Universite de Poitiers.

Haberland, H. (1986). Reported speech in Danish. In F. Coulmas (ed.), *Direct and indirect speech*. Berlin: Mouton de Gruyter.

Hall, N. & Robinson, A. (eds.) (1996). *Learning about punctuation.* Portsmouth, NH: Heinemann.

Heath, S. B. (1986). What no bedtime story means: narrative skills at home and school. In B. Schieffelin & E. Ochs (eds.), *Language socialization across cultures.* Cambridge: Cambridge University Press.

Hickmann, M. (1982). *The development of narrative skills. Pragmatic and metapragmatic aspects of discourse cohesion.* Doctoral dissertation. University of Chicago.

Hickmann, M. (1993). The boundaries of reported speech in narrative discourse: some developmental aspects. In J. A. Lucy (ed.), *Reflexive language: reported speech and metapragmatics.* Cambridge: Cambridge University Press.

Hildyard, A. & Hidi S. (1985). Oral-written differences in the production and recall of narratives. In D. R. Olson, N. Torrance & A. Hildyard (eds.), *Literacy, langauge, and learning: The nature and consequences of reading and writing.* Cambridge: Cambridge University Press.

Hjelmquist, E. & Gidlund, Å. (1985). Free recall of conversations. *Text*, 5: 169-185.

Homer (1988). *The Iliad.* (Translated by A. T. Murray.) London: Harvard University Press.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: relation to language input and gender. *Developmental Psychology*, 27: 236-248.

Jespersen, O. (1924). *The philosophy of grammar.* London: Allen and Unwin.

Kalepsky, T. (1899). Zür französischen Syntax. *Zeitschrift für Romanische Philologie* XXIII: 491-513.

Karmiloff-Smith, A. (1979). Micro- and macrodevelopmental changes in language acquisition and other representational systems. *Cognitive Science*, 3(2): 91-118.

Karmiloff-Smith, A. (1981). The grammatical marking of thematic structure in the development of language production. In W. Deutsch (ed.), *The child's conception of language.* London: Academic Press.

Kavanaugh, R. D., Eizenman, D. R. & Harris, P. L. (1997). Young children's understanding of pretense expressions of independent agency. *Developmental Psychology*, Vol. 33, 5: 764-770.

Knutson, U. (1998). [Interview with Eva Beckman]. *Månadsjournalen*, 5: 20-25.

Kotsinas, U. (1994). *Ungdomsspråk.* Uppsala: Hallgren och Fallgren Studieförlag.

REFERENCES

Kvavik, K. H. (1986). Characteristics of direct and reported speech prosody: Evidence from Spanish. In F. Coulmas (ed.), *Direct and indirect speech.* Berlin: Mouton de Gruyter.

Labov, W. (1972). *Language in the Inner City: Studies in the Black English Vernacular.* Philadelphia, PA: University of Pennsylvania Press.

Labov, W. & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helms (ed.), *Essays on the verbal and visual arts.* Seattle, WA: University of Washington Press.

Ledin, P. (1998). Att sätta punkt. Hur elever på låg- och mellanstadiet använder meningen i sina uppsatser. *Språk och Stil*, 8: 5-47.

Leech, G. N. & Short, M. (1981). *Style in fiction.* London: Longman.

Lehrer, A. (1989). Remembering and representing prose: Quoted speech as a data source. *Discourse Processes*, 12: 105-125.

Lerch, G. (1919). *Unegeintliche direkte Rede.* Diss. Munich.

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review*, 94, 412-426.

Leslie, A. M. (1988). Some implications of pretense for mechanisms underlying the child's theory of mind. In J. W. Astington, P. L. Harris & D. R. Olson (eds.), *Developing theories of mind.* Cambridge: Cambridge University Press.

Li, C. N. (1986). Direct and indirect speech: A functional study. In F. Coulmas (ed.), *Direct and indirect speech.* Berlin: Mouton de Gruyter.

Lieven, E. V. M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B. J. Richards (eds.), *Input and interaction in langauge acquisition.* Cambridge: Cambridge University Press.

Linell, P. (1982). The written language bias in Linguistics. *Studies in Communication*, 2. Department of Communication Studies, University of Linköping.

Lips, M. (1926). *Le style indirect libre.* Paris: Payot.

Lorck, E. (1921). *Die erlebte Rede.* Eine sprachliche Untersuchung. Heidelberg: Carl Winter.

MacWhinney, B. (1995). *The CHILDES project: Tools for analyzing talk.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Markman, E. M. (1977). Realizing you don't understand: A preliminary investigation. *Child Development*, 50: 643-655.

Mayer, M. (1969). *Frog, where are you?* New York: Dial Press

Maynard, S. K. (1986). The particle -*o* and content-oriented indirect speech in Japanese written discourse. In F. Coulmas (ed.), *Direct and indirect speech.* Berlin: Mouton de Gruyter.

Mead, G. H. (1934). *Mind, self and society: from the standpoint of a social behaviorist.* (Edited, with introduction, by Charles W. Morris.) Chicago: Chicago University Press.

Meehan, T. (1991). It's like, 'what's happening in the evolution of like'?: A theory of grammaticalization. *Kansas Working Papers in Linguistics*, 16: 37-51.

Michaels, S. & Collins, J. (1984). Oral discourse style: Classroom interaction and the acquisition of literacy. In D. Tannen (ed.), *Coherence in spoken and written discourse.* Norwood, NJ: Ablex.

Montgomery, D. E. & Montgomery, D. A. (1999). The influence of movement and outcome on young children's attributions of intention. *British Journal of Developmental Psychology*, 17: 245-261.

Nauclér, K. & Boyd, S. (1996). Turkish and Swedish pre-school children's language socialization – Implications for literacy development. In L. Huss (ed.), *Många vägar till tvåspråkighet.* Uppsala Multiethnic Papers 38, Uppsala University.

Nauclér, K. & Boyd, S. (1997). Telling a story. Norms for interaction with Turkish and Swedish children at home and in pre-school. In A. Sjögren (eds.), *Language and Environment.* Botkyrka: Mångkulturellt centrum.

Nelson, K. (1986). *Event knowledge: Structure and function in development.* Hillsdale NJ: Erlbaum.

Nelson, K. (ed.) (1989). *Narratives from the crib.* Cambridge: Harvard University Press.

Nelson, K. E., Carskaddon, G., & Bonvillian, J. (1973). Syntax acquisition: impact of experimental variation in adult verbal interaction with the child. *Child development*, 44: 497-504.

Nippold, M. A. (1998). *Later language development. The school-age and adolescent years.* (2nd ed.). Austin, TX: Pro-Ed.

Nordberg, B. (1985). Om ungdomars samtalsstil. Några preliminära iakttagelser. *Nysvenska studier*, 64: 5-27.

Nordberg, B. (1986). The use of onomatopoeia in conversational style of adolescents. *FUMS rapport*, 132. Uppsala.

Nordberg, B. (1992). Onomatopoetiska uttryck och ungdomars samtalsstil. In S. Strömquist (ed.), *Tal och samtal.* Lund: Studentlitteratur.

REFERENCES

Nordqvist, Å. (1998a). The development of (in)direct speech in spoken and written narratives. In S. Strömqvist & E. Ahlsén (eds.), *The Process of Writing - a progress report*. Göteborg University, Dept of Linguistics.

Nordqvist, Å. (1998b). Projecting speech to protagonists in oral and written narratives: a developmental study. *Psychology of Language and Communication*, Vol. 2., 2: 37-46.

Nordqvist, Å. (1998c). Investigating problems with understanding in interaction between child and adult. In *Proceedings from the XVIth Scandinavian Conference in Linguistics*. Åbo University, Dept of Linguistics.

Nordqvist, Å. (1998d). The influence of a physically present listener in a narrative situation. An experimental comparative study. Ms. Göteborg University, Dept of Linguistics.

Nordqvist, Å. (1999). The participation of tense in narrative construction: from speech to speech and writing. Paper presented at *VIIIth International Congress for the Study of Child Language*. University of the Basque Country.

Nordqvist, Å. (2000a). Anförda meningar: en granskning av SAG. In E. Engdahl & K. Norén (eds.), *Att använda SAG - 29 uppsatser om Svenska Akademiens Grammatik*. MISS 33. Göteborg University, Department of Swedish Language.

Nordqvist, Å. (2000b). Speech reporting, early language development and gender differences: Investigating production and input in longitudinal case studies of two siblings. Paper presented at *First International Gender and Language Association Conference*. Stanford University.

Nordqvist, Å. & Strömqvist, S. (1995). On the development of speech and writing in Swedish school children - an experimental study. Paper presented at *The XVth Scandinavian Conference of Linguistics*. Oslo University.

Olson, D. R. (1991). Literacy as a metalinguistic activity. In D. R. Olson & N. Torrance (eds.), *Literacy and orality*. Cambridge: Cambridge University Press.

Padden, C. (1986). Verbs and role-shifting in ASL. In C. Padden (ed.), *Proceedings of the Fourth National Symposium on Sign Language Research and Teaching*. National Association of the Deaf: Silver Spring, MD.

Page, N. (1988). *Speech in the English novel*. (2nd ed.) London: Mac Millan.

Palmer, F. R. (1976). *Semantics*. Cambridge: Cambridge University Press.

Partee, B. H. (1973). The syntax and semantics of quotation. In S. R. Anderson & P. Kiparsky (eds.), *A festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston.

Perera, K. (1996). Who says what? Learning to 'read' the punctuation of direct speech. In N. Hall & A. Robinson (eds.), *Learning about punctuation*. Portsmouth, NH: Heinemann.

Perner, J. (1988). Higher-order beliefs and intentions in children's understanding of social interaction. In J. W. Astington, P. L. Harris & D. R. Olson (eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.

Peterson, C. & McCabe, A. (1983). *Developmental psycholinguistics: three ways of looking at a child's narrative*. New York: Plenum.

Piaget, J. (1926). *The language and thought of the child*. London: Routledge and Kegan Paul.

Pine, J. M. (1994). The language of primary caregivers. In C. Gallaway & B. J. Richards (eds.), *Input and interaction in langauge acquisition*. Cambridge: Cambridge University Press.

Plunkett, K. & Strömqvist, S. (1992). The acquisition of Scandinavian languages. In D. I. Slobin (ed.), *The crosslinguistic study of language acquisition, Vol. 3*. Lawrence Erlbaum: Hillsdale, new Jersey.

Polanyi, L. (1982). Literary complexity in everyday storytelling. In D. Tannen (ed.), *Spoken and written language: Exploring orality and literacy*. Norwood, NJ: Ablex Publishing Corporation.

Pontecorvo, C. & Orsolini, M. (1996). Writing and written language in children's development. In C. Pontecorvo, M. Orsolini & B. Burge (eds.), *Children's early text construction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Poulin-Dubois, D. & Shultz, T. R. (1988). The development of the understanding of human behavior: From agency to intentionality. In J. W. Astington, P. L. Harris & D. R. Olson (eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.

Ragnarsdóttir, H. (1992). Episodic structure and interclausal connectives in Icelandic children's narratives. In R. Söderbergh (ed.), *Berättelser för och av barn*. Colloquium Paedolinguisticum Lundensis 1991. Lund University, Department of Linguistics. Paper no 8:33-45.

Reilly, J. S. (1992). How to tell a good story: The intersection of language and affect in children's narratives. *Journal of Narrative and Life History*, 2: 355-377.

Richards, B. J. (1994). Child-directed speech and influences on language acquisition: methodology and interpretation. In C. Gallaway & B. J.

REFERENCES

Richards (eds.), *Input and interaction in langauge acquisition*. Cambridge: Cambridge University Press.

Richthoff, U. (2000). *En svensk barnspråkskorpus. Uppbyggnad och analyser*. Licentiate Thesis. Department of Linguistics, Göteborg University.

Romaine, S. & Lange, D. (1991). The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech*, 66 (3): 227-279.

Sachs, J. (1983). Talking about the there and then: The emergence of displaced reference in parent-child discourse. In K. E. Nelson (ed.), *Children's language*. Hillsdale, NJ: Erlbaum.

Sacks, H., Schegloff, E. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Langauge*, 50: 696-735.

Sawyer, R. K. (1996). Role voicing, gender, and age in pre-school play discourse. *Discourse Processes*, 22: 238-256.

Shatz, M., Hoff-Ginsberg, E., & MacIver, D. (1989). Induction and the acquisition of English auxilliaries: the effects of differentially enriched input. *Journal of Child Language*, 16: 141-160.

Saussure, F. de (1916). *Cours de linguistique générale*. Paris: Payot.

Slobin, D. I. (1973). Cognitive prerequisities for the development of grammar. In C. A. Ferguson & D. I. Slobin (eds.), *Studies of child language development*. New York: Holt, Rinehart, & Winston.

Snow, C. E. & Ferguson, C. A. (eds.) (1977). *Talking to children. Language input and acquisition*. Cambridge: Cambridge University Press.

Snow, C. E., Perlmann, R., & Nathan, D. (1987). Why routines are different: towards a multiple-factors model of the relation between input and language acquisition. In K. E. Nelson & A. van Kleeck (eds.), *Children's langauge*, vol. VI. NJ: Erlbaum.

Sokolov, J. L. & Snow, C. E. (1994). The changing role of negative evidence in theories of langauge development. In C. Gallaway & B. J. Richards (eds.), *Input and interaction in langauge acquisition*. Cambridge: Cambridge University Press.

Strömqvist, S. (1984). *Make-believe through words. A linguistic study of children's play with a doll house*. Gothenburg Monographs in Linguistics 4. Department of Linguistics, Göteborg University.

Strömqvist, S. (1996). Discourse flow and linguistic information structuring: Explorations in speech and writing. *Gothenburg Papers in Theoretical Linguistics 78*. Dept of Linguistics, Göteborg University.

Strömqvist, S. (1998). Child language - An overview. In J. L. Mey (ed.), *Concise encyclopedia of pragmatics*. Amsterdam: Elsevier.

Strömqvist, S. (*to appear*). Language acquisition in early childhood. In G. Rickheit, T. Herrmann, and W. Deutsch (eds.), *International handbook of psycholinguistics*. Berlin: Walter de Gruyter.

Strömqvist, S. & Ahlsén, E. (eds.) (1998). *The Process of Writing - a progress report*. Göteborg University, Dept of Linguistics.

Strömqvist, S., Ahlsén, E. & Wengelin, Å. (1998). The production process in speech and writing. In S. Strömqvist & E. Ahlsén (eds.), *The Process of Writing - a progress report*. Göteborg University, Dept of Linguistics.

Strömqvist, S. & Day, D. (1993). On the development of narrative structure in child L1 and adult L2 acquisition. *Applied Psycholinguistics*, 14: 135-158.

Strömqvist, S. & Hellstrand, Å. (1994). Tala och skriva i lingvistiskt och didaktiskt perspektiv - en projektbeskrivning. *Didaktisk Tidskrift*, Göteborgs universitet, Nr. 1-2.

Strömqvist, S. & Malmsten, L. (1998). *ScriptLog Pro 1.04 – User's Manual*. Technical report, Department of Linguistics, Göteborg University.

Strömqvist, S., Nordqvist, Å. & Wengelin, Å. (*forthcoming*). Writing the frog story – developmental and cross-modal perspectives. In S. Strömqvist & L. Verhoeven (eds.), *The Frog Story Revisited*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Strömqvist, S., Richthoff, U. & Andersson, A.-B. (1993). Strömqvist's and Richthoff's Corpora - a guide to longitudinal data from four Swedish children. *Gothenburg Papers in Theoretical Linguistics* 66. Department of Linguistics, Göteborg University.

Strömqvist, S. & Verhoeven, L. (*forthcoming*). *The frog story revisited*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Svenska skrivregler (1991). Svenska språknämnden. Stockholm: Almqvist & Wiksell.

Söderbergh, R. (1973). *Project Child Language Syntax and Project Early Reading*. Department of Scandinavian Languages, Stockholm University.

Söderbergh, R. (1980). Story-telling, dramatic role play and displaced speech in play with dolls. *First Language*, Vol. 1, 3: 209-222.

Söderbergh, R. (1982). Linguistic effects by three years of age of extra contact during the first hour post partum. In C. E. Johnson & C. Larzon (eds.), *Proceedings of the Second International Congress for the Study of Child Language*. Washington D.C.: University Press of America.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (ed.), *Language typology and syntactic description: Vol. 3. Grammatical categories and the lexicon*. Cambridge: Cambridge University Press.

REFERENCES

Tannen, D. (1982). The oral/literate continuum in discourse. In D. Tannen (ed.), *Spoken and written language: Exploring orality and literacy*. Norwood, NJ: Ablex Publishing Corporation.

Tannen, D. (1985). Relative focus on involvement in oral and written discourse. In D. R. Olson, N. Torrance & A. Hildyard (eds.), *Literacy, langauge, and learning: The nature and consequences of reading and writing*. Cambridge: Cambridge University Press.

Tannen, D. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In F. Coulmas (ed.), *Direct and indirect speech*. Berlin: Mouton de Gruyter.

Tannen, D. (1987). The orality of literature and the literacy of conversation. In J. A. Langer (ed.), *Language, literacy, and culture: Issues of society and schooling*. Norwood, NJ: Ablex Publishing Corporation.

Tannen, D. (1989). *Talking voices. Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

The Merriam Webster Dictionary (1994). Springfield, Massachusetts: Merriam-Webster Inc.

Teleman, U., Hellberg, S. & Andersson, E. (1999). *Svenska Akademiens grammatik*. Stockholm: Norstedts förlag.

Toolan, M. (1998a). Narrative: Linguistic and structural theories. In J. L. Mey (ed.), *Concise encyclopedia of pragmatics*. Amsterdam: Elsevier.

Toolan, M. (1998b). Narrative, Natural. In J. L. Mey (ed.), *Concise encyclopedia of pragmatics*. Amsterdam: Elsevier.

Tobler, A. (1894). *Vermischte Beiträge zur französischen Grammatik II*. Leipzig.

Uccelli, P., Hemphill, L., Pan, B. A. & Snow, C. (*in press*). Telling two kinds of stories: Sources of narrative skill. In L. Balter & C. Tamis-LeMonda (eds.), *Child psychology: A handbook of contemporary issues*.

Vincent, D. & Perrin, L. (1999). On the narrative vs non-narrative functions of reported speech: A socio-pragmatic study. *Journal of Sociolinguistics*, 3/3: 291-313.

Voloshinov, V. N. (1973). *Marxism and the philosophy of langauge*. New York: Seminar Press.

Weijer, J. van de (1998). *Language input for word discovery*. Doctoral Dissertation. MPI Series in Psycholinguistics, 9.

Wellman, H. M. (1988). First steps in the child's theorizing about the mind. In J. W. Astington, P. L. Harris & D. R. Olson (eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.

Wengelin, Å. *(forthcoming)*. *The written language production process in adults with reading and writing difficulties.* Doctoral Dissertation. Gothenburg Monographs in Linguistics. Göteborg University.

Wiberg, B. & de Château, P. (1982). Long-term effect on mother-infant behavior of extra contact during the first hour post partum: Study design and method. In E. J. Anthony and C. Chiland (eds.), *The child in his family.* New York: John Wiley & Sons.

Wiberg, B., Humble, K. & de Château, P. (1989). Long-term effect on mother-infant behavior of extra contact during the first hour post partum: Follow-up at three years. *Scandinavian Journal of Social Medicine*, 2: 181-191.

Wimmer, H. and Perner, J.(1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.

Winberg, J. & de Château, P. (1982). Early social development: Studies of infant-mother interaction and relationships. *Review of Child Development Research*, 6: 1-44.

Wolf, D. & Hicks, D. (1989). The voices within narratives: the development of intertextuality in young children's stories. *Discourse Processes*, 12: 329-351.

Wolf, D., Rygh, J. & Altshuler, J. (1984). Agency and experience: Actions and states in play narratives. In I. Bretherton (ed.), *Symbolic play.* New York: Academic Press.

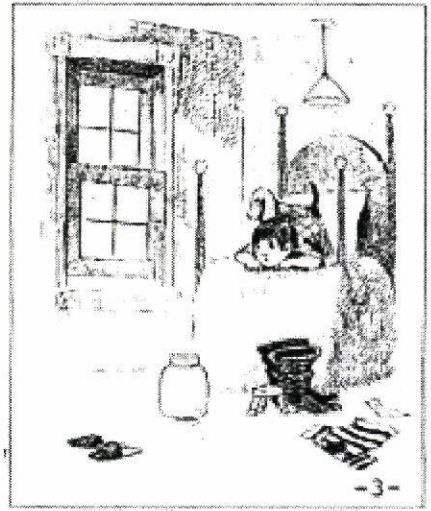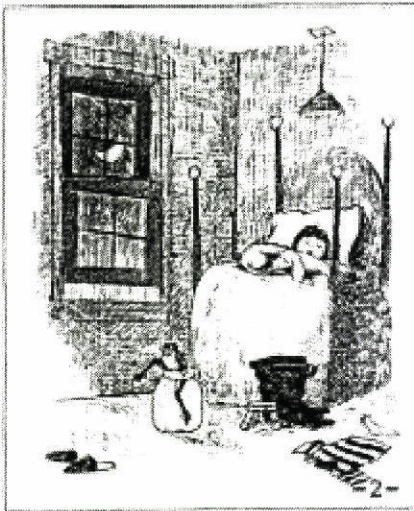Özyürek, A. (1996). How children talk about a conversation. *Journal of Child Language*, 23: 693-714.

APPENDIX

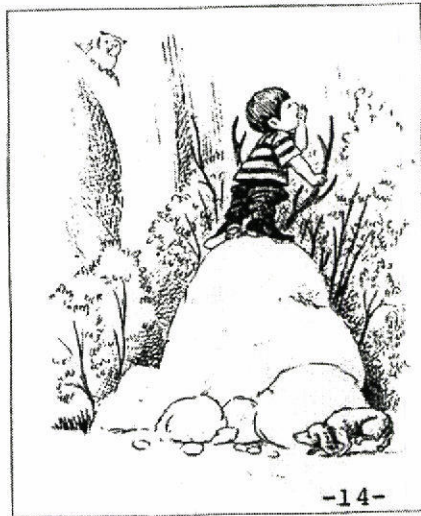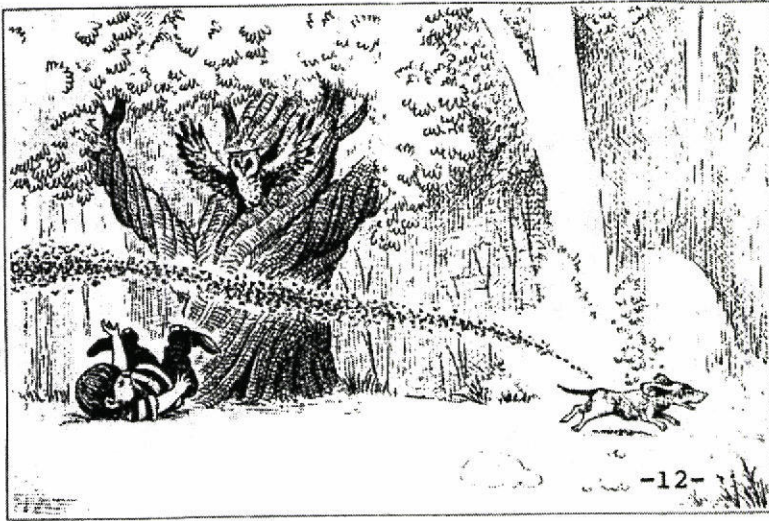*Frog, where are you?*[1]
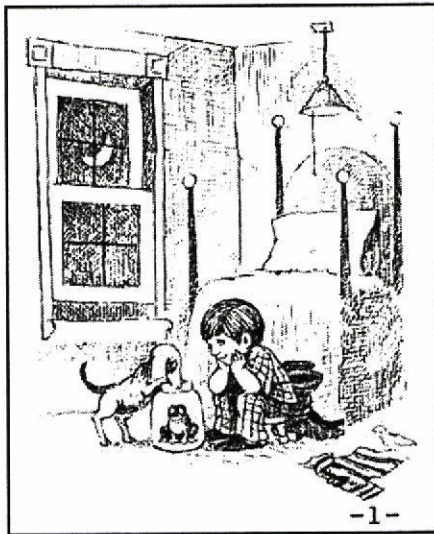
−4−

−5−

−8−

−9−

−10−

−11−

-12-

-13-

-14-
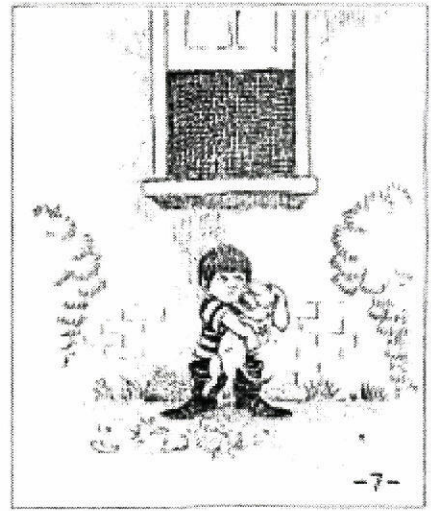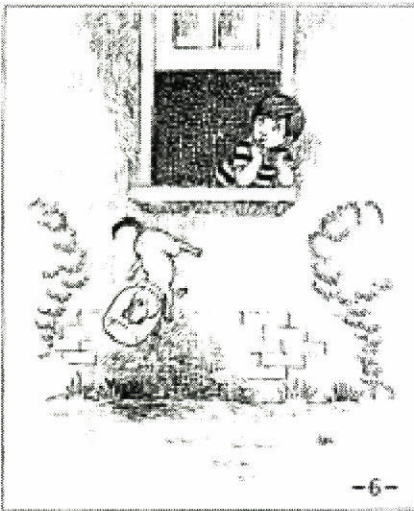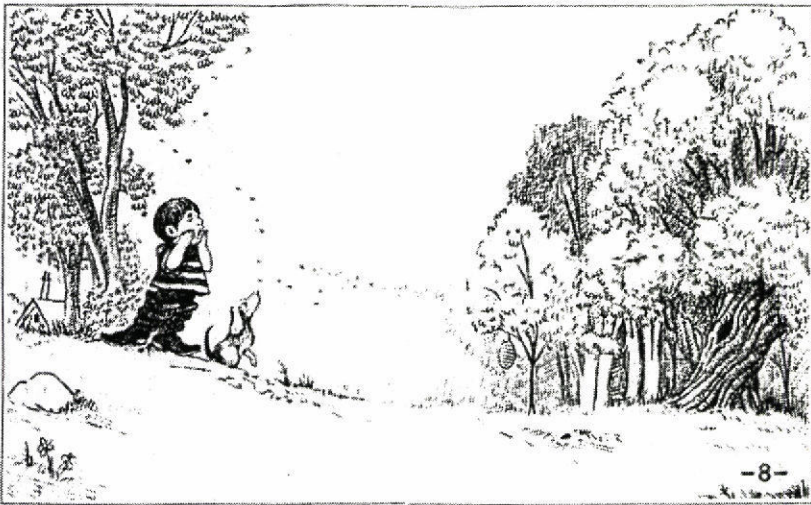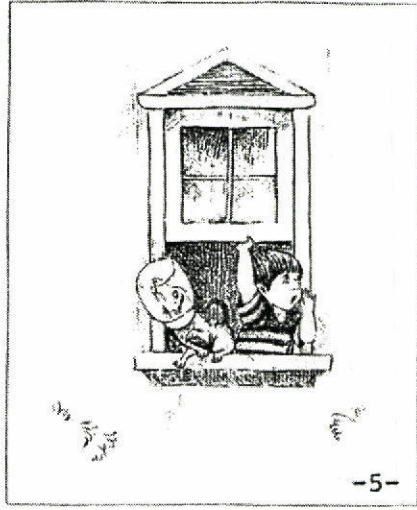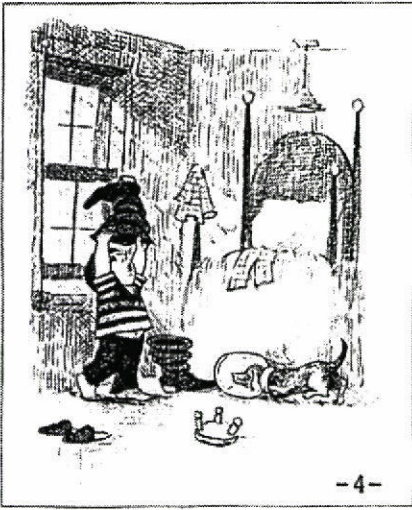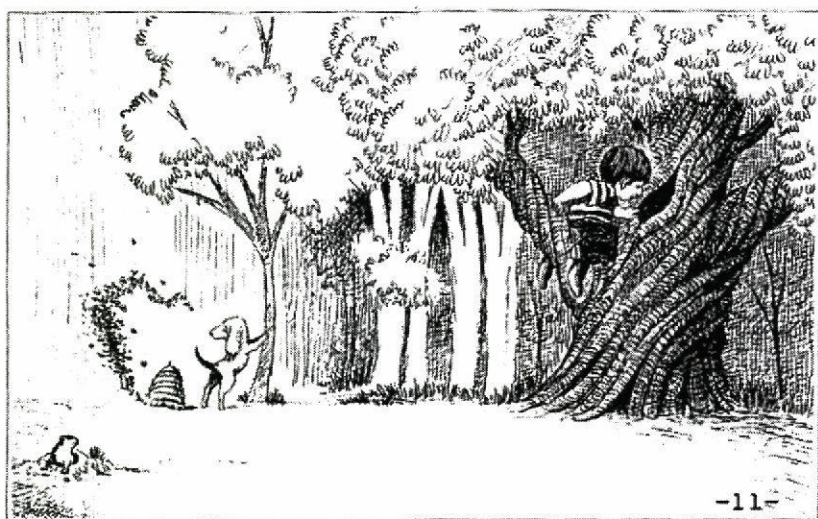
GOTHENBURG MONOGRAPHS IN LINGUISTICS
Doctoral dissertations from the Department of Linguistics
Göteborg University, Sweden

1. Lars-Gunnar Andersson, *Form and Function of Subordinate Clauses.* 1975

2. Jens Allwood, *Linguistic Communication as Action and Cooperation.* 1976

3: Pierre G. Javanaud, *The Vowel System of Lemosin - A Phonological Study.* 1981

4. Sven Strömqvist, *Make Believe through Words. A Linguistic Study of Children's Play with a Doll's House.* 1984

5. Elisabeth Ahlsén, *Discourse Patterns of Aphasia.* 1985

6. Sally Boyd, *Language Survival. A Study of Language Contact, Language Shift and Language Choice in Sweden.* 1985

7. Richard Hirsch, *Argumentation, Information and Interaction.* 1989

8. Sören Sjöström, *Spatial Relations - Towards a Theory of Spatial Verbs, Prepositions and Prenominal Adverbs in Swedish.* 1990

9. Anders Eriksson, *Aspects of Swedish Speech Rhythm.* 1991

10. Anders-Börje Andersson, *Second Language Learners' Acquisition of Grammatical Gender in Swedish.* 1992

11. Joakim Nivre, *Situations, Meaning and Communications. A Situation Theoretic Approach to Meaning in Language and Communication.* 1992

12. Kaarlo Voionmaa, *On the Semantics of Adult Verb Acquisition.* 1993

13. Mats Dahllöf, *On The Semantics of Propositional Attitude Reports.* 1995

14. Torbjörn Lager, *A Logical Approach to Computational Corpus Linguistics.* 1995

15. Karl Erland Gadelii, *Lesser Antillean French Creole and Universal Grammar.* 1997

16. Kerstin Nelfelt, *Simultaneous Sign and Speech: A Multimodal Perspective on the Communication of Hearing-Impaired Children.* 1998

17. Åsa Abelin, *Studies in Sound Symbolism.* 1999

18. Biljana Martinovska, *The Role of Repetitions and Reformulations in Court Proceedings - A Comparison of Sweden and Bulgaria.* 2000.

19. Åsa Nordqvist, *Speech about Speech: A Developmental Study on Form and Function of Direct and Indirect Speech.* 2001.