1 Introduction

> It is somewhat striking that so much research effort and interest has been focused on trying to understand how children come to learn the sounds, words, and syntax necessary to produce sentences in their native language, given that very little real language use is confined to the sentence level. (Pan & Snow 1999: 229)

Even though the situation has changed somewhat and a substantial body of research on phenomena above the sentence level is now available, the majority of these studies has been conducted on monolingual speakers and first language (L1) learners. Far fewer studies have targeted bilinguals and second language (L2) learners and still fewer have addressed learners in immersion (IM) programs.[1] By looking at the development of narrative discourse in an English immersion program at elementary school in Germany the present study seeks to contribute to filling this gap.

1.1 What is (narrative) discourse?

Discourse can be defined as the "use of language beyond a single sentence" (Bamberg & Moissinac 2003: 395). This definition encompasses not only written and oral mode but also a broad range of discourse types from *conversation* to more specific genres such as, for example, *narratives* (ibid.). *Discourse* used in this sense refers to the same phenomenon described by other authors as *text*.[2] Halliday and Hasan, for example, define text as "any passage, spoken or written, of whatever length, that does form a unified whole" (1976: 1). In the present study *discourse* and *text* will be used synonymously in the sense conveyed by both quotations. However, both terms will be employed as referring to *extended discourse* (Pan & Snow 1999), i.e. a sequence larger than just a couple of sentences.

Narrative as a particular discourse type includes subgenres such as (fairytale/ make-believe/ fictional) *stories* and *personal narratives* (e.g. Bamberg & Moissinac 2003, Sperry & Sperry 1996, Hicks 1991), i.e. narratives about personal experience. Consequently, *narrative discourse* can be defined as any spoken or written piece of extended discourse associated with the particular discourse genre *narrative*.[3,4]

---

1 I will be using the term *second language* as referring to any language learned in addition to the first one, regardless of context.

2 Yet other authors distinguish between *discourse* as a more dynamic and *text* as a static entity (e.g. Cutting 2002: 2, Johnstone 2002: 2, Hoey 1996, Clark 1994). Both views are yet again different from discourse in a Foucaultian tradition, where it is seen as a thematically-driven, superordinate communicative entity realized by a network of singular texts (e.g. Warnke 2008).

3 This definition of *narrative* is kept deliberately vague. For discussions of the ongoing debate about what characterizes narratives (as well as stories) see, for example, Her-

1.2 Why (narrative) discourse?

As the initial quote from Pan and Snow already indicates, discourse is an important part of human communication. Accordingly, discourse features are included in all influential models of language competence or performance (e.g. Bachman 1990) and in all major language assessment frameworks (e.g. Cambridge Certificate of Proficiency in English, TOEFL or the Common European Framework of Languages (Council of Europe 2001)), even if there is no general agreement on what exactly discourse competence encompasses and how to test it.

But what makes discourse special? The rules of grammar operate on a very local level, predominantly on clauses and sentences, and not usually across sentence borders. The rules of discourse production, on the other hand, operate (also) on larger stretches of spoken or written text. Since each rule system functions at a different level, local grammaticality is thus a priori unrelated to discourse requirements (cf. e.g. Givón 1995).[5] A series of sentences as in example (1.1) would more likely be accepted as discourse (here: a *story*), for example, than (1.2), even if the morphosyntax is target-like in (1.2) but not in (1.1):

(1.1) *Boy go school. Friend bad, take bike. Boy cry, no bike.*

(1.2) *The boy goes to school every day. He owns a bike. The boy's brother also has a bike.*

That is, discourse production—be it in an L1 or L2—may involve the formation of grammatical sentences, but more importantly it requires discourse-specific abilities. This ranges from the "social-cognitive sensitivity to communicative setting" (Berman 2008: 763), including familiarity with different genres, to command over the linguistic means of connecting stretches of speech or writing and the cognitive abilities to pre-plan for this linguistic as well as for a content-related, structural connectedness (Berman 2001).

In studying the (development of) discourse abilities of monolingual normal-developing, brain-damaged or language-impaired adults and children (e.g. Manolitsi & Botting 2011, Epstein & Phillips 2009, Reilly et al. 2004 & 1998, Norbury & Bishop 2003, Manhardt & Rescorla 2002, Berman & Slobin 1994, Reilly et al. 1998, Joanette & Brownell 1990), the study of narratives has proved especially useful, since, due to their strong socio-cultural importance (e.g. Bamberg & Moissinac 2003, Stein & Policastro 1984), narratives occur in conversation from a

man (2009), Renkema (2004: 191ff.), Richardson (2000), Boueke et al (1995) and Stein (1982). The present study follows Berman and Slobin's (1994) approach to defining narrative/story simply as the participants' productions in response to the picture-elicited storytelling task used for data collection (cf. ch. 4).

[4] In the following, the terms *narrative*, *narrative discourse* and *narrative text* will be used interchangeably.

[5] Even if the production of syntactically and morphologically target-like clauses and sentences facilitates understanding.

relatively early age (e.g. Nelson 1986, cf. also ch. 4).[6] Because of this, narratives are often considered the most important discourse genre (e.g. Bamberg & Moissinac 2003, Reilly et al. 1998).

Discourse competence deserves special attention in language acquisition studies, since it is a valuable asset in any educational context; comprehending and producing various discourse types is an important part of most curricula, be it "merely" oral and written stories or more specifically academic discourse types such as oral and written expository texts. Oral discourse skills in turn—especially oral narrative competence—have been found to be a significant predictor for literacy-related skills (e.g. Reese et al. 2010, Chang 2006, Griffin et al. 2004, Blankman et al. 2002, Dickinson & McCabe 2001, Snyder & Downey 1991, Norris & Bruning 1988).[7] Aspects of oral narrative competence can even be indicative for literacy-unrelated academic achievements, however: Fazio and colleagues (1996), for example, found that story retelling was the best single kindergarten predictor for the future academic status of their participants receiving academic remediation, while O'Neill and colleagues (2004) showed that the use of connectives is related to later mathematical achievement.

1.3 The present study: Goals and outline

The importance of (oral) discourse competence is, of course, not limited to monolingual education. On the contrary, immersion students and other L2 learners face an even greater challenge than monolinguals when asked to produce discourse in their L2, since even for young learners there may be a gap between cognitive and linguistic skills. At the same time immersion has been found to have an especially positive effect on participants' L2 conversational skills and willingness to communicate (e.g. Wode 2009: 38, Baker & MacIntyre 2003, Johnson & Swain 1997, Harley et al. 1990; cf. also Smit 2008, Lazaruk 2007, Genesee 1987). But what about the production of make-believe stories, which are considered a very challenging type of narrative discourse (cf. Berman 2004: 264ff.), since they require a largely autonomous construction of text? How does this type of discourse develop in an immersion program?

The present study investigates fictional adventure stories produced by 59 first and fourth graders (mean age 6;8 and 9;8) in an early partial immersion program

---

[6] As opposed to other discourse types, e.g. expository texts, which are introduced only later in formal schooling and whose development lags behind accordingly (Berman 2008, Berman & Verhoeven 2002: 18).

[7] This can be attributed to oral narratives' conceptual closeness to written discourse, i.e. their making use of many (linguistic) features otherwise associated with a written discourse style (Koch & Oesterreicher 1994). It should be kept in mind, however, that any discourse type's linguistic and content structure follows norms determined, in an educational context, by the respective (educated) majority culture and may therefore not apply to all parts of a population (cf. Gumperz et al. 1984, Scollon & Scollon 1984, Michaels & Collins 1984).

in the north of Germany, in which all subjects besides German language arts are taught in English. Learner variables collected were *grade* (first vs. fourth), *sex* (male vs. female), and *L2 preschool experience* (monolingual German vs. German-English bilingual group). Participants' stories were obtained through a picture-elicited oral storytelling task administered at the end of both school years. These stories were then analyzed in terms of two main discourse features: *Cohesion* and *coherence*, i.e. the linguistic connectedness of stretches of speech, for example via references or ellipses, and content connectedness through a global organization structure following an underlying narrative schema. It will be argued, furthermore, that these two measures represent aspects of participants' linguistic, and (respectively) cognitive development. In addition to investigating the development of cohesion and coherence from first to fourth grade this study explores differences attributable to participants' sex and/or preschool experience.

It should be emphasized that my study first and foremost gives a quantitative account of how coherence and cohesion develop and not a fine-grained analysis of developmental steps. At the same time, however, the detailed description of the categories of analysis in the methods chapter (ch. 5.2) presents the results of an initial in-depth qualitative analysis: Participants' texts were analyzed in detail as to the linguistic options they chose to realize the analysis categories provided by the two underlying frameworks (story grammar and Halliday and Hasan's (1976) approach to cohesive devices; cf. especially ch. 2 and 3). The qualitative results obtained were then quantified and the result of this latter step is described in the actual results section(s) of the present work.

In addition to contributing to the investigation of (the development of) narrative discourse produced by young L2 learners in an immersion setting, the present study also has a more concrete goal in relation to IM programs in general: Even though bilingual education is a very old phenomenon in Europe, over time monolingual L1 education came to be seen as "natural" in most European countries and it was only in the last third of the 20th century that this view started to change again (cf. Möller 2013 & 2009). Immense progress has been made especially in the last ten years, which shows in Germany, for example, in an increase in immersion and other bilingual education programs as well as in the introduction of at least some foreign language teaching in elementary school. Nevertheless, prejudices and reservations by parents and policymakers continue, as shown, for example, in recurring discussions on the importance of German as a national language (e.g. Spiegel Online 2008 & 2008a, Welt Online 2008). Therefore, the present study has a threefold aim:

1. To investigate how linguistic and content organization of (narrative) discourse develop over the four-year duration of an early partial immersion program.
2. To relate this development to participants' cognitive and linguistic development.
3. To relate the overall results to the effectiveness of the program.

These three goals will be pursued as follows: In chapter 2 the two frameworks used in the present study, story grammar (coherence) and Halliday and Hasan's

(1976) approach to cohesion, will be discussed critically in the light of the wide range of approaches to studying (narrative) discourse production and its development. It will be shown that these two approaches can make an important contribution to studying the development of narrative discourse—even in the light of more recent approaches—and that they are well suited for the purposes of the present study.

A simplified model of (narrative) discourse production will be outlined in chapter 3 in order to illustrate the challenges involved in telling a story from a picture book. This model leads up to a detailed description of the two fundamental dimensions of texts investigated in my study, namely coherence and cohesion. Coherence will be defined as a text's organization structure reflecting an underlying narrative schema. Cohesion, on the other hand, will be defined as the use of linguistic means to connect clauses and sentences into stretches of discourse. By continuity, coherence will be defined as a cognitive measure and cohesion as a linguistic measure. Both discourse measures will be described in detail and several research questions will be posed in relation to them.

In chapter 4 I will give an overview of the findings of previous studies with respect to narrative development. Studies will be presented on both monolingual L1 and, in a separate section, on L2 and bilingual learners. The overview of prior research will show that coherence and cohesion, as they were defined in chapter 3, can be studied fruitfully (a) within my participants' age range and (b) in L2 data. From the findings presented it will be argued that the analysis of L2 coherence allows insights into speakers' cognitive decelopment, since the studies presented show that the coherence of L2 productions and its development do not differ from those evident in L1 narratives once the speaker has acquired the necessary linguistic means in the L2. Chapter 4 concludes with several research hypotheses regarding the questions raised in the previous chapter.

The study's research design will be presented in chapter 5. Here, a short overview will be given of the immersion project in which my data was collected and then I will describe the participants and the data collection. Following this I will describe in detail the method of analysis for coherence (narrative components and their realizations) and cohesion (the subcategories of cohesive devices and their realizations) and will introduce the statistical methods that were applied. In the last section of this chapter I will address the dangers of the comparative fallacy with respect to my analysis.

In chapter 6 I will present the results obtained for narrative coherence with respect to the number of components produced and the individual narrative components identified in the task material. Then, complementing the latter analyses, the construction of a newly created index of global narrative structure will be described, and the results obtained with the help of this index will be given. Chapter 6 concludes with a summary of the coherence results.

The results obtained for narrative cohesion will be presented in chapter 7. First of all, I will give the results obtained for the texts' overall cohesive density and then those for the subcategories. After that the subcategories' degree of contribu-

tion to the overall cohesion of participants' stories will be described. Chapter 7 ends with a summary of the cohesion results.

Chapter 8 first of all details the results of a correlation analysis investigating the relationship between coherence and cohesion. In addition to this, the relationship between the development of these measures from first to fourth grade is explored.

In the last chapter, chapter 9, I will summarize and discuss the results of my study: I will address similarities and differences between coherence and cohesion results plus the relationship between the two measures. Then I will focus on two recurring themes of my study, namely the influence of the learner variables of grade, sex and L2 preschool experience and on participants' interindividual differences. In the light of my results I will then discuss the validity of coherence and cohesion as a cognitive and (respectively) linguistic measure as well as their usefulness for studying L2 data. After that I will address some general limitations of the present study. Chapter 9 ends with a discussion of conclusions regarding the effectiveness of the current program in particular and immersion teaching in general.