How do age, language, narrative task, language proficiency and exposure affect narrative macrostructure in German-Swedish bilingual children aged 4 to 6?

Josefin Lindgren and Ute Bohnacker
Uppsala University

Previous studies show mixed findings concerning whether higher-order story structure (macrostructure) is similar across bilinguals' two languages. It is not known how macrostructure is influenced by general language proficiency and amount of exposure. The present study investigates these issues in 46 German-Swedish bilingual 4- to 6-year-olds. Narratives were elicited in both languages with two picture-based tasks from the Multilingual Assessment Instrument for Narratives (MAIN) in the telling mode. We investigate to what extent the language of elicitation (Swedish vs German) influences bilingual children's macrostructure (story structure, episodic complexity) and explore effects of narrative task, age, narrative length, expressive vocabulary and estimated language exposure, both separately and combined, on macrostructure in the respective language. Results show that macrostructural skills developed measurably with age from 4 to 6 years in both languages, with no task effects. Story structure scores were higher in the majority language Swedish than in German and developed differently with age. The effect of narrative length on story structure was similar in the two languages. Language exposure did not have any significant effect. Macrostructure scores were significantly affected by expressive vocabulary in German only. Generally, the results may be linked to slightly higher language proficiency in Swedish.

Keywords: German, Multilingual Assessment Instrument for Narratives (MAIN), story structure, Swedish, vocabulary

https://doi.org/10.1075/lab.20020.lin | Published online: 28 January 2021
Linguistic Approaches to Bilingualism 12:4 (2022), pp. 479–508. ISSN 1879-9264 | E-ISSN 1879-9272

1. Introduction

Learning how to tell a story is not an easy task for children. Storytelling requires the child to use multiple linguistic and cognitive skills in concert. To tell an understandable story, a number of general language skills are needed, e.g. adequate vocabulary, syntactic structure, and morphology. The child needs to infer and communicate how story characters think and feel (see Burris & Brown, 2014), which requires Theory of Mind. Telling a story that is understandable for the listener also requires the child to structure the narrative events in a way that makes it clear what happened to whom and why. The story needs to suit the social and cultural context in which it is told. This means that narrative data from children can inform us about many different aspects of a child's linguistic, social and cognitive abilities, including how they structure complex discourse (Fiestas & Peña, 2004). It has also been proposed that narrative tasks constitute an ecologically valid method to assess children's language (e.g. Botting, 2002). As children grow older, oral narrative competence becomes increasingly important, for example for school achievements, as it is linked to literacy development and reading comprehension (e.g. Dickinson & Tabors, 2001; Griffin, Hemphill, Camp, & Wolf, 2004; Gutiérrez-Clellen, 2002). For all these reasons, it is important to study children's narratives.

Narratives are often divided into two levels, *macrostructure* and *microstructure* (e.g. Van Dongen & Westby, 1986). The macrostructure or *story structure* concerns the overall organization of narrative content (Stein & Glenn, 1979) and is thought to be largely independent of specific structures of the language. The *microstructure* comprises various linguistic structures that are language-specific, such as the use of adequate referential, temporal and causal linking devices (cohesive devices; Halliday & Hasan, 1976) or ritualized story openings and endings (e.g. *once upon a time, then they lived happily ever after*). Both levels need to be expressed adequately for a narrative to be understandable and well-formed. Macrostructure is foundational to storytelling though, as without it, there would be no story. In this paper, the focus is on macrostructure and its relationship and development across the two languages in bilingual children.

In recent years, it has become increasingly common to investigate children's language skills using narrative tasks. This growing interest in children's narratives has created a need for standardized narrative instruments in order to allow for comparisons across languages and groups. The present study employs the Multilingual Assessment Instrument for Narratives (MAIN; Gagarina et al., 2012, 2015, 2019), a picture-based narrative task battery suitable for children aged 4–10, which has been carefully constructed to allow for standardized assessment across

languages and groups of children, with a focus on macrostructure.[1] Several studies have used MAIN to investigate both mono- and bilingual children's production of narrative macrostructure (e.g. Bohnacker, 2016; Gagarina, 2016; Kunnari, Välimaa, & Laukkanen-Nevala, 2016; Lindgren, 2019), but data from more languages and language combinations are needed.

The present study investigates macrostructure in narratives elicited from German-Swedish bilinguals aged 4 to 6, a hitherto understudied group. Specifically, we investigate to what extent the language of elicitation (German, Swedish) influences bilingual children's production of macrostructure (story structure and episodic complexity), as it has been claimed that macrostructural skills are largely independent of language and transferrable between languages (see Section 2.2). We also explore potential effects of narrative task, age, narrative length, expressive vocabulary knowledge and estimated language exposure, separately and combined, on children's macrostructure in their respective language. These factors have rarely been investigated in bilingual children of any language combination.

2. Background

2.1 An overview of narrative macrostructure (story structure)

There are a number of different models of narrative macrostructure (e.g. Labov & Waletzky, 1967; Peterson & McCabe, 1983; Stein & Glenn, 1979). In models based on the so-called *story grammar* framework (e.g. Mandler & Johnson, 1977; Stein & Glenn, 1979; Stein & Policastro, 1984), which have often been used in studies of children's narratives (e.g. Fiestas & Peña, 2004; Trabasso & Nickels, 1992), a narrative has a *setting* (time and place of the events; introductions of story characters) and an *episodic system* (consisting of one or several episodes). The number and types of macrostructural components (e.g. goals, attempts, outcomes, reactions) in an episode varies somewhat between models, but, most importantly, the episode is seen as centered around the protagonist's goal (a *goal-based episode*).

The model on which MAIN, the narrative instrument used in the present study, is based (see Gagarina et al., 2012), was created from (goal-based) story grammar models in combination with Westby's (2005) decision-tree model of narrative structure (see below). The components of this model are shown in

1. MAIN was developed within the COST Action IS0804 *Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment* (2009–2013).

Table 1, and the picture-based stimuli of MAIN were specifically constructed to depict these components.[2]

Table 1. Macrostructural (story structure) components in the Multilingual Assessment Instrument for Narratives (MAIN; Gagarina et al., 2019)

Component	Description
Setting	Time and place of the events
Internal State as Initiating Event (IS as IE)	What does the character perceive/feel that sets the story events in motion?
Goal (G)	What does the character want?
Attempt (A)	What does the character do (in order to reach the goal)?
Outcome (O)	What is the result? What happens?
Internal State as Reaction (IS as R)	What does the character feel (in response to the outcome)?

In addition to measuring the extent to which story structure components are included in a child's narrative, the narrative's episodic complexity can be assessed. In Westby's (2005, p. 181) decision tree model (see also Gagarina et al., 2012, pp. 11–12), narrative episodes are classified as one out of the following:

- action/reaction sequence (AO)
- abbreviated episode (GA, GO)
- complete episode (GAO)
- no sequence

Action/reaction sequences (attempt-outcome sequences, AO) only contain an attempt and its outcome, as in (1). The more complex *abbreviated episodes* contain a goal but lack either the outcome (goal-attempt sequences, GA), as in (2), or the attempt (goal-outcome sequences, GO), as in (3). Finally, *complete episodes* (goal-attempt-outcome sequences, GAO), shown in (4), indicate a high level of episodic complexity (cf. Gagarina et al., 2012, p. 11, 2015; Stein & Glenn, 1979; Westby, 2005).[3] Episodes that do not contain at least two of these components

2. When eliciting a narrative, story content, including macrostructural components, can be presented visually in the stimulus materials, e.g. shown in pictures (as in the present study) or in video-clips (see Bishop, 1997, Chapter 7; Bishop & Adams, 1992; Lynch et al., 2008; Paris & Paris, 2003; van den Broek et al., 2005).

3. The examples in (1) to (4) come from the data of the present study. For each example, the child's code name, age, name of the story told and the language is given. The code name shows

do not contain any sequences of components, i.e. they can be classified as *no sequence*.

(1) **AO** (BiGer4–06, 4;5, Dog, German)
Den Hund jagt die Maus [A]. *Und der Hund stoßt sich an* [O].
'The dog chases the mouse. And the dog hits himself.'

(2) **GA** (BiGer4–09, 4;10, Dog, Swedish)
*Tror att hunden vill fånga muse*n [G]. *Hunden försöker att ta sig in där var musen är* [A].
'[I] think that the dog wants to catch the mouse. The dog tries to come in there where the mouse is.'

(3) **GO** (BiGer5–09, 5;4, Dog, Swedish)
*Sen kom hunden snabbt och ville ha alla korv*ar [G]. *Sen åt hunden upp all korv* [O].
'Then the dog came quickly and wanted to have all the sausages. Then the dog ate all the sausage.'

(4) **GAO** (BiGer5–02, 5;11, Dog, German)
Ein Hund will eine Maus fangen [G], *und rennt hinter der Maus her* [A], *und die Maus rennt in ein Loch im Baumstamm. Der Hund stößt sich den Kopf am Baum* [O].
'A dog wants to catch a mouse, and runs after the mouse, and the mouse runs into a hole in the tree trunk. The dog hits his head against the tree.'

In the present study, the children's narratives are scored for story structure following the components of the model presented in Table 1, and their episodic complexity is investigated using the classification described above (see also Section 4.4).

2.2 Factors influencing children's production of story structure

A number of different factors that are thought to influence children's story structure have been investigated, for example age, language, vocabulary knowledge, and narrative length; these factors will also be investigated in the present study. Existing studies of children's narratives differ substantially in the methods used to elicit narratives (e.g. telling vs retelling, supported by pictures or not), and what is included in measures of macrostructure also varies. This makes it somewhat difficult to compare the results. However, two findings for narrative macrostructure

---

the age group, so that BiGer**4**-06 is a bilingual German-Swedish **4**-year-old. For a description of the materials, see Section 4.2.

are consistent across studies. First, *age* is an important factor in explaining children's production of narrative macrostructure: a steep increase in development takes place between age 3–4 and 6–7 (e.g. Berman & Slobin, 1994; Bohnacker, 2016; Lindgren, 2019; Pearson, 2002; Schneider, Hayward, & Dubé, 2006; Trabasso & Nickels, 1992; Trabasso & Rodkin, 1994; Trabasso, Stein, Rodkin, Munger, & Baughn, 1992). After age 4, children increasingly start to link events together temporally and causally to form episodic sequences and begin to include goals in their narratives. Older children produce more complex episodic structure, including complete episodes. However, Trabasso et al. (1992) found that although goals started to appear in children's narratives at age 5, proportions of goals similar to those of adults were not produced until age 9.

A second finding regarding the acquisition of narrative macrostructure that is relatively consistent across studies is the following. Bilinguals tend to perform at a similar level in their two *languages* and performance in the two languages is often correlated (e.g. Bohnacker, 2016; Fiestas & Peña, 2004; Iluz-Cohen & Walters, 2012; Pearson, 2002; Uccelli & Páez, 2007). For example, Akinci, Jisa and Kern (2001) found that Turkish-French bilinguals ($N = 42$, ages 5, 7, and 10) produced equally complex Frog-story narratives in their two languages. Similar results were reported for the same type of narratives by 12 Spanish-English bilingual children aged 4;0–6;11 in Fiestas & Peña (2004). As most studies found no significant difference between bilinguals' languages, this led researchers to put forward the hypothesis that story structure/narrative macrostructure is language independent (e.g. Bohnacker, 2016; Iluz-Cohen & Walters, 2012; Paradis, Genesee, & Crago, 2011; Pearson, 2002). The present study tests this hypothesis.

Relatively little is known about narrative development in Swedish-speaking children, especially bilinguals. Apart from the studies using MAIN (see below), there are a number of other studies of narratives by Swedish-speaking children (Nordqvist, 2001; Reuterskiöld, Hansson, & Sahlén, 2011; Reuterskiöld Wagner, Sahlén, & Nettelbladt, 1999; Strömqvist & Day, 1993; Viberg, 2001), but only one, Nordqvist (2001), analyzed narrative macrostructure. Following the seminal study by Berman and Slobin (1994), Nordqvist (2001) analyzed the production of the core plot components onset, unfolding, and resolution in narratives by Swedish monolingual children aged 3, 4, 5, 9, and 15 years as well as adults ($N = 84$), and found age development, especially between age 4 and 5, for all three components (Nordqvist, 2001, pp. 201–203).

A range of different languages and language combinations have been investigated using MAIN, e.g. Dutch monolinguals/bilinguals (Blom & Boerma, 2016; Boerma, Leseman, Timmermeister, Wijnen, & Blom, 2016), English-Swedish bilinguals (Bohnacker, 2016), Finnish monolinguals/Swedish-Finnish bilinguals (Kunnari et al., 2016), Polish monolinguals/Polish-English bilinguals

(Otwinowska, Mieszkowska, Białecka-Pikul, Opacki, & Haman, 2020), Russian-German bilinguals (Gagarina, 2016), Swedish monolinguals (Lindgren, 2019), Swedish monolinguals, German-Swedish and Turkish-Swedish bilinguals (Lindgren, 2018),[4] Turkish-Swedish bilinguals (Öztekin, 2019) and Turkish-German bilinguals (Maviş, Tunçer, & Gagarina, 2016). Previous studies using MAIN analyzed macrostructure with the story structure score described in the MAIN scoring protocol (see Section 4.4; Gagarina et al., 2019). Their results are mainly in line with those of studies using other stimuli, showing similarities in performance across bilinguals' languages and development from age 3/4–5 to 6–7. To give an example, Bohnacker (2016) found for 52 English-Swedish bilinguals that 6–7-year-olds performed better than 5-year-olds, irrespective of language. Kunnari et al. (2016), who studied Finnish monolinguals and Swedish-Finnish bilinguals growing up in Finland aged 5;0–6;7 ($N=32$), found a significant effect of age on the story structure score. The Swedish-Finnish bilinguals performed equally well in their two languages.

With respect to story structure development from around age 6 onwards, results suggest a slower rate of development. For example, while Gagarina (2016) found that Russian-German bilingual preschoolers (age 2;7–4;4) performed significantly lower than children in Grade 1 (age 6;5–7;5) and Grade 3 (age 7;11–10;6) in both languages, there were no differences between Gagarina's two older groups. Similar results have also been found in a longitudinal study of 17 Swedish monolingual children aged 4 to 7 (Lindgren, 2019). Lindgren (2019) found a steep increase in story structure scores from age 4;4 to age 5;10, but little further development until age 7;4, indicating that children may reach a plateau in their development of narrative macrostructure around age 6, at least for these stories. Similarly, in another longitudinal study, Blom and Boerma (2016) found no significant difference in story structure performance at age 5;9 and 6;9 in typically-developing Dutch monolinguals.

Some studies have also investigated episodic complexity. However, the specific measures used to analyze episodic complexity vary substantially. In some studies, the child's best performance was assessed (Gagarina, 2016; Roch, Florit, & Levorato, 2016), giving the child a complexity score based on the most complex episode produced in the narrative. Other studies compared episodes in which a goal was produced with those without (Kunnari et al., 2016), or gave each child a total complexity score based on all episodes (Maviş et al., 2016); alternatively, they analyzed proportions of different types of sequences (Lindgren, 2019), or investigated the frequency of GAO-sequences (Bohnacker, 2016). The common thread

4. Note that the German-Swedish bilinguals in Lindgren (2018) are the same as in the present study.

throughout the results is that there is development with age in macrostructural complexity (Bohnacker, 2016; Gagarina, 2016; Kunnari et al., 2016; Lindgren, 2019), but that children aged 4–7 produce relatively few episodes that contain goals, such as complete episodes (GAOs). For example, English-Swedish 5-year-olds were found to produce only 7% GAOs, and only slightly more at age 6–7 (12%) (Bohnacker, 2016) and Swedish monolinguals produced only 18% GAOs at age 7 (Lindgren, 2019).

With respect to differences in performance between different MAIN stories (see Section 4.2 for a description of the materials), with the exception of the first author's PhD thesis (Lindgren, 2018) and a related PhD thesis (Öztekin, 2019), both under the supervision of the second author, there are no studies comparing Cat/Dog and Baby Birds/Baby Goats for the production of macrostructure. However, Bohnacker and Lindgren (in press) found for Swedish monolinguals aged 4–6 that *narrative comprehension* was higher for Cat/Dog than for Baby Birds/Baby Goats. Lindgren and Bohnacker (2020a) and Bohnacker, Lindgren and Öztekin (2020) found the same for German-Swedish and Turkish-Swedish bilinguals respectively, indicating that the Cat/Dog stories and the Baby Birds/Baby Goats stories may not quite have the same level of difficulty in comprehension. Whether this is also the case for narrative production, which is investigated in the present study, remains an open question. Generally, differences between different stories (i.e. stimulus materials) have not been investigated to any larger extent.

A number of studies have found significant relationships between, on the one hand, children's *vocabulary knowledge* or the *length of the narrative* (linguistic productivity), and, on the other, story structure or a similar measure (e.g. Bitetti & Hammer, 2016; Mäkinen, 2014; Pearson, 2002; Uccelli & Páez, 2007). For example, Mäkinen (2014), in a study of Finnish monolinguals aged 4–8, found a significant effect of story length (total number of words, TNW) on the amount of event content (i.e. story information) expressed in the narrative. Note, however, that amount of event content is not the same as story *structure*, and it may well be the case that Mäkinen's (2014) measure is more sensitive to story length as it concerns the amount of story information. In their longitudinal study of narratives by Spanish-English bilinguals ($N=24$) at age 5–6 and 6–7, Uccelli and Páez (2007) found significant correlations between both narrative length and the scores on an expressive vocabulary test and story structure in both languages. Pearson (2002), in a study of Frog story narratives by Spanish-English bilinguals, found moderate correlations between measures of the child's vocabulary knowledge and the story score. Using the Narrative Scoring Scheme (NSS), Heilmann, Miller, Nockerts, and Dunaway (2010) found a correlation between narrative vocabulary and their NSS score, which contains macrostructural aspects. In a longitudinal study of 81

Spanish-English bilinguals, Bitetti and Hammer (2016) found an effect of receptive vocabulary (PPVT) on the children's NSS scores. However, there are also a number of studies where no clear relationship was found between story structure and various microstructural measures in the narratives (e.g. Iluz-Cohen & Walters, 2012; Rodina, 2017), which could be due to the relatively small sample sizes in these studies.

Some studies have analyzed the effects of narrative exposure, e.g. joint book reading activities, on children's narratives (e.g. Bitetti & Hammer, 2016). The effect of the amount of language input on story structure in bilinguals, which is investigated in the present study, has to our knowledge only been investigated in one previous study, namely Haman, Wodniecka et al. (2017). They investigated the effects of L1 and L2 cumulative language exposure on L1 language skills, including on story structure in 53 Polish-English bilingual 4- to 7-year-olds (unusually however, Haman, Wodniecka et al. (2017) averaged the scores of two MAIN narratives in different elicitation modes, telling and retelling). They found an effect of cumulative language exposure in both languages on story structure in Polish; these results seem to indicate that increased language exposure in general leads to better narrative skills.

3. Aims, research questions and predictions

The present study investigates oral narratives in both languages of 4- to 6-year-old German-Swedish bilinguals. The following research questions are asked:

- Do language (Swedish vs German) and narrative task (MAIN Cat/Dog vs Baby Birds/Baby Goats) influence story structure and episodic complexity?
- How does age affect story structure and episodic complexity?
- How do vocabulary knowledge and language exposure/input relate to story structure?
- Is there any effect of narrative length (linguistic productivity) on story structure?

As earlier studies using MAIN have found that the production of story structure develops during the preschool years, from age 4 to 7, and that bilingual children score similarly in their two languages (e.g. Bohnacker 2016; Kunnari et al., 2016), we predict that the same will be the case for the German-Swedish bilinguals in the present study. As the narrative tasks were constructed to be parallel in terms of story structure (Gagarina et al., 2012), we expect the children to perform similarly irrespective of task. Based on previous studies (Uccelli & Páez, 2007; Pearson, 2002), we expect an effect of vocabulary knowledge and narrative length on the

story structure. Since only one study (Haman, Wodniecka et al., 2017) has investigated the effects of input on story structure, it is difficult to make predictions with respect to this factor.

4. Methods

4.1 Participants

The participants were 46 German-Swedish bilinguals (31 girls, 15 boys) aged 4;0–6;11 ($M = 5;7$, $SD = 0;11$) growing up in Sweden. All parents signed a consent form and filled in a detailed questionnaire about the child's language development, language use in the family, and the language skills and educational backgrounds of the parents. The information presented below is based on data from the questionnaires.

Most children (74%, 34 children) had lived in Sweden from birth, and they all came from mid- to high-SES backgrounds; all parents had at least some tertiary education, which is typical for German-speaking adults in Sweden. No child had been diagnosed with language impairment. All children spoke both Swedish and German well enough to complete the narrative tasks in each language. They had been exposed to German (the minority language) in the home from birth. Most children (69.6%, 32 children) received regular Swedish input since birth or before age 1;0; only four children (8.7%) had started to receive regular exposure to Swedish (the majority language) after age 2. Around half the children (25 children, 54.3%) had one German native speaker parent and one Swedish native speaker parent. Most of the remaining children (18 children, 39.1%) had two German native speaker parents. Two children had one native German parent and one parent with another native language, and one had parents who were both German-Swedish bilinguals.

Around a third of the children (14 children, 30.4%) attended a Swedish-German bilingual (pre)school and the other children (32 children, 69.6%) attended regular Swedish-medium (pre)schools. The children attended (pre)school (or an after-school program organized by the school) for a major part of the day, something that is typical in Sweden. The younger children (aged 4–5) attended preschools, and most of the older children (aged 6) attended *förskoleklass* 'preschool class', a preparatory year before Grade 1 of primary school. Only a few of the oldest children were in Grade 1 where they had started to learn to read and write.

The parents were asked to estimate their child's daily input in the two languages on a scale ranging from 95% German/5% Swedish to 5% German/95%

Swedish, with the additional option of writing down their own distribution. Based on this estimate, the children on average received two thirds (64.1%) daily input in Swedish and one third (34.5%) daily input in German.[5] No child received more than 60% German daily input. The higher proportion of Swedish input was mirrored in how the children's expressive language skills were rated by the parents: Close to half of the children (48%) were rated to have somewhat higher expressive language skills in Swedish, whereas only a smaller group (17%) were seen as more proficient in German. The other children (35%) were rated to be equally proficient in both languages.[6]

4.2 Materials

Oral narratives were elicited from the children in German and Swedish using picture-based narrative tasks from the Multilingual Assessment Instrument for Narratives (MAIN; Gagarina et al., 2019). Each of the four MAIN stories (Cat, Dog, Baby Birds, Baby Goats) consists of six coloured pictures showing a story with three episodes. Each episode contains a goal-attempt-outcome sequence (see Gagarina et al., 2012). The four stories are parallel in length and in story grammar components, and two each are strictly parallel with respect to their plotlines (Cat/Dog; Baby Birds/Baby Goats). There is also a difference in the number of characters between the stories: while Cat/Dog has three characters, Baby Birds/Baby Goats has five. In this paper, Cat/Dog is referred to as one narrative task and Baby Birds/Baby Goats as the other. In addition to the picture sequences and a standardized procedure for administering the tasks, MAIN also contains a scheme for scoring narrative macrostructure (see Section 4.4) and ten comprehension questions.

To get a measure of the children's vocabulary knowledge, we administered the German and Swedish versions of the Cross-linguistic Lexical Tasks (LITMUS-CLT, hereafter CLT; Haman, Łuniewska, & Pomiechowska, 2015), CLT-DE (Rinker & Gagarina, 2014) and CLT-SWE (Ringblom, Håkansson, & Lindgren, 2014). The CLTs are suitable for preschool children and are the only existing vocabulary tasks that are comparable for Swedish and German. Each CLT language version has four different parts: noun comprehension, verb comprehension, noun production, and verb production. As the children's scores on the receptive (comprehension) parts were close to ceiling in both languages, it was

5. The averages for the two languages do not sum up to 100% because four children also received input in a third language. The input in this third language (English, Chinese) was always the lowest.

6. For more information about the participants, see Lindgren (2018, Chapter 3).

decided to only include the scores on the expressive (production) parts in the present study.[7] In the CLT expressive parts, which are picture naming tasks, the child is asked to name an object ('What is this?' for nouns) or an action ('What is she doing?' for verbs). In each language, the maximum expressive vocabulary score is 60 points.

4.3 Procedure

The children were tested in German and Swedish in two separate sessions on different days. The average time between sessions was eight days; in all cases, it was at least three days. Half of the children were tested in German first, and the other half in Swedish first. The sessions, which were video and audio recorded in their entirety, took place in a quiet room at their (pre)schools, or in a few cases at home or at the university. The experimenters were native speakers of the language of testing (for Swedish: the first author; for German: either a female research assistant or the second author). They only spoke to the child in the language of the testing and acted as if they had no knowledge of the other language. The children received stickers after each task and a diploma upon completion of their participation in the study.

In each language, each child told two different MAIN stories and did the vocabulary task (CLT). After a short warming-up phase in which the experimenter asked the child some questions (e.g. about favorite things or spare time activities), the tasks were administered in the following order: MAIN Cat/Dog – CLT – MAIN Baby Birds/Baby Goats. No child received the same story twice, and stories were counter-balanced across languages.[8] The order of the CLT parts were counterbalanced following Haman, Łuniewska et al. (2015).

The CLTs were administered as paper-and-pencil tasks following standardized procedure (Haman, Łuniewska et al., 2015). The MAIN narrative tasks were administered using the 'telling mode' procedure as described in Gagarina et al. (2019). This procedure is as follows. The child is presented with three envelopes and picks one, takes out the picture sequence inside and looks at all the pictures, after which the pictures are folded back so that only the first two are visible and the child begins telling the story. When the child finishes telling about the first two

7. For a study investigating the same children's receptive and expressive CLT scores in both languages, see Lindgren and Bohnacker (2020b).

8. Counter-balancing was done so that for the Cat/Dog task, half of the children told Cat in Swedish and Dog in German and the other half Dog in Swedish and Cat in German; similarly, half of the children told Baby Birds in Swedish and Baby Goats in German and vice versa. For further details of the counter-balancing system used, see Lindgren (2018, Chapter 3).

pictures, the next two pictures and finally all pictures are unfolded. At no point before or during the story telling are the pictures visible to the experimenter; there is no shared visual attention. After having told the story, the child answers the accompanying comprehension questions. Results for the comprehension questions are not included in the present study, but have been reported elsewhere (Lindgren & Bohnacker, 2020a).

4.4 Coding and analysis

All narratives were transcribed orthographically in CHAT-format (MacWhinney, 2000) by the first author (Swedish) and a female research assistant (German) and were carefully checked. For one child, data from the Swedish Cat/Dog task had to be discarded as this child switched to German in the middle of her narrative and did not switch back to Swedish (despite efforts from the experimenter).

All narratives were scored for story structure following the MAIN scoring protocol (Gagarina et al., 2019). The MAIN *story structure score* is a sum of episodic (macrostructural) components and assesses the child's ability to verbally realize these components. The story structure score includes setting (maximum 2 points, 1 each for time and place), and one point for each of five macrostructural components in the three episodes: internal state as initiating event, goal (G), attempt (A), outcome (O), internal state as reaction (3 × 5 components). The maximum score is 17 points. Table 2 gives an overview of setting and episode 1 for the Baby Goats story.

Table 2. Macrostructural components in Baby Goats, setting and episode 1 (constructed examples)

Component	Baby goats
Setting	Once upon a time… *(time)*
	…by a lake *(place)*
Episode 1	
IS as IE	The mother goat saw that the baby goat was in danger
Goal	She wanted to rescue it
Attempt	She ran down into the water…
Outcome	…and pushed the baby out of the water
IS as R	The mother goat was happy

Note. IS = internal state, IE = initiating event, R = reaction.

The first author scored all narratives, following the MAIN manual (Gagarina et al. 2012, 2019). As the data varied in how the narrative content was expressed, and the manual did not provide a scoring decision for all cases, the second author and her research group (including the first author) developed additional scoring guidelines. The guidelines, which included both general scoring principles as well as decisions on a large number of unclear cases, were based on extensive and systematic discussions of unclear cases from the narratives of 286 mono- and bilingual children. The first author checked the scoring thoroughly for consistency against these guidelines. All unclear cases were discussed with the second author. Three principles were central for scoring. First, only those words produced by the child that were comprehensible in the target language were considered in the scoring. Second, the child's verbal realization of the action or emotion had to be clear enough; non-verbal cues (e.g. pointing, iconic gestures) were not considered. (Recall that the child tells the story without the listener being able to see the pictures.) An action/emotion was considered clear enough when it was understandable and could be linked to a specific character and point in the story. The production of relatively specific verbs and the inclusion of appropriate agents/patients were thus necessary to score a point.[9] Third, we used a principle of lenient scoring of reference. This was done to prevent that the (lack of) ability to correctly refer to story characters/objects would influence the child's score. Each child thus received a *story structure score* (Max = 17) for each narrative task in German and Swedish (four scores per child).

Next, based on the different components of the story structure scoring, as a measure of *episodic complexity*, each episode (i.e. 3 episodes per narrative) was coded as one of the following types of sequences of macrostructural components: *GAO-sequence* (complete episode, G+A+O are produced in the episode), *GA/GO-sequence* (incomplete episode, either G+A or G+O are produced in the episode), or *AO-sequence* (action/reaction sequence, A+O are produced in the episode). Episodes that did not contain any of these combinations of components were coded as *no sequence*.

The following measures were included as independent variables (predictors) in the analyses. The child received a CLT *expressive vocabulary score* (Max = 60) in German and another one in Swedish. The *length of each narrative* in total number of words (TNW) was calculated in CLAN (MacWhinney, 2000) using the *freq*

9. To give an example, in the Baby Goats story, *he takes* (zero points) cannot be considered clear enough, as it is unclear who takes whom and there are multiple points in the story this could refer to, whereas *the fox catches him* can be clearly linked to a specific point in the story and scores a point for outcome (in Episode 2).

function. Only words in the target language were included in this calculation.[10] Hesitation markers (e.g. *uhm*), repetitions and false starts were marked in the transcripts (using the CHAT-conventions) and excluded from the total number of words. The values for the child's *estimated percentage of daily input* in German and Swedish (on a scale from 0 to 100) were based on information in the parental questionnaire (see Section 4.1 for details). Finally, the child's *age in months* (range: 48–83 months) as well as the *language of the first testing* (German or Swedish) were included in the analyses.

To investigate the effects of language (German vs Swedish), narrative task (Cat/Dog vs Baby Birds/Baby Goats), age, narrative length, vocabulary knowledge (CLT production scores), and language input on the children's story structure scores and episodic complexity, we carried out a series of linear and logit mixed effect models in R using the functions lmer and glmer of the lme4 package. For details of the specific models, see Section 5.

5. Results

5.1 Story structure score

In Table 3, results for the story structure score are shown for the two tasks in each language, with somewhat lower mean scores in German than in Swedish. On average, the German-Swedish bilinguals produce around one third of the maximal number of story structure components (Max = 17) in both languages. However, the range is relatively large; the children with the lowest scores receive only 0–3 points and the children with the highest scores 10–12 points.

Table 3. Story structure scores (Max = 17)

	German		Swedish	
	Cat/Dog (*N* = 46)	*BB/BG* (*N* = 46)	*Cat/Dog* (*N* = 45)	*BB/BG* (*N* = 46)
Mean (SD)	5.3 (2.5)	5.4 (2.6)	6.4 (2.0)	6.6 (2.4)
Range	0–10	1–12	3–11	1–11

Note. BB = Baby Birds, BG = Baby Goats.

10. Some children produced words in the other language in their narratives. This was more common in the German narratives (2.3% of the words), which was more often the children's weaker language (see Section 4.1), than in the Swedish narratives (0.4% of the words). These words were not included in the narrative length measure. For more information about code-switching in the data, see Lindgren (2018, Chapter 4).

As expected from the mean scores (Table 3), a linear mixed-effects model (with a random effect for Child)[11] show a significant effect of Language ($B = 1.26$, $SE = 0.36$, $p < .001$), but no effect of Task ($B = -0.07$, $SE = 0.36$, $p = .86$) and no significant interaction effect ($B = -0.19$, $SE = 0.51$, $p = .71$). The children thus score significantly higher in Swedish than in German, but there is no difference in story structure scores between the two tasks Cat/Dog and Baby Birds/Baby Goats.

5.2 Narrative length and vocabulary knowledge

Before investigating the effects of narrative length (in total number of words, TNW) and vocabulary knowledge on the story structure score, we report results for these variables separately. Table 4 shows the results for narrative length for the two narrative tasks in German and Swedish.

Table 4. Narrative length (TNW)

	German		Swedish	
	Cat/Dog	*BB/BG*	*Cat/Dog*	*BB/BG*
Mean (SD)	97.8 (48.9)	74.7 (33.7)	76.0 (26.3)	68.5 (26.3)
Range	31–213	30–201	26–138	26–151

Note. BB = Baby Birds, BG = Baby Goats.

A linear mixed-effects model (with a random effect for Child) show a significant effect of Language ($B = -21.93$, $SE = 5.70$, $p < .001$), with longer narratives in German, and of Task ($B = -23.09$, $SE = 5.66$, $p < .001$), with longer narratives for Cat/Dog, and a borderline significant interaction effect ($B = 15.71$, $SE = 8.03$, $p = .05$), indicating that the difference in length between the tasks is somewhat larger for German, and that the difference in length between the languages is larger for Cat/Dog than for Baby Birds/Baby Goats. It is also worth noticing that the ranges are wider in German; whereas the shortest narratives were of similar length in the two languages, the longest ones were substantially longer in German.

In Table 5, the children's expressive vocabulary scores (CLT) in German and Swedish are reported. The linear mixed effects model (with a random effect for Child) show a significant effect of Language ($B = 5.11$, $SE = 1.72$, $p < .01$); scores in Swedish are significantly higher than in German. The range is also larger in German than in Swedish; whereas the highest scores are comparable, the lowest scores are considerably lower in German. This indicates that there is a larger

11. Note that adding random slopes to the models was not possible, since no models containing random slopes converged.

spread in the children's vocabulary knowledge in the minority language German than in the majority language Swedish.[12]

Table 5. Expressive vocabulary scores (CLT) (Max = 60)

	German	Swedish
Mean (SD)	40.1 (11.0)	45.2 (6.39)
Range	17–56	28–57

5.3 Effects of age, narrative length, expressive vocabulary knowledge and estimated daily input on the story structure score

Next, we investigate the effects of age, narrative length (TNW), the child's expressive vocabulary score (CLT) and the estimated proportion of daily language input on the story structure score in each of the two languages. Before reporting the results of our multivariate analyses, we describe the relationship between each of the factors and the story structure score separately in the form of a series of scatterplots that also show regression lines and R^2-values, which indicate the importance of the factor in explaining the variation in the story structure score.[13]

Figure 1 shows the relationship between the child's age and the story structure score. It is clear that this relationship is much stronger in Swedish than in German (as shown by the size of their respective R^2-values). In German, the scores are more scattered, and there are a number of older children who score relatively low, even below the youngest children, whereas in Swedish, the lowest-scoring children are all among the youngest.

In Figure 2, the relationship between narrative length (TNW) and the story structure score is shown. This relationship is of comparable strength in the two languages. However, as was also mentioned in Section 5.2, there is more variation in narrative length in German. There are relatively many narratives that are very short (50 words or less), but also a fair number that are long (150 words or longer). In Swedish, the narratives are more homogenous with respect to length. In both languages, there is more variation in the story structure scores of the short narratives, and some of the shorter narratives receive relatively high scores, indicating that a narrative does not need to be very long in order to receive an average story

12. For a detailed analysis of the children's vocabulary scores (both receptive and expressive) on the Swedish and German CLTs, see Lindgren and Bohnacker (2020b).

13. The scatterplots are based on two story structure values (Cat/Dog and BB/BG) for each child in each language.

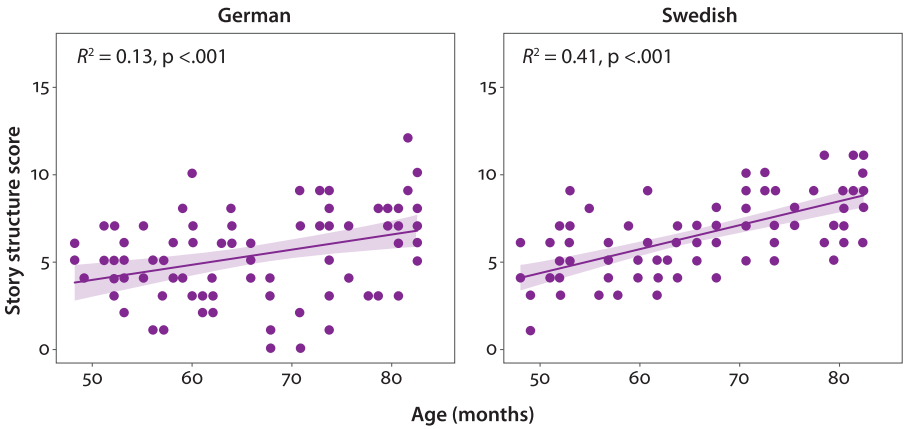

Figure 1. The relationship between the child's age in months and the story structure score (Max = 17)

structure score or higher. However, in both languages, the longest narratives are among those receiving the highest scores.

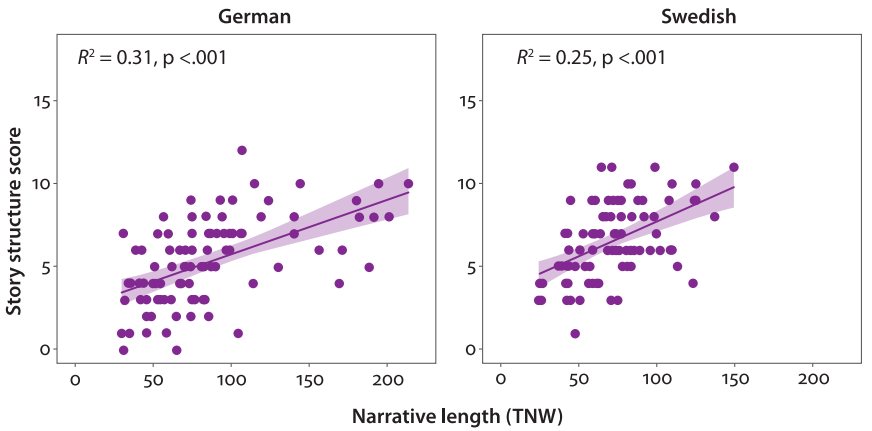

Figure 2. The relationship between narrative length (TNW) and the story structure score (Max = 17)

In Figure 3, we see the relationship between the children's expressive vocabulary scores (CLT) and the story structure score. A number of observations can be made. First, as already shown in Section 5.2, there is more individual variation in the German vocabulary scores, and the lowest scores are substantially lower in German than in Swedish. Second, there is a much stronger relationship between expressive vocabulary knowledge and story structure in German than in Swedish

(as seen by the size of the R^2-values). In German, there is a clear tendency for children with higher vocabulary scores to also score higher on story structure; in Swedish, this is the case to a lesser degree. In Swedish, the variation in story structure scores between children with similar vocabulary scores is larger.

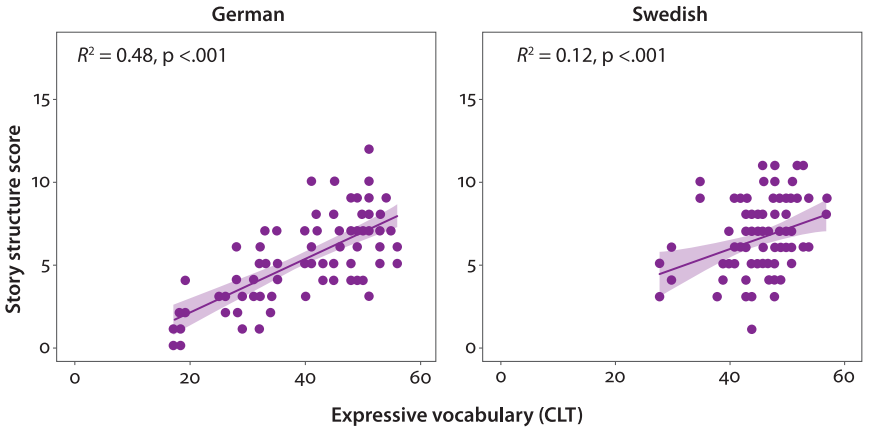

Figure 3. The relationship between expressive vocabulary score (CLT; Max = 60) and story structure score (Max = 17)

Finally, Figure 4 shows the relationship between the estimated daily input (in %) and the story structure score for the two languages. Here, we also see a difference between the languages. There is no significant relationship between estimated daily input in Swedish and the Swedish story structure, whereas the relationship is significant, but not very strong in German. Recall that the children had more exposure to the majority language Swedish. In such a setting, more exposure to the minority language German may make a measurable difference for the children's performance. However, in both languages, there is substantial variation in story structure scores between children receiving the same amount of daily input in that language.

After having investigated the relationships between each of the factors age, narrative length (TNW), expressive vocabulary (CLT), and input and the story structure score, we carried out multivariate analyses to investigate the effect of all variables simultaneously. Table 6 shows the results for the linear mixed-effect models (with a random effect for Child) for Swedish and German. In these models, we also control for the language of the first testing (Test 1).

In Swedish, there are significant effects of the child's age and narrative length (TNW); children who are older or tell longer narratives have higher Swedish story structure scores. In German, there are significant effects of the child's age,

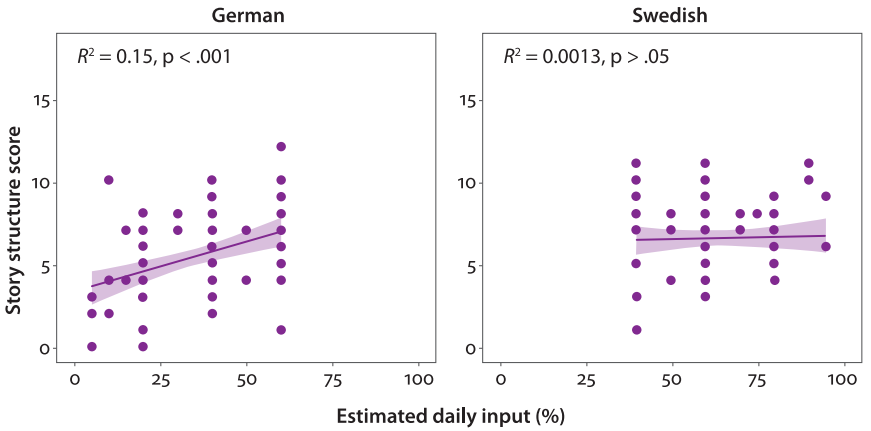

Figure 4. The relationship between the child's estimated daily input and the story structure score (Max = 17)

Table 6. Summary of linear mixed-effect models for macrostructure production scores of the narrative tasks in Swedish and German

	German		Swedish	
	B	*SE*	B	*SE*
Intercept	−4.17**	1.32	−2.94	1.75
Age (months)	0.04*	0.02	0.11***	0.02
Narrative length	0.02**	0.01	0.02**	0.01
Vocabulary	0.12***	0.03	0.02	0.04
Input	0.01	0.02	0.01	0.01
Test 1	0.87*	0.40	−0.52	0.41

Note. *$p < .05$; **$p < .01$; ***$p < .001$. Narrative length = total number of words (TNW) in the target language in the narrative; Vocabulary = expressive CLT score in the language of the narrative task; Input = estimated proportion of daily input in the language of the narrative task; Test 1 = language of the first testing; the model shows the effect when the language of the first testing is Swedish.

narrative length (TNW), the language of the first testing, and expressive vocabulary (CLT); children who are older or tell longer narratives, or are tested in Swedish first, or have higher German vocabulary scores have higher German story structure scores. Note that the age effect is stronger in Swedish than in German, as seen by the size of the coefficients, whereas the effect of narrative length is of the same size.

5.4 Episodic complexity

Finally, we investigated episodic complexity in the form of different types of macrostructural sequences, i.e. types of episodic structures. Figure 5 shows the percentages of different types of episodic structures (out of all possible episodes in the narratives) produced in each of the two languages for the two narrative tasks separately.

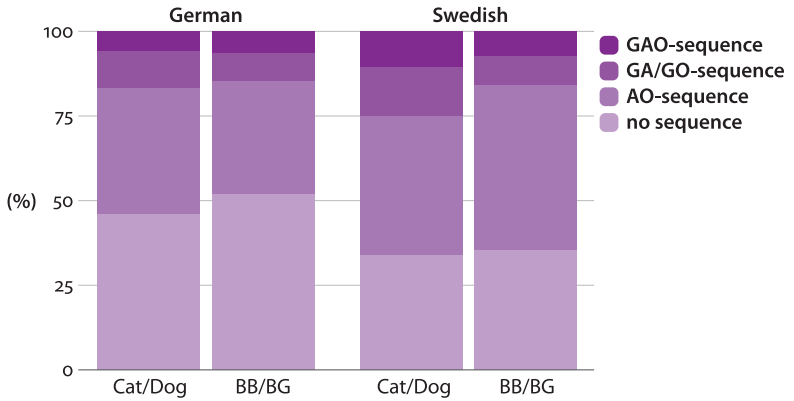

Figure 5. Types of macrostructural sequences, by language and narrative task ($N = 549$)

In Figure 5, we see that in both languages and for both narrative tasks, the most common types are AO-sequences (33%–49%) and occurrences of no sequence (34%–52%). Sequences containing goals, i.e. GA/GO-sequences (abbreviated episodes) and GAO-sequences (complete episodes), were much less frequently produced (GA/GO-sequences: 8%–15%, GAO-sequences: 6%–10%). The overall episodic complexity is thus relatively low; only rarely do these children reach the higher level of episodic complexity in their narratives, as goals are seldom produced.

In the following statistical analyses, in which we investigate effects of language and narrative task, we focus on the production of *no sequences* and *GAO-sequences*, i.e. on the levels with the lowest and highest complexity, as these are cases when the children fail to produce any type of sequence of macrostructural components within an episode or produce complete episodic structures, respectively. We also include age in the analyses.

The logit mixed-effects model (with a random effect for Child) for the production of *no sequences* shows a significant effect of Language ($B = -0.68$, $SE = 0.19$, $p < .001$), with a lower proportion of episodes with no sequence in Swedish than in German, and of Age ($B = -0.05$, $SE = 0.01$, $p < .001$); older children are less likely to produce an episode that do not contain any sequence. The effect

of Task is not significant ($B = 0.19$, $SE = 0.19$, $p = .32$). As the production of GAOs is generally low across languages and tasks, running a mixed effects model on these data was deemed unsuitable. Instead, we ran a logistic regression analysis (using the glm function in R). The model shows a significant effect of Age ($B = 0.06$, $SE = 0.02$, $p < .001$), with older children being more likely to produce a GAO, but no significant effects of Language ($B = 0.39$, $SE = 0.33$, $p = .24$) or Task ($B = -0.17$, $SE = 0.33$, $p = .60$).

6. Discussion

The present study investigated oral narratives (elicited with MAIN; Gagarina et al., 2019) in both languages of 46 German-Swedish bilinguals aged 4 to 6. We analysed effects of language, narrative task (Cat/Dog vs Baby Birds/Baby Goats) and age on story structure and episodic complexity, and effects of narrative length (TNW), children's expressive vocabulary knowledge (CLT production scores) and estimated daily input in the language on story structure. Below we summarize and discuss the results for each of these factors in turn.

With respect to differences between the *languages*, contrary to our prediction, we found significantly higher story structure scores in Swedish than in German, as well as a somewhat higher episodic complexity in Swedish with fewer episodes that contain no sequence of the core macrostructural components (goals, attempts and outcomes). However, there was no significant difference between the languages concerning the proportion of complete episodes (GAOs). These results indicate that, in contrast to findings from previous studies (e.g. Bohnacker, 2016; Fiestas & Peña, 2004; Paradis et al., 2011; Pearson, 2002), macrostructure in bilinguals is not completely independent of language; our results do not support this hypothesis. However, it seems to be the case that the production of complete episodes (GAOs) is less influenced by language.

We found no difference in performance between the two *tasks* (Telling, Cat/Dog vs Baby Birds/Baby Goats), neither for the story structure score nor for episodic complexity, which was not surprising considering that the tasks were constructed to be parallel in terms of macrostructure (see Gagarina et al., 2012) and mainly differed in the number and types of characters (three in Cat/Dog, including one human, five in Baby Birds/Baby Goats, all animals). Our results were different from previous findings for story comprehension on the same tasks, where differences were found between the MAIN Cat/Dog and Baby Birds/Baby Goats tasks (Bohnacker & Lindgren, in press; Bohnacker et al., 2020; Lindgren & Bohnacker, 2020a).

Age was found to be a significant predictor of the children's story structure scores in both languages, although the effect was smaller for German than for Swedish. Moreover, older children produced narratives that showed a higher level of episodic complexity, e.g. were more likely to include complete episodes (although such episodes were generally infrequent, ranging from 6%–10%, depending on the language and task). The finding that the older children in our study produced narratives with higher macrostructure scores is in line with previous studies (e.g. Kunnari et al., 2016) and is not surprising: Older children are cognitively more mature and are therefore better at interpreting and verbalizing the story content shown in the pictures. They are also likely to have more experience with narratives and storytelling activities, have acquired a more clearly formulated story schema and may be more aware of what they are expected to tell to a listener than younger children. Older children also generally have more well-developed language skills and this may also influence their narrative production. So why then is age a weaker predictor of narrative macrostructure (story structure score) in German and a stronger one in Swedish? The reason could lie in the Swedish-speaking setting that the children are growing up in, where links between age, language exposure and more well-developed language skills are stronger for the majority language Swedish than for the minority language German. Since the majority of children attend a Swedish-medium (pre)school full-time, they might also have received more exposure to narratives in Swedish (e.g. during joint book reading activities), a factor that may influence children's ability to tell stories (Bitetti & Hammer, 2016; Sénéchal & LeFevre, 2001).

Expressive *vocabulary* scores predicted the children's story structure scores only for German, not for Swedish. This could be linked to the relationship between age and vocabulary, where age significantly affected expressive vocabulary in Swedish, but not in German (see Lindgren & Bohnacker, 2020b). In Swedish, age thus influences both story structure scores and expressive vocabulary scores, which is not the case in German. The finding that expressive vocabulary is a significant predictor of macrostructure in German but not in Swedish may also be linked to the larger variation in vocabulary scores in German. A child needs a certain level of expressive vocabulary in order to tell a story. In Swedish, no child had very low expressive vocabulary scores, which may mean that they all had vocabulary knowledge above the threshold needed to narrate the MAIN stories. In German on the other hand, some children scored very low on expressive vocabulary, and these children may experience difficulties in narrating and expressing the macrostructure in a comprehensible way, even though they are older (and cognitively more mature).

With respect to *narrative length* (linguistic productivity), there was a comparable effect in the two languages; narrative length predicted the story structure

score in both languages, and longer narratives tended to receive higher story structure scores. This result was in line with our expectation based on previous studies (e.g. Uccelli & Páez, 2007).

When *input* was investigated as a separate factor in our preliminary analyses, the proportion of daily input in German as estimated by parents was found to have a significant effect on the German story structure scores. However, when included in a linear mixed effects model together with age, narrative length, vocabulary knowledge and the language of the first testing, input was not significant. Instead, as discussed above, the children's German expressive vocabulary scores significantly predicted their German story structure scores. This suggests that it is not the daily input per se that influences children's ability to express macrostructural components in their narratives, but rather the language skills (such as vocabulary) that may be linked to input.

In addition to the factors discussed above, we also controlled for the effect of the *language of the first testing*. In German, story structure scores were significantly higher when the child was tested in Swedish first. This practice effect was not found for Swedish, suggesting that the children's weaker language may benefit from having already carried out the same type of narrative tasks in the stronger language first, but that the opposite is not the case.

Apart from presenting new results from an understudied language combination (Swedish/German), the current study expands our knowledge in the following ways: It confirms that bilingual children's macrostructural skills, in terms of both story structure and episodic complexity in fictional narratives, develop measurably with age from 4 to 6 years in both their languages, and that the chosen elicitation materials (MAIN; Gagarina et al., 2019) do not lead to (unwanted) task effects in narrative production. This is worth mentioning, as the materials are relatively new and task effects have been found for narrative comprehension (recall Section 2.2). Interestingly however, in the two languages, the children's macrostructure scores were not the same, nor was their development with age; rather, they were stronger in the majority language Swedish. In the minority language German, language proficiency (operationalised here as the expressive vocabulary scores) affected macrostructural performance more than in Swedish. (In Swedish, the effect of expressive vocabulary was not significant.) This echoes findings concerning bilinguals' performance in other domains of language (vocabulary, narrative comprehension) where age often is a strong predictor for majority language performance, but less so for performance in the minority language (regarding vocabulary, see e.g. Gathercole, Thomas, Roberts, Hughes, & Hughes, 2013; Lindgren & Bohnacker, 2020b; Oller & Eilers, 2002; regarding story comprehension, see e.g. Bohnacker et al., 2020). Importantly, our German/Swedish results indicate, in contrast to findings from previous studies (e.g. Akinci

et al., 2001; Bohnacker, 2016; Fiestas & Peña, 2004; Pearson, 2002), that macrostructure in bilinguals is not independent of language. The results can therefore not be interpreted to fully support the assumption that story structure is invariant across a bilingual child's two languages at a given age.

References

Akinci, M.-A., Jisa, H., & Kern, S. (2001). Influence of L1 Turkish on L2 French narratives. In L. Verhoeven & S. Strömqvist (Eds.), *Narrative development in a multilingual context* (pp. 189–208). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.23.07aki

Berman, R.A., & Slobin, D.I. (1994). Narrative structure. In R.A. Berman & D.I. Slobin (Eds.), *Relating events in narrative: A crosslinguistic developmental study* (pp. 39–84). Hillsdale, NJ: Lawrence Erlbaum.

Bishop, D.V.M. (1997). *Uncommon understanding: Development and disorders of language comprehension in children*. Hove/New York: Psychology Press.

Bishop, D.V.M., & Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential meaning. *Journal of Speech Language and Hearing Research*, *35*(1), 119–129. https://doi.org/10.1044/jshr.3501.119

Bitetti, D., & Hammer, C.S. (2016). The home literacy environment and the English narrative development of Spanish-English bilingual children. *Journal of Speech, Language, and Hearing Research*, *59*(5), 1159–1171. https://doi.org/10.1044/2016_JSLHR-L-15-0064

Blom, E., & Boerma, T. (2016). Why do children with language impairment have difficulties with narrative macrostructure? *Research in Developmental Disabilities*, *55*, 301–311. https://doi.org/10.1016/j.ridd.2016.05.001

Boerma, T., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2016). Narrative abilities of monolingual and bilingual children with and without language impairment: Implications for clinical practice: A narrative as diagnostic tool. *International Journal of Language & Communication Disorders*, *51*(6), 626–638. https://doi.org/10.1111/1460-6984.12234

Bohnacker, U. (2016). Tell me a story in English or Swedish: Narrative production and comprehension in bilingual preschoolers and first graders. *Applied Psycholinguistics*, *37*(1), 19–48. https://doi.org/10.1017/S0142716415000405

Bohnacker, U., & Lindgren, J. (in press). MAIN story comprehension: What can we expect of a typically developing child? In S. Armon-Lotem & K.K. Grohmann (Eds.), *LITMUS in Action: Cross-Comparison Studies across Europe*. Amsterdam: John Benjamins. Preprint retrieved from http://uu.diva-portal.org/smash/get/diva2:1348899/FULLTEXT01.pdf (18 June, 2019).

Bohnacker, U., Lindgren, J., & Öztekin, B. (2020). Bilingual Turkish-Swedish children's understanding of MAIN picture sequences: Individual variation, age, language and task effects. In U. Bohnacker & N. Gagarina (Eds.), *Developing narrative comprehension: Multilingual Assessment Instrument for Narratives* (pp. 99–147). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.61.04boh

Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching and Therapy*, *18*(1), 1–21. https://doi.org/10.1191/0265659002ct224oa

Burris, S. E., & Brown, D. D. (2014). When all children comprehend: Increasing the external validity of narrative comprehension development research. *Frontiers in Psychology*, *5* (Article 168). https://doi.org/10.3389/fpsyg.2014.00168

Dickinson, D., & Tabors, P. (2001). *Beginning literacy with language*. Baltimore, MD: Brookes.

Fiestas, C. E., & Peña, E. D. (2004). Narrative discourse in bilingual children: Language and task effects. *Language, Speech, and Hearing Services in Schools*, *35*(2), 155–168. https://doi.org/10.1044/0161-1461(2004/016)

Gagarina, N. (2016). Narratives of Russian–German preschool and primary school bilinguals: Rasskaz and Erzählung. *Applied Psycholinguistics*, *37*(1), 91–122. https://doi.org/10.1017/S0142716415000430

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, *56*, 1–140.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2015). Assessment of narrative abilities in bilingual children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 243–269). Bristol: Multilingual Matters. https://doi.org/10.21832/9781783093137-011

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U., & Walters, J. (2019). MAIN: Multilingual Assessment Instrument for Narratives – Revised. *ZAS Papers in Linguistics*, *63*, 1–36. https://doi.org/10.21248/zaspil.63.2019.516

Gathercole, V. C. M., Thomas, E. M., Roberts, E. J., Hughes, C. O., & Hughes, E. K. (2013). Why assessment needs to take exposure into account: Vocabulary and grammatical abilities in bilingual children. In V. C. M. Gathercole (Ed.), *Issues in the assessment of bilinguals* (pp. 20–55). Bristol: Multilingual Matters. https://doi.org/10.21832/9781783090105-004

Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral discourse in the preschool years and later literacy skills. *First Language*, *24*(2), 123–147. https://doi.org/10.1177/0142723704042369

Gutiérrez-Clellen, V. F. (2002). Narratives in two languages: Assessing performance of bilingual children. *Linguistics and Education*, *13*(2), 175–197. https://doi.org/10.1016/S0898-5898(01)00061-4

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). Designing cross-linguistic lexical tasks (CLTs) for bilingual preschool children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from Language Impairment* (pp. 196–239). Bristol: Multilingual Matters. https://doi.org/10.21832/9781783093137-010

Haman, E., Wodniecka, Z., Marecka, M., Szewczyk, J., Białecka-Pikul, M., Otwinowska, A., Mieszkowska, K., Łuniewska, M., Kołak, J., Miękisz, A., Kacprzak, A., Banasik, N., & Foryś-Nogala, M. (2017). How does L1 and L2 exposure impact L1 performance in bilingual children? Evidence from Polish-English migrants to the United Kingdom. *Frontiers in Psychology*, *8* (Article 1444). https://doi.org/10.3389/fpsyg.2017.01444

Heilmann, J., Miller, J. F., & Nockerts, A. (2010). Sensitivity of narrative organization measures using narrative retells produced by young school-age children. *Language Testing*, *27*(4), 603–626. https://doi.org/10.1177/0265532209355669

Iluz-Cohen, P., & Walters, J. (2012). Telling stories in two languages: Narratives of bilingual preschool children with typical and impaired language. *Bilingualism: Language and Cognition*, *15*(1), 58–74. https://doi.org/10.1017/S1366728911000538

Kunnari, S., Välimaa, T., & Laukkanen-Nevala, P. (2016). Macrostructure in the narratives of monolingual Finnish and bilingual Finnish–Swedish children. *Applied Psycholinguistics*, *37*(1), 123–144. https://doi.org/10.1017/S0142716415000442

Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle, WA: University of Washington.

Lindgren, J. (2018). *Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish children aged 4–6*. Uppsala: Acta Universitatis Upsaliensis.

Lindgren, J. (2019). Comprehension and production of narrative macrostructure in Swedish: A longitudinal study from age 4 to 7. *First Language*, *39*(4), 412–432. https://doi.org/10.1177/0142723719844089

Lindgren, J., & Bohnacker, U. (2020a). Inferential comprehension, age and language: How German-Swedish bilingual preschoolers understand picture-based stories. In U. Bohnacker & N. Gagarina (Eds.), *Developing narrative comprehension: Multilingual Assessment Instrument for Narratives* (pp. 61–98). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.61.03lin

Lindgren, J., & Bohnacker, U. (2020b). Vocabulary development in closely-related languages: Age, word type and cognate facilitation effects in bilingual Swedish-German preschool children. *Linguistic Approaches to Bilingualism*, *10*(5), 587–622. https://doi.org/10.1075/lab.18041.lin

Lynch, J. S., van den Broek, P., Kremer, K. E., Kendeou, P., White, M. J., & Lorch, E. P. (2008). The development of narrative comprehension and its relation to other early reading skills. *Reading Psychology*, *29*(4), 327–365. https://doi.org/10.1080/02702710802165416

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (Third Edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mäkinen, L. (2014). Narrative language in typically developing children, children with specific language impairment and children with autism spectrum disorder (PhD Thesis). University of Oulu, Oulu.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*, 111–151. https://doi.org/10.1016/0010-0285(77)90006-8

Maviş, I., Tunçer, M., & Gagarina, N. (2016). Macrostructure components in narrations of Turkish–German bilingual children. *Applied Psycholinguistics*, *37*(1), 69–89. https://doi.org/10.1017/S0142716415000429

Nordqvist, Å. (2001). Speech about speech: A developmental study on form and function of direct and indirect speech (PhD Thesis). University of Gothenburg, Gothenburg.

Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children*. Clevedon: Multilingual Matters. https://doi.org/10.21832/9781853595721

Otwinowska, A., Mieszkowska, K., Białecka-Pikul, M., Opacki, M., & Haman, E. (2020). Retelling a model story improves the narratives of Polish-English bilingual children. *International Journal of Bilingual Education and Bilingualism*, *23*(9), 36–76. https://doi.org/10.1080/13670050.2018.1434124

Öztekin, B. (2019). *Typical and atypical language development in bilingual Turkish-Swedish preschool children*. Uppsala: Acta Universitatis Upsaliensis.

Paradis, J., Genesee, F., & Crago, M.B. (2011). *Dual language development and disorders: A handbook on bilingualism and second language learning* (2nd ed). Baltimore, MD: Paul H. Brookes.

Paris, A.H., & Paris, S.G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly*, *38*(1), 36–76. https://doi.org/10.1598/RRQ.38.1.3

Pearson, B.Z. (2002). Narrative competence among monolingual and bilingual school children in Miami. In D.K. Oller & R.E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 135–174). Clevedon: Multilingual Matters. https://doi.org/10.21832/9781853595721-008

Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. New York, NY: Plenum Press. https://doi.org/10.1007/978-1-4757-0608-6

Reuterskiöld, C., Hansson, K., & Sahlén, B. (2011). Narrative skills in Swedish children with language impairment. *Journal of Communication Disorders*, *44*(6), 733–744. https://doi.org/10.1016/j.jcomdis.2011.04.010

Reuterskiöld Wagner, C., Sahlén, B., & Nettelbladt, U. (1999). What's the story? Narration and comprehension in Swedish preschool children with language impairment. *Child Language Teaching and Therapy*, *15*(2), 83–93.

Ringblom, N., Håkansson, G., & Lindgren, J. (2014). *Cross-Linguistic Lexical Task: Swedish version (CLT-SWE)*. Unpublished material.

Rinker, T., & Gagarina, N. (2014). *Cross-Linguistic Lexical Task: German version (CLT-DE)*. Unpublished material.

Roch, M., Florit, E., & Levorato, C. (2016). Narrative competence of Italian–English bilingual children between 5 and 7 years. *Applied Psycholinguistics*, *37*(1), 49–67. https://doi.org/10.1017/S0142716415000417

Rodina, Y. (2017). Narrative abilities of preschool bilingual Norwegian-Russian children. *International Journal of Bilingualism*, *21*(5), 617–635. https://doi.org/10.1177/1367006916643528

Schneider, P., Hayward, D., & Dubé, R.V. (2006). Storytelling from pictures using the Edmonton Narrative Norms Instrument. *Journal of Speech-Language Pathology and Audiology*, *30*(4), 224–238.

Sénéchal, M., & LeFevre, J.-A. (2001). Storybook reading and parent teaching: Links to language and literacy development. *New Directions for Child and Adolescent Development*, *92*, 39–52. https://doi.org/10.1002/cd.14

Stein, N.L., & Glenn, C.G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Discourse processing: Multidisciplinary perspectives* (pp. 53–120). Norwood, NJ: Ablex.

Stein, N.L., & Policastro, M. (1984). The concept of a story: A comparison between children's and teacher's viewpoints. In H. Mandl, N.L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 113–155). Hillsdale, NJ: Lawrence Erlbaum.

Strömqvist, S., & Day, D. (1993). On the development of narrative structure in child L1 and adult L2 acquisition. *Applied Psycholinguistics*, *14*(2), 135–158. https://doi.org/10.1017/S0142716400009528

Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes*, *15*(3), 249–275. https://doi.org/10.1080/01638539209544812

Trabasso, T., & Rodkin, P.C. (1994). Knowledge of goal/plans: A conceptual basis for narrating "Frog where are you?". In R.A. Berman & D.I. Slobin (Eds.), *Relating events in narrative: A crosslinguistic developmental study* (pp. 85–106). Hillsdale, NJ: Lawrence Erlbaum.

Trabasso, T., Stein, N. L., Rodkin, P. C., Munger, M. P., & Baughn, C. R. (1992). Knowledge of goals and plans in the on-line narration of events. *Cognitive Development*, *7*(2), 133–170. https://doi.org/10.1016/0885-2014(92)90009-G

Uccelli, P., & Páez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language Speech and Hearing Services in Schools*, *38*(3), 225–236. https://doi.org/10.1044/0161-1461(2007/024)

van den Broek, P., Kendeou, P., Kremer, K., Lynch, J. S., Butler, J., White, M. J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. In S. Stahl & S. Paris (Eds.), *Children's reading comprehension and assessment* (pp. 107–130). Mahwah, N.J.: Erlbaum.

Van Dongen, R., & Westby, C. E. (1986). Building the narrative mode of thought through children's literature. *Topics in Language Disorders*, *7*(1), 70–83. https://doi.org/10.1097/00011363-198612000-00009

Viberg, Å. (2001). Age-related and L2-related features in bilingual narrative development in Sweden. In L. Verhoeven & S. Strömqvist (Eds.), *Narrative development in a multilingual context* (pp. 87–128). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.23.04vib

Westby, C. E. (2005). Assessing and remediating text comprehension problems. In H. W. Catts & A. G. Kamhi (Eds.), *Language and reading disabilities*. (2nd edition, pp. 157–232). Boston, MA: Pearson.

Address for correspondence

Josefin Lindgren
Uppsala University
Department of Linguistics and Philology
Box 635
751 26 Uppsala
Sweden

lindgren.josefin@gmail.com
https://orcid.org/0000-0003-3995-8315

Co-author information

Ute Bohnacker
Uppsala University
Department of Linguistics and Philology

ute.bohnacker@lingfil.uu.se
https://orcid.org/0000-0002-6715-7470

Publication history

Date received: 9 April 2020
Date accepted: 20 November 2020
Published online: 28 January 2021