Storytelling in bilingual Turkish-Swedish children

Effects of language, age and exposure on narrative macrostructure

Ute Bohnacker, Josefin Lindgren, and Buket Öztekin
Uppsala University

The empirical evidence for whether narrative macrostructure skills are shared between a bilingual child's two languages is inconclusive, and it is not known how macrostructure (overall story structure) is influenced by general language proficiency and amount of exposure. The present study investigates these issues in 100 Turkish-Swedish bilingual 4-to-7-year-old children growing up in Sweden. Oral narratives were elicited in both Turkish and Swedish with two picture-based tasks from the Multilingual Assessment Instrument for Narratives (MAIN) in the telling mode. We investigated to what extent the language of elicitation influences bilingual children's macrostructure (story structure, episodic complexity), and explored effects of age, narrative task, narrative length, expressive vocabulary and language exposure, both separately and combined, on macrostructure in the respective language. Story structure and episodic complexity were found to increase similarly with age in both Turkish and Swedish from 4 to 7 years. Scores did not differ between the two MAIN storytelling tasks. Expressive vocabulary and narrative length influenced story structure scores positively and similarly in both languages. Daily language exposure and length of exposure to Swedish did not show any significant effect. The results can be interpreted in support of a carry-over of narrative macrostructural skills between the two languages.

Keywords: episodic complexity, story structure, Swedish, Turkish, vocabulary

https://doi.org/10.1075/lab.20057.boh | Published online: 16 September 2021
Linguistic Approaches to Bilingualism 12:4 (2022), pp. 413–445. ISSN 1879-9264 | E-ISSN 1879-9272

1. Introduction

Oral narratives are a cultural universal, and we tell, listen to, and communicate through stories every day. Narratives are a tool for making sense of the world and sharing our experience of it. Narrative competence involves interrelated cognitive and linguistic skills that allow us to build a coherent representation of a story and to convey it to a listener. To develop oral narrative competence is important for social reasons, but also for academic success, as narrative skills have been found to be linked to children's literacy development and reading comprehension (Roth et al., 1996; Dickinson & Tabors, 2001; Griffin et al., 2004).

What about storytelling in children growing up with more than one language? How does narrative competence develop? Do aspects of narrative competence carry over between a bilingual's languages, or are such relationships contingent on language proficiency, language exposure and age of the child?

Narratives can be characterised on two levels, the organisation of the story, the so-called *macrostructure* or *story structure*, and the specific linguistic forms that make up the story, the *microstructure* (e.g. Van Dongen & Westby, 1986; Hickmann, 2004). Both levels are needed for there to be a narrative. Macrostructure is foundational though, as it concerns the overall structure of narrative content. It is thought to be largely independent of the structural properties of a language and to follow a common set of principles, where characters engage in goal-directed actions in an attempt to solve a problem (Stein & Glenn, 1979; Mandler et al., 1980; Berman & Slobin, 1994).[1] Note that macrostructure is not simply the number of information units or propositions included in the story, or the combined information contained in these units. Rather, narrative macrostructure is the underlying structural organisation of such narrative content. Macrostructural components encode different types of information which serve specific functions in the story structure, and together form the core of the narrative. Macrostructure is often analysed in terms of thematic units or story grammar components (see 2.1).

Microstructure, on the other hand, includes a range of lexical and grammatical forms employed by the storyteller, for instance referential, temporal and causal linking devices to promote cohesion (Berman, 1988). Microstructural devices build cohesion not only in narratives, but in any kind of discourse. Microstructure

1. Notwithstanding culture-specific differences in storytelling conventions and styles (e.g. Gorman et al., 2011), it is generally assumed that the macrostructure of narratives has, in Paradis et al.'s words (2011, p. 125), "a large conceptual component not tied to any specific language".

is often analysed through measures of language-specific referring expressions, connectives, tense, and choice of vocabulary.[2]

As microstructure is embedded in specific language knowledge, there is agreement that narrative microstructural skills are affected by general language proficiency (vocabulary, morphosyntax), and language exposure (Pearson, 2002; Iluz-Cohen & Walters, 2012; Squires et al., 2014; Govindarajan & Paradis, 2019). Macrostructure, on the other hand, is assumed to be far less dependent on specific knowledge of lexis and grammar, and also less dependent on exposure (Pearson, 2002; Hipfner-Boucher et al., 2015). Macrostructure has been hypothesised to be shared across languages in multilinguals. Second-language (L2) learners who know how to coherently structure a narrative in their first language (L1) may use this knowledge when telling stories in the other language as well, i.e. transfer it. For children acquiring more than one language, this would mean that despite uneven language exposure, and even when proficiency is lower in one language than the other, macrostructural skills could carry over between languages. This view has indeed been put forward by a number of researchers (e.g. Berman, 2001: 420–426; Pearson, 2002; Paradis et al., 2011: 125–127).

Despite an increasing number of studies on the narrative abilities of bilinguals, the empirical evidence is inconclusive as to whether macrostructure skills are shared between a bilingual child's two languages, and it is not known how macrostructure is influenced by general language proficiency and amount of exposure.

To shed light on these questions, the present study explores the narratives of 100 Turkish-Swedish bilingual children aged 4–7. To ensure comparability, strictly parallel elicitation and scoring are used in both languages. We report on two different measures of macrostructure: story structure scores, and episodic complexity. We investigate to what extent the language of elicitation influences the children's macrostructure, and explore effects of age, task, testing order, narrative length, daily language exposure, length of exposure, and independent measures of expressive vocabulary knowledge, both separately and combined. These factors have rarely been investigated in bilingual children of any language combination in relation to narrative macrostructure.

The paper proceeds as follows. Section 2 provides some background on macrostructure and factors that may affect its development in children. Section 3 describes the aim and research questions. Section 4 presents the methodology, and Section 5 the results. Finally, Section 6 contains a discussion and conclusion.

2. Sometimes measures of productivity (e.g. the number of words or utterances), lexical diversity, syntactic complexity, and grammatical accuracy are also subsumed under microstructure.

2. Background

2.1 Narrative macrostructure

There are a number of different models of narrative macrostructure (e.g. Labov & Waletzky, 1967; Stein & Glenn, 1979; Peterson & McCabe, 1983; Berman & Slobin, 1994). The *story grammar* framework (e.g. Mandler & Johnson, 1977; Stein & Glenn, 1979) has often been used in studies of children's fictional narratives (e.g. Trabasso et al., 1992; Schneider et al., 2006). Common to all models of story grammar is that they treat narratives as having a *setting* that provides the context, specifies time and place and introduces the protagonists, and an *episodic system*. This episodic system consists of one or several episodes, each of which is made up of a behavioural sequence that is causally and temporally related, centering around the animate protagonist's goal (a *goal-based episode*): an initiating event or internal state of the story character, the character's goal, an attempt or action to reach the goal, an outcome or consequence of the attempt in terms of the goal, and an internal state as a reaction to the outcome. The components of the model on which the narrative instrument used in this study is based (MAIN, Gagarina et al., 2012, 2019, see 4.2) are shown in Table 1.

Table 1. Macrostructural (story structure) components in MAIN (Gagarina et al., 2012, 2019)

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

Macrostructure is commonly investigated by analysing which story grammar components are included in the narration and counting them, resulting in a story structure score (e.g. Mandler & Johnson, 1977; Stein & Glenn, 1979; Schneider et al., 2006; Hayward et al., 2009); this is also done in the present study. Alternative macrostructural measures determine the level of complexity of component combinations in an episode (Stein & Policastro, 1984; Hughes et al., 1997; Westby,

2012). Here, researchers often classify episodes as belonging to one of the following four complexity levels, on the basis of Goal (G), Attempt (A) and Outcome (O); this is also done in the present study:

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

Episodes in which fewer than two components are expressed contain *no sequence* of components. Combinations of an attempt and its outcome are *action/reaction sequences* (attempt-outcome sequences, AO), as illustrated 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). *Complete episodes* (goal-attempt-outcome sequences, GAO), as shown in (4), indicate a high level of episodic complexity (Stein & Glenn, 1979; Gagarina et al., 2012: 11–12; Westby, 2012: 211). Examples (1)–(4) come from the Turkish child data of the present study.[3]

(1) **AO** (5;8, BiTur5–11)
bir tane kedi kelebek gördü. sonra kedi atladı [A]. *kelebek kaçtı* [O].
'A cat saw a butterfly. Then the cat jumped [A]. The butterfly escaped [O].'

(2) **GA** (4;5, BiTur4–05)
bir kedi bir kelebek gördü ve onu yakalamak istedi [G]. *o zıpladı ve kelebeğin arkasında(n) gitti* [A].
'A cat saw a butterfly and wanted to catch it [G]. And it jumped and went after the butterfly [A].'

(3) **GO** (5;8, BiTur5–23)
kedi şunu yakalayacaktı [G]. *kedi yapamadı, kedi bu düştü* [O].
'The cat was going to catch this [G]. The cat couldn't do it, the cat fell down [O].'

(4) **GAO** (4;10, BiTur4–12)
kedi geliyor onu kelebeği yakalamak istiyor [G], *sonra şöyle hopluyor, sonra şöyle kelebeği yakalamaya geliyor* [A], *sonra kelebek uçuyor* [O].
'The cat comes and wants to catch the butterfly [G], and then (it) jumps like this, and then (it) comes to catch the butterfly like this [A]. Then the butterfly flies [O].'

3. All examples are from the MAIN Cat story; age (years;months) and child code name are given in parentheses.

2.2 Factors influencing children's production of narrative macrostructure

2.2.1 *Age*

Age has been pointed out as a strong contributor to children's storytelling skills, where macrostructural growth is in evidence between age 3–4/5 and 6–7 (Berman, 1988; Trabasso & Nickels, 1992; Trabasso et al., 1992; Berman & Slobin, 1994; Pearson, 2002; Schneider et al., 2006; Squires et al., 2014; Bohnacker, 2016; Gagarina, 2016; Lindgren, 2019). Macrostructure, being a cognitive-linguistic skill, is of course to be expected to improve with age – alongside increasing general cognitive maturity, processing abilities and children's expanding world knowledge and exposure to stories.

Berman and Slobin (1994) were some of the first to observe that children's narrative macrostructure depends more on age than on the language a story is told in. In their well-known study of monolingual English-, German-, Hebrew-, Spanish- and Turkish-speaking children age 3 to 9 and adults telling the Frog Story from a wordless picture book, they found strong commonalities across languages for the same age groups, but striking differences between the 3/4-, 5- and 9-year-olds and adults. From age 5, children increasingly started to link events together temporally and causally to form episodic sequences and began to include some motivations in their narratives. Although goals started to appear in children's narratives at age 5, proportions of goals similar to those of adults were not reached until age 9 (Trabasso et al., 1992). Older children produced more complex episodic structure, including complete episodes.

For bilinguals, chronological age has also been found to be a strong predictor of macrostructural skills, but as few studies have investigated both languages in relation to age, it is less clear whether age contributes equally in both languages. Pearson (2002) analysed the Frog stories told by 160 Spanish-English bilingual second- and fifth-graders (age 7–8, 10–11). Narrative quality was operationalised as a 'language score' (a composite of lexicon, syntax and referencing) and a 'story score' (a composite of macrostructural elements, sequencing and internal states). Pearson's cross-sectional results show large gains in story structure in both Spanish and English from second to fifth grade. Uccelli and Páez (2007) studied 24 Spanish-English bilinguals who told a 3-picture story at age 5–6, and again one year later. Using an adaptation of Pearson's (2002) scoring, they found significantly improved narrative skills in both Spanish and English at age 6–7. However, in English, the children's scores were higher and gains over time were greater than for the minority language Spanish (p. 230, 233). The same pattern was found by Squires et al. (2014) who studied 21 Spanish-English children retelling Frog stories at age 5, and again one year later, even though the elicitation method (retell) and the measures of macrostructure were different. Other studies have found sim-

ilar macrostructural gains with age in both languages of bilingual children from age 4–5 to 6–7. However, in a cross-sectional study of 58 children with a narrow age range (5–6) and late age of L2 onset, chronological age was not a good predictor of macrostructural skills in the L2, whereas age of onset was (Govindarajan & Paradis, 2019). (L1 macrostructure was not investigated.) The present study investigates the effect of age on the macrostructural skills of Turkish-Swedish bilingual children in both languages.

2.2.2 *The two languages*

Although there is general agreement that macrostructural skills can be transferred or shared, the empirical evidence regarding the relationship between the narrative abilities in a bilingual's two languages is inconclusive. A number of studies report similar macrostructure scores in both languages, suggesting carry-over (Akinci et al., 2001; Pearson, 2002; Iluz-Cohen & Walters, 2012; Squires et al., 2014; Bohnacker, 2016; Gagarina, 2016; Kunnari et al., 2016). Other studies have found poorer macrostructure skills in one of the languages (Gutiérrez-Clellen, 2002; Kapalková et al., 2016; Roch et al., 2016).

It is plausible to assume that these inconclusive results are due to methodological differences. First, studies use a variety of materials and elicitation modes (e.g. retelling of an aurally presented story vs. story generation/telling, with or without visual support), with divergent findings. Second, the participants' ages vary from preschoolers to school-aged children. The relationship between the narrative skills in the two languages may change as a function of age and exposure. Third, studies operationalise macrostructure in different ways and measure different components. Some measurements are relatively coarse, where children may already be credited with full points when having mentioned the main protagonist(s), an initiating event, one attempt and one outcome/consequence, so that ceiling is easily reached (e.g. Akinci et al., 2001). Different measures may also yield contradictory results; e.g., Gagarina (2016) found Russian-German children's macrostructure measured on a story structure score to correlate significantly in the two languages, but not when measured as episodic complexity.

Studies that employ the same elicitation tool as the present study (e.g. Gagarina, 2016; Kunnari et al., 2016; Lindgren, 2018) often report findings that accord with those of studies using other stimuli: development with age, similar macrostructure in mono- and bilinguals and in bilinguals' two languages. For instance, Bohnacker (2016) found for 52 English-Swedish bilinguals that 6–7-year-olds had higher story structure scores than 5-year-olds, irrespective of language. Age effects were also found when macrostructure was measured via episodic complexity, though complete episodes (GAOs) were relatively rare at

age 4–7. The 5-year-olds in Bohnacker (2016) produced only 7% GAOs, and only slightly more at age 6–7 (12%).

Interestingly, the macrostructural scores of the oldest children in these studies were nowhere near ceiling, and much lower than adults. Gagarina et al. (2019) analysed MAIN stories told by 69 German-, Russian- and Swedish-speaking adults, and their mean story structure scores were 50%–80% higher than the children's (age 6–7). As for episodic complexity, the adults produced three to five times as many complete episodes (GAOs) as the children in the studies above, on the same storytelling tasks (pp. 198–201). Thus, 6–7-year-olds still have some way to go to reach adult-like levels on these tasks. (For similar results regarding the Frog story, see Berman, 1988; Trabasso et al., 1992; Berman & Slobin, 1994.)

Some studies that use MAIN to investigate macrostructure in bilingual children have found divergent performance, with significantly lower scores in one of the languages (Kapalková et al. (2016) for 40 Slovak-L1/English-L2 5-to-6-year-olds; Roch et al. (2016) for 62 Italian-L1/English-L2 5-to-7-year-olds). In both studies, the youngest children had limited L2 exposure (one or two years of English as an L2 via school). They may not have had sufficient language resources to display their macrostructural knowledge in the L2. Neither study investigated language proficiency (such as vocabulary) and/or exposure in relation to macrostructural performance.

2.2.3 *Exposure, language proficiency and other factors*

The effect of language exposure on bilingual children's macrostructure has rarely been investigated. To our knowledge, there is only one such study with MAIN, namely Haman et al. (2017). They studied the effects of L1 and L2 cumulative language exposure on L1 Polish language skills, including story structure, in 53 Polish-English bilingual 4-to-7-year-olds. (Unusually however, Haman et al. averaged the scores for narratives elicited in different modes, telling and retelling.) They found an effect of cumulative exposure on Polish story structure, which may indicate that increased language exposure leads to better narrative skills. (The other language, English, was not investigated.)

Using Frog story elicitation materials, Pearson (2002) studied 160 Spanish-English bilingual children, separated into groups according to amount of language exposure and SES. Pearson found no effects of exposure (nor of SES) on the narrative quality when measured via story structure, but clear effects when measured via lexis, syntax and referencing, which suggests that exposure has an effect on microstructure, but not on macrostructure. Hipfner-Boucher et al. (2015) studied three groups of 25 Canadian 4-to-5-year-old children (with different L1s), with different amounts of exposure to English, retelling the Renfrew Bus story. They found that the bilingual group with the least exposure to English

included fewer scorable, lexically specific information units in their retells; however, story grammar scores did not significantly differ from the other two groups with more exposure to English (pp. 686–687). (The children's other languages were not investigated.) Govindarajan and Paradis (2019) administered yet another type of narrative stimuli, the ENNI storytelling task, to 63 bilingual 5-to-6-year-olds who on average had slightly less than 2 years of exposure to English. Length of exposure to English and language richness were found to be good predictors of the children's L2-English story grammar scores. (The L1s were not investigated.) In sum, the few studies on exposure and macrostructure in bilingual children present conflicting results.

Some researchers have studied how bilingual children's *lexical knowledge* (as measured on an independent vocabulary test) affects narrative macrostructure. Pearson (2002), in the aforementioned study of 160 Spanish-English bilinguals, found moderate correlations between the results of standardised vocabulary tests and the 'story score' (which includes aspects of macrostructure) in both languages. In a longitudinal study of 81 Spanish-English bilinguals, Bitetti and Hammer (2016) found an effect of receptive vocabulary (PPVT) on the children's Frog story scores according to the Narrative Scoring Scheme (NSS), a composite measure that includes aspects of macrostructure.

A few studies have also found significant relationships between measures of *narrative vocabulary* or the *length of the narrative* and macrostructure (or a similar measure) in bilinguals. Whilst lexical diversity and longer narratives do not necessarily mean better macrostructure, Uccelli and Páez (2007) found that story structure correlated with narrative length and the scores on an expressive vocabulary test in both languages for their 24 Spanish-English bilinguals. Bitetti et al. (2020) found, for 200 Spanish-English bilingual preschoolers, that lexical density in Frog stories predicted NSS scores for both languages. Other studies have found no relationships between story structure and various microstructural measures (e.g. Iluz-Cohen & Walters, 2012; Rodina, 2017), which could be due to the relatively small sample sizes in these studies ($N=17$; $N=16$). Different materials and measures make comparisons difficult. The present study investigates the effect of age, language, exposure, narrative length and vocabulary skills (as a proxy for general language proficiency) on the story structure of Turkish-Swedish bilingual children.

3. Aim and research questions

This study investigates oral narratives in 4–7-year-old Turkish-Swedish bilingual children. Story structure and episodic complexity are analysed for 4 different sto-

rytelling tasks (two in Turkish, two in Swedish, MAIN, see 4.2). The following research questions are asked:

a. Do language (Turkish vs. Swedish) and narrative task (MAIN Cat/Dog vs. Baby Birds/Baby Goats) influence story structure and episodic complexity?
b. How does age relate to story structure and episodic complexity?
c. To what extent does language proficiency as measured on an independent vocabulary task and as narrative length affect story structure?
d. How does language exposure/input relate to story structure?

4. Method

4.1 Participants

The participants were 100 4-to-7-year-old Turkish-Swedish bilingual children growing up in urban areas of eastern central Sweden. The children had Turkish as their home language, and Swedish as the language of schooling. Basic inclusion criteria were: (i) ability to speak both Turkish and Swedish, (ii) 4–7 years, and (iii) no diagnosis of language impairment or neuropsychiatric disorder. The participants came from 50 different preschools and schools and were part of a larger multilingualism research project. The present study includes only children for whom we have both storytelling and vocabulary data in both languages. Table 2 gives an overview of these 100 children aged 4;0–8;1 (M = 73.2 months, SD = 14.0, range 48–97 months).[4]

Table 2. Participants

	4 years	5 years	6 years	7 years	Total
N	24	22	26	28	100
Girls/boys	12/12	12/10	15/11	14/14	53/47
Age range	4;0–4;11	5;0–5;11	6;1–6;11	7;0–8;1*	4;0–8;1

Note. * Three children in the 7-year-old group had just turned 8 years at the time of testing.

Informed parental consent was obtained in writing. Families and children could terminate participation at any time. Background information was collected

4. The breakdown by age group in Table 1 is simply intended to show that children were relatively evenly distributed across the ages. In the remainder of the paper, age is treated as a continuous variable.

via a parental questionnaire, administered in Turkish or Swedish. Questionnaire data were available for all 100 children.

Nearly all children, 92% (92/100), had lived in Sweden from birth.[5] By contrast, only 22% of the parents had been born and grown up in Sweden. Most parents had been born and grown up in Turkey and migrated to Sweden as adults (with residence lengths of 0.6 to 41 years). For most children (90%) then, both parents were first-generation immigrants from Turkey or one first-generation and one second-generation parent. A large majority of parents (75%) reported Turkish as their native language (L1), 5% Swedish as their L1, and 0.5% (one parent) both Turkish and Swedish as L1s. Very few children (8/100) had a parent who was a native speaker of Swedish. 20% of the parents reported a L1 different from Turkish or Swedish; in nearly all cases, this language was Kurdish.[6]

With Turkish so strongly represented as parental L1, virtually all children in the study (96%, 96/100) had received Turkish input from their parents since birth. Only for 4 children did exposure to Turkish start after 1;0, as they were first exposed to Kurdish via their Kurdish-L1 parents. Onset of exposure to Swedish varied more than for Turkish. Most children (85%) had started to hear Swedish regularly before age 3;0. For 15% of the participants, exposure to Swedish began after age 3, and for very few after age 4; these were exclusively children born in Turkey who had later moved to Sweden.[7]

A majority of parents reported that they spoke almost only or mostly Turkish to their child (67%) and to each other (68%). Some spoke Turkish and another language (often Kurdish) to each other (11%) and to the child (7%). The remaining parents spoke both Turkish and Swedish to each other (14%) and to the child (30%). Only two couples spoke only/mostly Swedish to each other (2%), and none spoke only/mostly Swedish to their child (0%). The majority of children (66%) were reported to speak almost only/mostly Turkish to both parents; 12% children spoke only/mostly Swedish to their parents, and the remainder a mix of both languages. Communication between siblings was reported to be only/

5. The background description in this section is similar to that provided in Bohnacker, Öztekin & Lindgren (2020), a study of narrative comprehension skills. However, since the participants in the present study are not fully identical to those included in Bohnacker et al. (2020), the percentages reported are not identical.

6. Children who were exposed to an additional third language in the home, such as Kurdish, were not excluded, since they form an integral part of the Turkish-Swedish population and did not generally perform differently on our Turkish and Swedish language measures than the strictly bilingual children.

7. We did not exclude children born in Turkey (with at least two years exposure to Swedish), since they did not generally perform differently on our measures of Swedish and Turkish than the children born in Sweden.

mostly Turkish for 34% of the children, for 47% both Swedish and Turkish, and for 16% only/mostly Swedish, which is very different from parental language use and suggestive of a generational shift of language preference.

The questionnaire also asked parents to estimate the proportion of daily language input to the child. Parents estimated this on a scale ranging from 95% Turkish/5% Swedish to 5% Turkish/95% Swedish, with the option of writing down their own distribution. Based on this estimate, the daily input mean for the whole group was 46.3% Turkish ($SD = 20.7$, range 5–95, median = 40), and 55.6% Swedish ($SD = 19.5$, range 5–95, median = 60). The majority of children (68%) were reported to receive approximately equal amounts of Turkish and Swedish during the day (40:60, 50:50, 60:40). For 19 children (20%), parents reported daily input to be at least 80% Swedish; for 9 children (9%), at least 80% Turkish input was reported. The few remaining children had other input distributions, including a third language.

All children were active bilinguals. 6% of the children (6/100) also spoke a third language: Kurdish (Kurmanji), Zaza (Dimili), German or English, but for none of them was this third language rated as strong as Turkish or Swedish. A few other children did not speak a third language but were exposed to it; this mainly occurred in families where the parents had Kurdish as their L1 and spoke a mix of Turkish and Kurdish to each other.

The majority of parents rated their children's receptive and expressive language skills in both Turkish and Swedish as 'very good' or 'good', though some rated one language as 'so-so' or even 'bad'. 41% of the children were said to have equal language skills in both languages; for 35%, the parents stated that the child's language skills were higher in Turkish than in Swedish. For a smaller but substantial group (23%), parents rated their children to be more proficient in Swedish.

Virtually every child attended institutional childcare for a major part of the day, as is typical in the Swedish context. On average, children had started preschool at 23 months (range 12–48 months), with a few late starts (60–72 months) in cases of recent relocation from Turkey. Of the children that we had attendance information for ($N = 93$), the large majority (84%) were reported to regularly attend (pre)school for 30–46 hours per week, 10% attended 20–26 hours, and 6% 6–9 hours/week.[8] All schools and preschools were run in Swedish, though on closer inspection during our data collection visits, other children and/or a staff member also spoke Turkish.

8. Preschool in Sweden is normally attended until age 5 or 6, after which children attend one year of 'preschool class' (*förskoleklass*) to prepare them for school proper. At age 7, children start grade 1 of primary school.

Most children lived in linguistically and culturally diverse urban areas. The individual children came from a wide variety of socio-economic backgrounds, regarding both parental occupations and education, where all levels from non-completed primary school to doctorate were represented. This variation can be considered typical of Turkish-speaking families in Sweden. Many parents had completed secondary education but did not have any tertiary education.

4.2 Materials

4.2.1 *MAIN*

Oral narratives were elicited from the children in Turkish and Swedish using the Multilingual Assessment Instrument for Narratives (MAIN, Gagarina et al., 2012, 2019). MAIN consists of 4 picture sequences (*Cat, Dog, Baby Birds, Baby Goats*) with 6 pictures each. Each of these stories is made up of three episodes. The four picture sequences are designed to be parallel in length and story grammar components, and two each (*Cat* and *Dog; Baby Birds* and *Baby Goats*) are strictly parallel with regard to their plotline and number of story characters. Because of this parallelism, we refer to *Cat/Dog* as one narrative task, and *Baby Birds/Baby Goats* as the other. There is a standardised procedure for administering the tasks and a scheme for scoring (see 4.3, 4.4).

4.2.2 *Vocabulary: CLT*

To get a measure of the children's vocabulary knowledge, picture-based Cross-linguistic Lexical Tasks were administered in both Turkish and Swedish (CLT, Haman, Łuniewska, & Pomiechowska, 2015). The CLTs are the only existing vocabulary tasks that are directly comparable for Turkish and Swedish (CLT-TR, Ünal-Logacev, Tunçer & Ege, 2013; CLT-SWE, Ringblom, Håkansson & Lindgren, 2014). Each version consists of 4 parts (noun comprehension, verb comprehension, noun production and verb production), with 30 vocabulary items in each part (plus 2 practice items). As the children performed close to ceiling on the comprehension parts, we only included the scores for vocabulary production in the present study, in order to be able to explore variation. CLT vocabulary production is tested via picture naming, where the experimenter asks the child to name a depicted object ("What is this?" for noun production) or action ("What is she doing?" for verb production). In each language, the maximum total vocabulary production score is 60 points.

4.3 Procedure

Each child told two MAIN stories per language and did the CLT vocabulary tasks in each language. Turkish and Swedish were tested on different occasions with at least four days in between ($M = 13.6$ days, $SD = 7.9$, range 4–36 days). The order of the languages was counterbalanced; half of the children in each age group were seen in Turkish first and the other half in Swedish first. Experimenters were native speakers of Turkish and Swedish (the authors and trained research assistants). Sessions took place in a quiet room in the home or (pre)school. Experimenters only spoke to the child in the language of testing and acted as if they had no knowledge of the other language. The entire procedure was audio and video recorded.

The stories were counterbalanced across ages and languages, and children never received the same story twice. Cat or Dog was always done first, followed by the CLT, which in turn was followed by Baby Birds or Baby Goats. The stories were administered according to the 'telling mode' procedure as described in Gagarina et al. (2019). The child was presented with three envelopes, picked one, took out the picture sequence inside and looked at the pictures. Then the pictures were folded up again so that only the first two were visible when the child was prompted to tell the story. More pictures were unfolded as the child kept telling, until all six pictures were unfolded. This procedure of previewing the pictures and telling the story with the pictures visible reduces the cognitive effort required in narration. As the pictures were held away from the view of the experimenter, there was no shared visual attention, so the child could not rely on joint eye gaze or pointing when telling the story.[9] The CLTs were administered according to standard procedure (Haman et al., 2015).

4.4 Coding and analysis

All narratives were transcribed orthographically in CHAT-format (MacWhinney, 2000) by the second and third authors and five trained research assistants (native speakers of Turkish and Swedish) and carefully checked. The three authors and a research assistant then scored the narratives for story structure, following the MAIN scoring protocol (Gagarina et al., 2019). Macrostructural scoring focuses on story content and not on the language-specific linguistic skills of the child, such as choice of words and grammatical correctness. The MAIN *story structure*

9. Afterwards, the experimenter looked at the pictures together with the child and asked 10 standardised comprehension questions. Comprehension results are not reported here (see Bohnacker et al., 2020).

score is a sum of macrostructural components and assesses the child's ability to express these components. The score includes setting (max 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, attempt, outcome, internal state as reaction (3 × 5 components). The maximal score is 17 points. Table 3 gives an overview of setting and episode 1 for the Cat story.

Table 3. Macrostructural components, setting and episode 1 (constructed examples)

Component	Cat
Setting	*Once upon a time…* (time) *… by a lake/ in a meadow* (place)
Episode 1	
Internal state as initiating event	*A cat saw a butterfly (sitting on a bush)*
Goal	*… and wanted to catch it*
Attempt	*The cat jumped forward/ started to chase it …*
Outcome	*The butterfly flew away/ the cat missed it /fell into the bush*
Internal state as reaction	*The cat was angry*

As children often expressed story content in ways that the manual did not provide scoring decisions for, additional scoring guidelines were developed.[10] Unclear cases were discussed and scored on the basis of these guidelines. Three principles were central for scoring. First, only those words produced by the child that were comprehensible in the target language were considered for scoring. Second, the child's verbal realisation 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. 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 story in Turkish and Swedish (four scores per child).

10. The guidelines (Bohnacker 2018b) included both general scoring principles and decisions on a large number of unclear cases, based on extensive, systematic discussions of the narratives of 286 mono- and bilingual children.

On the basis of the different story structure components, *episodic complexity* was calculated, following the scoring protocol (Gagarina et al., 2019). Each episode was coded as one of the following types of sequences of components (recall 2.1): *GAO-sequence* (G+A+O are produced in the episode), *GA/GO-sequence* (either G+A or G+O), or *AO-sequence* (A+O). Episodes that did not contain any of these combinations were coded as *no sequence*. The maximum for each narrative was 3 GAOs, as each picture sequence contained 3 episodes.

CLT vocabulary responses were documented on paper during testing and later checked against the audio and video recordings. The 1,200 responses were scored as 1 (correct) or 0 (incorrect) by the three authors and a trained research assistant (native or near-native speakers of Turkish and Swedish) and carefully checked for consistency and accuracy.[11]

In addition to story structure and episodic complexity, we included the following measures as independent variables (predictors) in the analyses: The *length of each narrative* in total number of words (TNW) was calculated in CLAN (MacWhinney, 2000) using the *freq* function. Only words in the target language were included in this calculation.[12] Each child received a CLT *vocabulary production score* (max = 60) in Turkish and another one in Swedish. The child's chronological *age in months* (range: 48–97 months) was included in the analyses. *Length of exposure to Swedish* in months (range: 14–94 months) was calculated based on the information about age of onset in the parental questionnaire. The values for the child's *estimated percentage of daily input* in Turkish and Swedish (on a scale from 0 to 100) were based on information in the questionnaire. Finally, the *language of the first testing* (Swedish or Turkish) was included in the analyses.

To investigate the effects of language (Turkish vs Swedish), narrative task (Cat/Dog vs Baby Birds/Baby Goats), age, narrative length, vocabulary knowledge (CLT production scores), length of exposure and daily 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.

---

11. We followed the CLT scoring guidelines developed by the first author and her research team on the basis of data from 220 monolingual and bilingual children (Bohnacker 2018a), to complement the short information contained in the CLT test materials.

12. Some children produced words in the other language. This was more common in Turkish (2% of the words) than in the Swedish narratives (< 0.05%). These words were not included in the narrative length measure.

5. Results

The 100 children told two stories each in Turkish and two in Swedish (except one child who only told one), yielding 399 narratives that could be analysed for story structure.

5.1 Story structure score

Table 4 shows the children's story structure scores for Turkish and Swedish. Mean scores were generally low and did not even reach a third of the maximum. The mean scores and score ranges were similar across languages and tasks.

Table 4. Story structure scores (max = 17)

	Turkish		Swedish	
	Cat/Dog (*N* = 100)	*BB/BG* (*N* = 100)	*Cat/Dog* (*N* = 100)	*BB/BG* (*N* = 99)
Mean (SD)	4.6 (2.3)	5.1 (2.7)	5.3 (2.4)	5.0 (2.7)
Range	0–10	0–11	0–12	0–11

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

A linear mixed-effects model (with a random effect for Child)[13] showed no significant effects of Language ($\beta = 0.04$, $SE = 0.28$, $p = .89$) and Task ($\beta = 0.25$, $SE = 0.28$, $p = .37$), and the interaction effect was also not significant ($\beta = -0.69$, $SE = 0.40$, $p = .08$). The children thus scored similarly in the two languages and there was no difference in story structure scores between the two tasks Cat/Dog and Baby Birds/Baby Goats. (For scores in relation to age, see 5.4.)

5.2 Narrative length

Table 5 shows the results for narrative length (TNW) for the two tasks in Turkish and Swedish. The range in TNW is very large, suggesting considerable individual variation.

A linear mixed-effects model (with a random effect for Child) showed a significant effect of Language ($\beta = -11.93$, $SE = 3.38$, $p < .001$), but not of Task ($\beta = 1.95$, $SE = 3.38$, $p = .57$); the interaction effect was also not significant ($\beta = -3.27$, $SE = 4.77$, $p = .49$). Overall, Turkish narratives thus contained fewer words than the

13. Adding random slopes to the models was not possible; no models containing random slopes converged.

Table 5. Narrative length (TNW)

	Turkish		Swedish	
	Cat/Dog	*BB/BG*	*Cat/Dog*	*BB/BG*
Mean (SD)	68.4 (23.6)	69.8 (27.0)	83.6 (33.7)	77.7 (26.7)
Range	13–143	14–165	18–237	30–227

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

Swedish narratives, but this should not be taken to mean that the children were more productive in their Swedish storytelling. Rather, the different word counts are due to the agglutinative nature of Turkish, where one inflected Turkish word conveys the same information as several Swedish words together. Within each language, the two narrative tasks show similar lengths.[14]

5.3 Expressive vocabulary knowledge

Table 6 shows the children's vocabulary production scores (CLT) in the two languages.

Table 6. Vocabulary production scores (CLT)

	Turkish	Swedish
Mean (SD)	39.3 (11.7)	35.0 (11.1)
Range	3–60	8–54

A linear mixed-effects model (with a random effect for Child) showed a significant effect of Language ($\beta = 29.40$, $SE = 7.81$, $p < .001$), with higher scores in Turkish, and Age ($\beta = 0.49$, $SE = 0.07$, $p < .001$), with higher scores for older children, but also a significant interaction between Language and Age ($\beta = -0.34$, $SE = 0.11$, $p < .01$). As an earlier study (Öztekin, 2019) found a difference in the development of vocabulary with age in the two languages, we included age as well as the interaction between age and language in this analysis. As shown by the linear regressions in Figure 1, there was a significant effect of age on vocabulary production in Swedish, but not in Turkish. The ranges in scores in both languages are very large, indicating substantial individual variation.[15]

14. In the present study, we are not interested in narrative length per se, but whether there is a relationship between macrostructure scores and narrative length, see 5.4.

15. Whilst language dominance is not the topic of this paper, a reviewer suggested that we investigate dominance by subtracting the Swedish vocabulary production scores from the Turk-

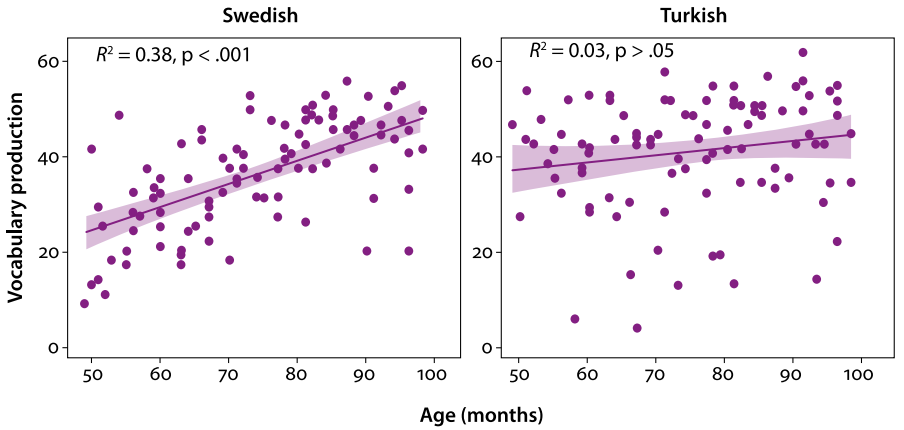

Figure 1. The relationship between the child's age in months and the vocabulary production score (max = 30)

5.4 Effects of age, linguistic factors and exposure on story structure

Next, we investigated the effects of age, narrative length (TNW), the child's vocabulary production (CLT), length of exposure to the majority language Swedish, and proportion of daily language input on the story structure score in each language. First, we show the relationship between the story structure score and each of the factors separately with a series of scatterplots that include regression lines (simple linear regressions) and R^2-values. The R^2-values show how much of the variation in the story structure score is explained by the factor.[16]

The first factor included is age. Figure 2 shows that in both languages, there is a significant relationship between the child's age in months and the story structure score. The relationship is slightly stronger in the majority language Swedish than in the minority language Turkish. The relationship between age and Swedish

ish ones. A positive score indicates a higher score in Turkish (and possibly Turkish dominance), and a negative score indicates a higher score in Swedish (and possibly Swedish dominance). The mean 'dominance score' for the group as a whole was 4.34 (SD = 17.45), which is the same as the difference in the mean vocabulary scores in Table 6. There was a significant effect of age in months on the dominance score, even though it was small (r^2 = 0.075, p < .01). The correlation between age and Turkish dominance was negative ($r(98) = -0.274$, $p < .01$), which means that older children were less Turkish-dominant than younger children. However, not even the oldest children were Swedish-dominant in terms of vocabulary, but relatively balanced ($M_{7\text{-years}} = -0.36$, SD = 16.31). Individual variation was large at all ages.

16. The scatterplots are based on two story structure scores (one for Cat/Dog and one for Baby Birds/Baby Goats) in each language for every child in the sample (except for one child where only one Swedish and two Turkish stories were available).

story structure (Figure 2, left) is reminiscent of the pattern found for age and Swedish vocabulary (Figure 1, left). For Turkish however, despite the lack of a relationship between age and Turkish vocabulary (Figure 1, right), there is a significant relationship between age and Turkish story structure (Figure 2, right).

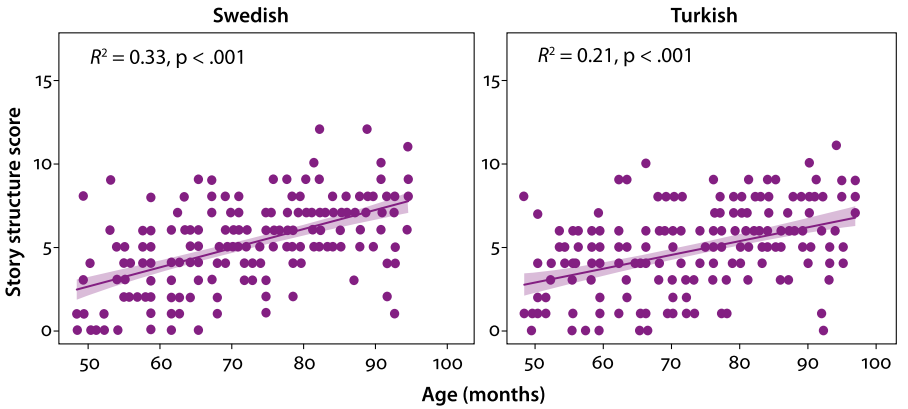

Figure 2. The relationship between the child's age in months and the story structure score (max=17)

The next factor is narrative length (TNW). As shown in Figure 3, narrative length has a significant effect on the story structure scores. The effect is similar in both languages.

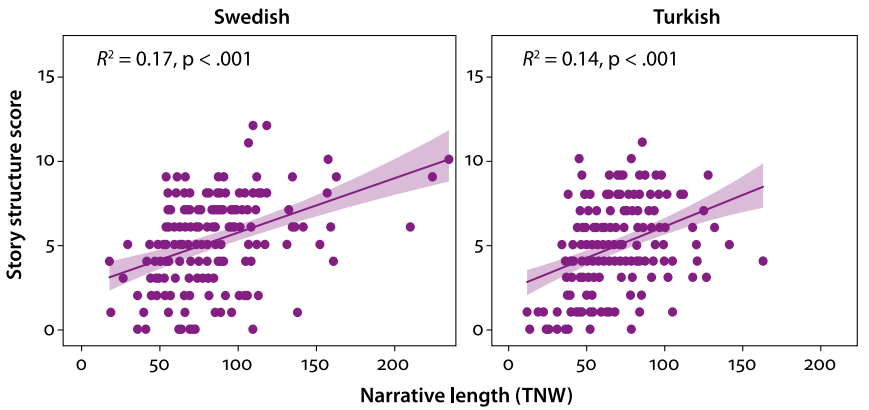

Figure 3. The relationship between narrative length (TNW) and the story structure score (max=17)

Next, the relationship between the children's vocabulary production scores (CLT) and story structure scores is investigated (Figure 4). In both languages,

there is a significant relationship, and it is somewhat stronger in the majority language Swedish.

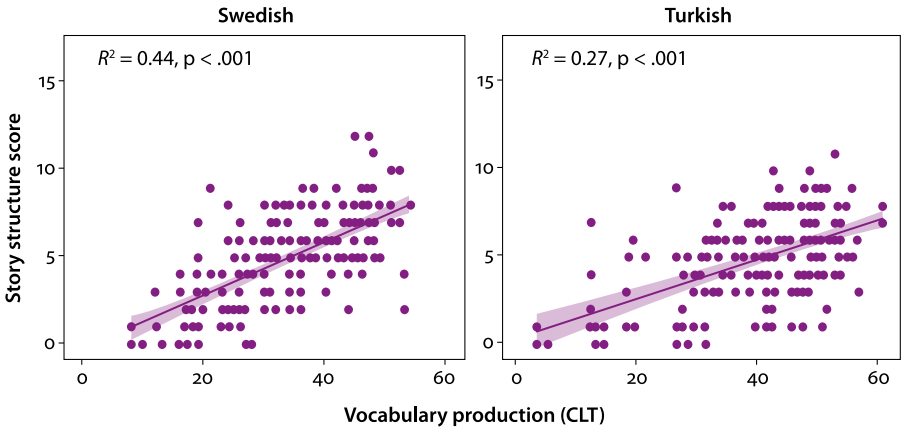

Figure 4. The relationship between vocabulary production score (CLT; max=60) and story structure score (max=17)

Concerning the child's length of exposure (LoE) to the majority language Swedish (Figure 5), there is a positive effect of increasing length of exposure, in both languages, but this effect is much stronger in Swedish and negligible in size in Turkish. The slight effect manifest in Turkish may be due to LoE correlating with chronological age. (As will be shown below, in the multivariate analysis, LoE is no longer significant for Turkish, but age is.) Note that there is no negative effect, i.e. longer exposure to Swedish does not lead to lower Turkish scores.

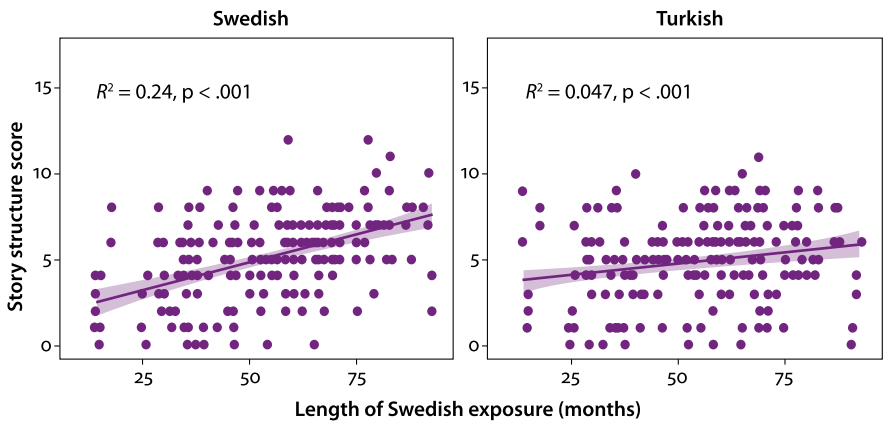

Figure 5. The relationship between the child's length of Swedish exposure (in months) and the story structure score (max=17)

Finally, we explore the effect of the children's estimated current daily input in one language (in %) on the story structure score in that language (Figure 6). Here, there is a significant effect for Swedish but not for Turkish, though even in Swedish the effect is very small. This is noteworthy, as there is considerable variation in the daily input patterns of the children. (As we will see below, in the multivariate analyses daily input is no longer significant.)

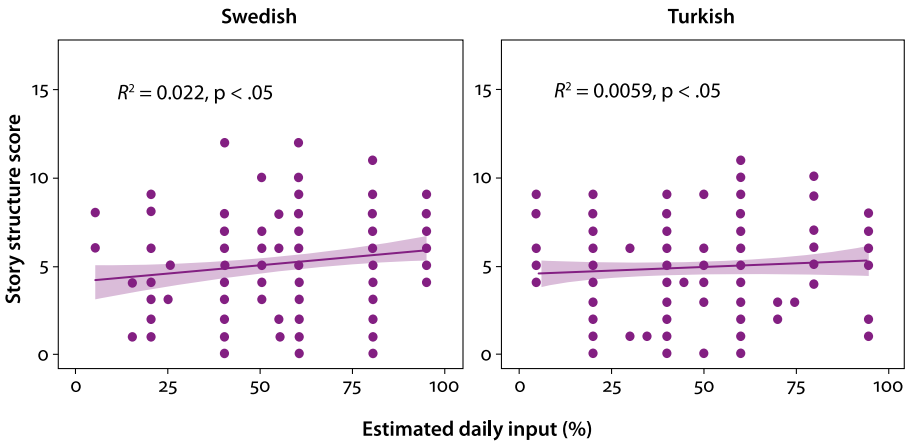

Figure 6. The relationship between the child's estimated daily input in one language (%) and the story structure score in that language (max = 17)

Having investigated the relationships between each of the factors age, narrative length vocabulary production scores, length of exposure to Swedish, daily language input and the story structure score separately, we carried out multivariate analyses to investigate the effect of all variables simultaneously. Table 7 shows the results for the linear mixed-effect models (with a random effect for Child) for Turkish and Swedish. In these models, we also controlled for the language of the first testing (Test 1).

In both languages, there were significant effects of the child's age (older children generally scored higher), narrative length (longer narratives generally received higher scores), and vocabulary (children with higher vocabulary production scores scored higher). In neither language were there any effects of the length of exposure to Swedish, nor of the estimated proportion of daily input. Altogether, this indicates that story production in the minority language Turkish and in the majority language Swedish is affected by the same factors, and in much the same way.

Only one factor included in the above models, namely order of presentation of the narrative tasks, yielded a difference between the languages. Here, the effect

Table 7. Summary of linear mixed-effect models for macrostructure production scores of the narrative tasks in Swedish and Turkish

	Turkish		Swedish	
	β	*SE*	β	*SE*
Intercept	−4.28***	0.92	−3.44***	0.84
Age (months)	0.05***	0.01	0.05***	0.01
Narrative length	0.02***	0.01	0.02***	<.01
Vocabulary	0.10***	0.01	0.10***	0.02
LoE	0.02	0.01	<.01	0.01
Input	<.01	0.01	<.01	0.01
Test1	0.85**	0.28	−0.40	0.27

Note. ** $p < .01$ *** $p < .001$.
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; LoE = length of exposure to the majority language Swedish in months; Input = estimated proportion of daily input in the language of the narrative task; Test1 = language of the first testing; the models show the effect when the language of the first testing is Swedish.

of language of the first testing (Test 1) was significant for Turkish but not for Swedish. This means that children who were tested in Swedish first tended to receive higher story structure scores for their Turkish narratives. (We discuss this further in Section 6.)

5.5 Episodic complexity

The narrative data contained 1,200 potential episodes (100 children × 4 stories × 3 episodes each). As one Swedish story was missing for one child, there were 1,197 episodes in total that could be analysed for complexity.

As can be seen in Figure 7, the children's narratives generally showed a relatively low level of complexity. For both languages and tasks, episodes that contained *no sequence* of the central macrostructural components were by far the most frequent type (54%–67% out of all episodes). The second most common episode type was the *AO-sequence* (20%–30% of the episodes). Episodes containing goals as part of macrostructural sequences were relatively infrequent. This was the case both for *GA/GO-sequences* (7%–11%) and complete episodic structures, *GAO-sequences* (4%–8%).

Attempts and outcomes are clearly depicted in the MAIN picture sequences and might thus be described more easily than other episodic components, such as

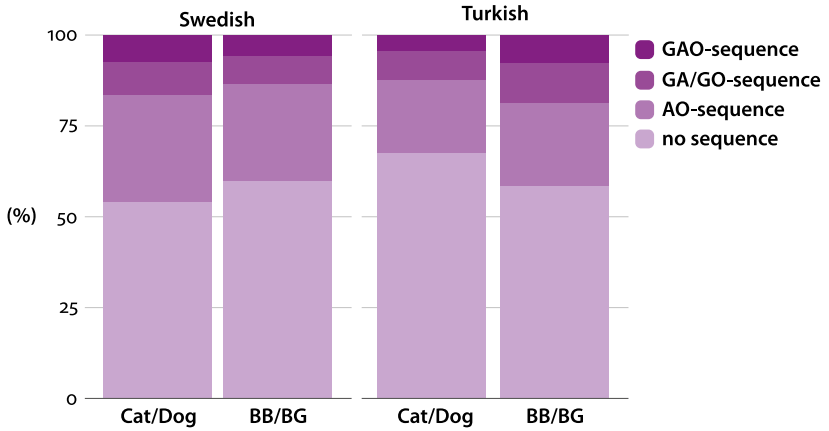

Figure 7. Types of macrostructural sequences, by language and narrative task ($N = 1{,}197$)

goals, which need to be inferred. However, it is striking that there were so many instances of *no sequence* in the data. This suggests that whilst the children include story structure components in their narratives (as we have seen in their story structure scores in Section 5.1), they often do not combine these components into sequences. They may express an attempt, or an outcome, or, though more rarely, a goal, but not together with other scorable components in that episode. Alternatively, they may omit to mention an episode altogether. Whilst *no sequence* is common in both languages, it is even more common in Turkish (for statistics see below). The following examples illustrate this. In Example (5), from a seven-year-old's telling of the Cat story, of all the events in episode 3, only the outcome (O) is mentioned, whilst neither goal nor attempt are described.

(5) Cat story, Episode 3, BiTur7-05, Turkish (only O, no sequence)
kedi de bütün balıkları yedi
cat too all fish-PL-ACC eat-PAST-3SG
'The cat ate all the fish.'

In Example (6), from a five-year-old's telling of the same story, in episode 1, only the goal is mentioned, but neither attempt nor outcome.

(6) Cat story, Episode 1, BiTur5–15, Turkish (only G, no sequence)
kedi bi(r) tane kelebek görmüş, onu yakalamak istemiş,
cat one butterfly see-PAST.3SG it-ACC catch want-PAST.3SG
'The cat saw a butterfly, wanted to catch it.'

In Example (7), from another five-year-old's telling of the Cat story, Episode 1 is alluded to, but only in vague terms ('a cat comes, comes towards it'). As neither the cat's goal (wanting to catch the butterfly) nor the attempt or outcome are men-

tioned, this first episode is scored as *no sequence*. Episode 2 is narrated cursorily, there is no mention of the boy having a ball, it rolling into the water, him wanting to get it back and attempting to fish it out with his fishing rod. Only the outcome of that episode is mentioned ('took the ball', 1 point), but this outcome is very hard for a listener to understand without any preceding context. Episode 2 is thus also an instance of *no sequence*. Regarding Episode 3, the child makes some reference to it ('then takes, eats them') but without any previous mention of the boy carrying the fish, putting them down and the cat stealing them, 'then takes, eats them' remains opaque to the listener (who cannot see the pictures), so episode 3 is also scored as no sequence.

(7) Cat story, BiTur5–12, Turkish (no sequences in Episodes 1-2-3)
bir kelebek ağaçta ve kedi geliyo(r), onun üstüne
a butterfly tree-LOC and cat come-PRES.3SG it-POSS top-GEN-DAT
geliyo(r) ve oğlan geliyo(r).
come-PRES.3SG and boy come.PRES.3SG
'A butterfly is on a tree and cat comes, comes towards it and a boy comes.'
sonra alıyo(r), onları yiyo(r) ve topu da almış
then take-PRES.3SG them-ACC eat-PRES.3SG and ball-ACC too take-PAST.3SG
oluyo(r)
AUX-PRES.3SG
'then takes, eats them and took/happened to take the ball.'

Whilst a more detailed qualitative analysis of macrostructural complexity lies outsides the confines of this paper, we noticed that many children, especially the youngest in their Turkish narratives, appeared to focus on the outcome of an episode, as if they were rushing through the events, leaving out goal and attempt to get to the outcome. This results in a high incidence of *no sequence*. In Swedish, there is also a focus on the outcome, but to a lesser degree, and many instances of *no sequence* are due to rudimentary language and formulations that are so unclear that they do not score a point, especially in the youngest children (age 4). Older children (age 6–7) produce many more scorable components, as well as AO-sequences.

In our statistical analyses of macrostructural complexity, we focused on the children's production of *no sequences* and *GAO-sequences*. These represent the levels with the lowest and highest complexity. For the production of *no sequences*, the logit mixed-effects model (with a random effect for Child) showed significant effects of Language ($\beta = 0.28$, $SE = 0.13$, $p < .05$), with a higher proportion of episodes with no sequence in in Turkish than in Swedish, and of Age ($\beta = -0.04$, $SE = 0.01$, $p < .001$); older children produced fewer episodes without any sequence, but no significant effect of Task ($\beta = 0.06$, $SE = 0.13$, $p = .62$). The production of

GAOs was generally low in both languages and for both tasks. For this reason, using a mixed effects model was deemed unsuitable. A logistic regression analysis (with the `glm` function in R) showed a significant effect of Age ($\beta = 0.04$, $SE = 0.01$, $p < .001$), with older children more frequently producing GAOs, but no significant effects of Language ($\beta = -0.12$, $SE = 0.24$, $p = .62$) or Task ($\beta = -0.12$, $SE = 0.24$, $p = .62$).

6. Discussion and conclusion

The present study investigated macrostructure in the oral narratives of 100 Turkish-Swedish children aged 4–7. In both languages, story structure scores and levels of episodic complexity were low, considerably lower than what has been reported for adults telling the same stories (Gagarina et al., 2019). This suggests that even the oldest of our participants (age 7) still have some way to go to achieve adult-like macrostructure and full-blown accounts of story events.

In both Turkish and Swedish, *older* children generally produced more complete episodes (GAO) and fewer episodes without any sequence. Older children also achieved higher overall story structure scores. This is in line with earlier research that used the same or different elicitation materials and methods of macrostructural scoring with bilingual children (e.g. Pearson, 2002; Uccelli & Páez, 2007; Squires et al., 2014; Bohnacker, 2016; Gagarina, 2016; Kunnari et al., 2016). Older, cognitively more mature, children are better at interpreting and expressing the story content shown in the pictures. Their expressive language skills may also be more advanced, which may positively influence narrative production. Older children are also likely to have more experience with narratives and story-telling activities and may have more well-developed mental story schemata.

It is worth noting that story structure increased with age in *both* the majority language and the minority language. This finding differs from a range of studies covering *other domains* of language, in particular vocabulary and sometimes also morphosyntax, where the minority language tends to show a flat growth curve or no increase with age (e.g. Oller & Eilers, 2002, for Spanish; Gathercole et al., 2013, for Welsh; Montanari et al. 2018, for Russian). The pattern observed for story structure also differs from the one observed for expressive vocabulary in our own cross-sectional data, where Swedish vocabulary scores strongly increased with age, but Turkish vocabulary scores did not (see 5.3). We know that vocabulary development is heavily dependent on input, and lexical knowledge cannot straightforwardly be transferred between languages (unless they share many cognates). Knowledge of story structure is a more language-general ability and may lend itself more easily to transfer (see below).

In both Turkish and Swedish, *longer narratives* generally received higher story structure scores, and children with *higher vocabulary production* also scored higher. In neither language were there any effects of the estimated proportion of current daily input, nor of the length of exposure to the majority language Swedish. These results accord well with Pearson's (2002) large-scale study of Spanish/English second- and fifth-graders, and Hipfner-Boucher et al.'s (2015) of bilingual 4-to-5-year-olds, despite differences in age and elicitation methods. In the present study, length of exposure to Swedish did affect macrostructure scores when this factor was investigated on its own, but when included in a linear mixed effects model together with age, narrative length, vocabulary knowledge and language of the first testing, length of exposure was no longer significant. Our results differ from Govindarajan and Paradis' (2019), who found length of exposure to the majority language English to be a significant predictor of story grammar scores, whereas chronological age was not. However, Govindarajan and Paradis' participants were L2 learners with generally shorter exposure lengths and a narrower age span than our Turkish-Swedish children. In the present study, children's expressive vocabulary scores predicted their story structure scores in the respective language. This suggests that it is not the amount of input per se that influences the children's ability to express macrostructural components in their narratives, but rather the language skills (such as vocabulary) that may be a result of the amount of input.

Taken together, our results indicate that macrostructure production in the minority language Turkish and in the majority language Swedish is affected by the same factors, and in much the same way. This is an important finding, as few studies have explored both languages of bilingual children in this area (cf. Lindgren & Bohnacker 2021). It should be interesting to see whether future studies of other minority/majority language pairings will arrive at similar results.

We found no difference in performance between the two MAIN *tasks* (Cat/Dog, Baby Birds/Baby Goats), neither for story structure score nor for episodic complexity, which was expected as the tasks were constructed to be parallel in terms of macrostructure (Gagarina et al., 2012).[17]

We also controlled for the effect of the language of the first testing. In Turkish, story structure scores were significantly higher when the child had been presented with the tasks in Swedish first. There was no such practice effect in the other direction. Whilst our participants do not generally appear to be dominant in Swedish (see 5.3, fn. 15)), most of them do spend a large part of the day in Swedish

17. Our results were different from earlier findings for story *comprehension* by the same children that documented task differences between Cat/Dog and Baby Birds/Baby Goats (Bohnacker et al., 2020).

(pre)school settings, where curriculum activities include shared book reading and storytelling. They may thus be more exposed to narratives in Swedish, which can influence children's ability to tell stories (Bitetti & Hammer, 2016). In the minority (and home) language Turkish, activities that foster narrative abilities are likely to be more varied from child to child. Future research might profitably include investigations of both preschool and home-based storytelling and shared book reading activities and their relation to story production skills.

Overall, we found no language difference in the story structure scores nor in episodic complexity, when measured as the proportion of complete GAO-episodes (though there were more instances of *no sequence*, the lowest complexity level, in Turkish than in Swedish). These findings can be taken to support the assumption that bilinguals in their two languages have a similar understanding of story events and similar awareness of the goal-directed behaviour of the protagonists at a given age (Akinci et al., 2001; Berman, 2001; Pearson, 2002; Paradis et al., 2011). Bilingual children should readily transfer such conceptual, language-general, narrative abilities from one language to another and profit from them. Yet our finding of equivalent macrostructural scores in the two languages at group level does not mean that every individual child is equivalent in both languages; recall the significant effects of vocabulary knowledge and narrative length on macrostructure scores. A child may be hard pressed to display their macrostructural knowledge if the expressive skills (e.g. vocabulary) in one language are very limited. The resulting narrative may be very short and omit most macrostructural components – or it may be longer but lexically so unspecific and hard to follow that few, if any, components are scorable. In consequence, the child may receive a very low macrostructure score in that language, lower than in the other language. Thus, our group results (and the group results of previous studies) cannot be taken as evidence that macrostructural skills are independent of language proficiency, as has sometimes been suggested. What we can say is that macrostructure is more dependent on age than on a specific language (in keeping with e.g. Mandler et al., 1980; Berman & Slobin, 1994; Pearson, 2002). Transfer of macrostructural knowledge may be contingent on the child's having attained a threshold level of proficiency.

Funding

Research funded by Vetenskapsrådet (VR 421-2013-1309) and by Riksbankens jubileumsfond (P19-0644:1) to Ute Bohnacker.

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). John Benjamins. https://doi.org/10.1075/sibil.23.07aki

Berman, R.A. (1988). On the ability to relate events in narrative. *Discourse Processes*, *11*(4), 469–497. https://doi.org/10.1080/01638538809544714

Berman, R.A. (2001). Narrative development in multilingual contexts: A cross-linguistic perspective. In L. Verhoeven & S. Strömqvist (Eds.), *Narrative development in a multilingual context* (pp. 419–428). John Benjamins. https://doi.org/10.1075/sibil.23.16ber

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). Lawrence Erlbaum.

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

Bitetti, D., Hammer, C.S., & López, L.M. (2020). The narrative macrostructure production of Spanish-English bilingual preschoolers: Within- and cross-language relations. *Applied Psycholinguistics*, *41*, 79–106. https://doi.org/10.1017/S0142716419000419

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. (2018a). *Guidelines for Scoring CLT*. Uppsala University, Version March 2018 [unpublished material].

Bohnacker, U. (2018b). *Guidelines for Scoring Macrostructure in MAIN*. Uppsala University, Version November 2018 [unpublished material].

Bohnacker, U., Öztekin, B., & Lindgren, J. (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). John Benjamins. https://doi.org/10.1075/sibil.61.04boh

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

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., Bohnacker, U., & Lindgren, J. (2019). Macrostructural organization of adults' oral narrative texts. *ZAS Papers in Linguistics*, *62*, 190–208. https://doi.org/10.21248/zaspil.62.2019.449

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., 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). Multilingual Matters. https://doi.org/10.21832/9781783090105-004

Gorman, B. K., Fiestas, C. E., Peña, E. D., & Clark, M. R. (2011). Creative and stylistic devices employed by children during a storybook narrative task: A cross-cultural study. *Language, Speech and Hearing Services in Schools*, *42*(2), 167–171. https://doi.org/10.1044/0161-1461(2010/10-0052)

Govindarajan, K. & Paradis, J. (2019). Narrative abilities of bilingual children with and without Developmental Language Disorder (SLI): Differentiation and the role of age and input factors, *Journal of Communication Disorders*, *77*(1), 1–16. https://doi.org/10.1016/j.jcomdis.2018.10.001

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

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). 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*, 1444. https://doi.org/10.3389/fpsyg.2017.01444

Hayward, D., Schneider, P., & Gillam, R. B. (2009). Age and task-related effects on young children's understanding of a complex picture story. *The Alberta Journal of Educational Research*, *55*(1), 54–72.

Hickmann, M. (2004). Coherence, cohesion, and context: Some comparative perspectives in narrative development. In S. Strömqvist, & L. Verhoeven (Eds.), *Relating events in narrative* (pp. 281–306). Psychology Press.

Hipfner-Boucher, K., Milburn, T., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2015). Narrative abilities in subgroups of English language learners and monolingual peers. *International Journal of Bilingualism*, *19*(6), 677–692. https://doi.org/10.1177/1367006914534330

Hughes, D., McGillvray, L. & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessments*. Eau Claire: Thinking Publications.

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

Kapalková, S., Polišenská, K., Marková, L., & Fenton, J. (2016). Narrative abilities in early successive bilingual Slovak–English children: A cross-language comparison. *Applied Psycholinguistics*, *37*(1), 145–164. https://doi.org/10.1017/S0142716415000454

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). University of Washington.

Lindgren, J. (2018). *Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish children aged 4–6*. Studia Linguistica Upsaliensia 19. 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. (2021). How do age, language, narrative task, language proficiency and exposure affect narrative macrostructure in German-Swedish bilingual children aged 4 to 6? *Linguistic Approaches to Bilingualism*. https://doi.org/10.1075/lab.20020.lin

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum Associates.

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

Mandler, J.M., Scribner, S., Cole, M., & DeForest, M. (1980). Cross-cultural invariance in story recall. *Child Development*, *51*, 19–26. https://doi.org/10.2307/1129585

Montanari, E.G., Abel, R., Graßer, B. & Tschudinovski, L. (2018). Do bilinguals create two different sets of vocabulary for two domains? Vocabulary development and overlap in the first years of schooling. *Linguistic Approaches to Bilingualism*, *8*(4), 502–522. https://doi.org/10.1075/lab.16021.mon

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

Öztekin, B. (2019). *Typical and atypical language development in Turkish-Swedish bilingual children aged 4–7*. Studia Linguistica Upsaliensia 25. 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.). Brookes.

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). 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*. Plenum Press. https://doi.org/10.1007/978-1-4757-0608-6

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

Roth, F.P., Speece, D.L., Cooper, D.H. & de la Paz, S. (1996). Unresolved mysteries: How do metalinguistic and narrative skills connect with early reading? *Journal of Special Education*, *30*(3), 257–277. https://doi.org/10.1177/002246699603000303

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.

Squires, K.E., Lugo-Neris, M.J., Peña, E.D., Bedore, L.M., Bohman, T.M. & Gillam, R.B. (2014). Story retelling by bilingual children with language impairments and typically developing controls. *International Journal of Language & Communication Disorders*, *49*(1), 60–74. https://doi.org/10.1111/1460-6984.12044

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). 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). Lawrence Erlbaum.

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., 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)

Ünal-Logacev, Ö., Tunçer, A.M., & Ege, P. (2013). Cross-Linguistic Lexical Tasks: Turkish version (CLT-TR). Unpublished material.

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

Westby, C.E. (2012). Assessing and remediating text comprehension problems. In A.G. Kamhi & H.W. Catts (Eds.), *Language and reading disabilities* (3rd ed., pp. 163–225). Pearson.

Address for correspondence

Ute Bohnacker
Department of Linguistics & Philology
Uppsala University
Box 635
SE-751 26 Uppsala
Sweden

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

Co-author information

Josefin Lindgren
Department of Scandinavian Languages
Uppsala University
josefin.lindgren@nordiska.uu.se
https://orcid.org/0000-0003-3995-8315

Buket Öztekin
Department of Linguistics & Philology
Uppsala University
buket.oztekin@lingfil.uu.se

Publication history

Date received: 9 July 2020
Date accepted: 3 August 2021
Published online: 16 September 2021