Psychology of Language and Communication 2022, Vol. 26, No. 1

sciendo

DOI: 10.2478/plc-2022-0007

Sara Košutar, Matea Kramarić, Gordana Hržica

Department of Speech and Language Pathology, Faculty of Education and Rehabilitation Sciences University of Zagreb, Croatia

The relationship between narrative microstructure and macrostructure: Differences between six- and eight-year-olds

The current study aimed to investigate age-related differences in narrative abilities at the macrostructural and microstructural levels and to examine which microstructural aspects explain narrative macrostructure at ages six and eight. Oral narratives were elicited from 89 Croatian monolingual children using the Multilingual Assessment Instrument for Narratives (MAIN). At the microstructural level, the measure of lexical diversity D, clausal density, and mean length of clause were assessed. Macrostructure was assessed using the standardized MAIN scoring procedure. We found differences between the two age groups in lexical diversity, clausal density, and macrostructure, with eight-year-olds scoring higher on all measures. Variance in the macrostructure was explained to a significant extent by lexical diversity in the case of six-year-olds, and by both lexical diversity and clausal density in the case of eight-year-olds. Our results suggest that six-year-olds rely mostly on lexical abilities when telling a story, while eight-year-olds also draw on syntactic abilities.

Key words: microstructure, macrostructure, narrative analysis, children, age-related differences

Address for correspondence: Sara Košutar, Department of Speech and Language Pathology, Faculty of Education and Rehabilitation Sciences University of Zagreb, Zagreb, Croatia

E-mail: sara.kosutar@erf.unizg.hr

Narration is a major form of discourse used in everyday communication, and it allows us to share our personal experiences, feelings, and thoughts (Karmiloff & Karmiloff-Smith, 2002). Narrating requires the development and integration of at least two types of knowledge (Aksu-Koç & Aktan-Erciyes, 2018). First, knowledge about the overall story structure at the conceptual level, with all the elements that entails (e.g., setting, goal, attempt, and outcome). Second, knowledge about different language aspects, such as vocabulary, morphosyntax, and cohesion in order for story events to relate to one another at the linguistic level. Relying simultaneously on conceptual and linguistic knowledge to produce a well-formed narrative is a complex task, especially for children. Although narratives can be found even in children's early language production, it takes time to develop the necessary knowledge and the ability to integrate different types of knowledge successfully.

Narrative analysis has been used extensively as a tool for tracing the development of narrative competence (e.g., Botting, 2002; Justice et al., 2006). Information about the overall story structure and linguistic aspects can be obtained at two levels underlying the narrative structure: microstructure and macrostructure (Gagarina et al., 2012; Gagarina, Klop, Kunnari, et al., 2019). At the level of macrostructure, the focus is on assessing children's cognitive abilities to produce a coherent, highly structured narrative. At the level of microstructure, narratives are assessed in terms of language abilities necessary for producing a cohesive narrative.

Although micro- and macrostructure are usually considered to be interrelated, their relationship has not been studied in depth (Hickmann, 2003; Karmiloff & Karmiloff-Smith, 2002). It has been argued that macrostructure is not language-specific, is less dependent on language skills, and is grounded in general cognitive processes (Berman, 2001; Paradis et al., 2011; Trabasso & Nickels, 1992). However, a separate line of research demonstrates a close relationship between micro- and macrostructure (e.g., Berman & Slobin, 1994; Peterson & McCabe, 1991). It has been shown that certain language skills correlate with the story structure at the higher level of discourse organization (Bishop & Donlan, 2005; Fernández, 2013; Heilmann et al., 2010; Mäkinen et al., 2014). Although this line of research has yielded interesting results, the role of age in the relationship between language abilities and story structure has not been thoroughly investigated yet.

In the present study, we examined the relationship between macrostructure and microstructure in two age groups of Croatian-speaking children: six- and eight-year-olds. Given that both macro- and microstructural abilities develop with age, the relationship between them might also change during development as the children acquire more sophisticated language skills. The role of language abilities in building story structure in children of different ages has not been extensively studied yet, and many questions remain unanswered.

Narrative Macrostructure

The macrostructure of the narrative refers to the overall story structure. As speakers need to link the utterances together and to the theme of the story, maintaining a coherent macrostructure is a challenging task. Therefore, the development of macrostructure has been an important research topic (e.g., Berman & Slobin, 1994; Hickmann, 2003). Research has focused mainly on two analytical frameworks: the approach proposed by Labov and Waletzky (1967), also known as high-point analysis (see Peterson & McCabe, 1983), and the story grammar approach (Stein & Glenn, 1979). Research on personal narratives often uses high-point analysis, while fictional narratives are more often analyzed using the story grammar model (Aksu-Koç & Aktan-Erciyes, 2018). In high-point analysis, three main story components can be discerned: complicating actions (what happened), orientations (who, when, and where), and evaluations (emotional references, exclamations, judgments, and perceptions; Peterson & McCabe, 1983). According to the story grammar model, the structure of the story consists of the setting and episodes (Hickmann, 2003; McCabe & Bliss, 2003). The setting serves to introduce the time, place, and characters of the story. Next, each episode consists of an initiating event, which is some problem that the main character needs to resolve. The initiating event triggers the character's internal response or reaction to the problem, which is related to their goal, guiding their actions. The attempt represents the action that the main character undertakes to resolve the problem. This leads to some outcome, which then results in a reaction or the character's emotional response. The Multilingual Assessment Instrument for Narratives (MAIN; Gagarina et al., 2012; Gagarina, Klop, Kunnari, et al., 2019) uses the story grammar model to analyze macrostructure. The MAIN builds on the ideas of story grammar components to propose a new system of story structure at the macrostructural level (Gagarina, Klop, Kunnari, et al., 2019). The MAIN macrostructure evaluation differentiates between story structure (setting and three episodes), which reflects story grammar quantity; and story complexity (combinations of episode components such as goal, attempt, and outcome), which reflects story structure quality. Complete episodes containing a goal, attempt, and outcome (GAO) indicate the highest level of story complexity. The present study focused on elements of story structure and evaluated them using standardized MAIN scoring.

Already by the age of three or four, children can produce narratives in response to a fictional storytelling task, albeit with help from adults in the form of encouragement or questions (Berman & Slobin, 1994; Capone Singleton & Shulman, 2014). First narratives are produced in simple structures, with complex story grammar components appearing only later. Young children's narratives are primarily descriptive, without temporal or clausal links, and are referred to as descriptive sequences (Berman & Slobin, 1994; Nicolopoulou & Richner, 2007; Peterson & McCabe, 1983; Stein & Policastro, 1984; Trabasso & Rodkin, 1994; Westby, 2012). In parallel with descriptions, sequences appear in which simple

events are chained together. These are generally referred to as action sequences and consist of routine actions with occasional usage of general connectives such as and. Older children begin to use reactive sequences, establishing causal relations between events (Berman & Slobin, 1994; Peterson & McCabe, 1983; Stein & Albro, 1997; Stein & Policastro, 1984; Westby, 2012). The most advanced structures to appear are GAO sequences. Thus, as children grow older, they move from descriptive and action sequences to complex episodic organization involving causal connections, in which the characters' reactions are made explicit to the listener (Bohnacker, 2016).

Although most studies have found differences in macrostructural abilities across children aged three to nine (e.g., Berman & Slobin, 1994; Bohnacker, 2016; Lindgren, 2018; Soodla & Kikas, 2010), some studies have reported contradictory results. In a longitudinal study on Swedish-speaking children using the MAIN to assess narrative, Lindgren (2019) found differences in macrostructure production between ages four and five, but no differences between ages five and seven. Similar results were obtained in a longitudinal study by Blom and Boerma (2016) using the same narrative procedure in Dutch: they found no improvements in macrostructural abilities from ages six to seven.

Microstructural Abilities

Microstructure analysis focuses on the internal linguistic organization of the story. Different aspects of microstructure correspond to several relatively independent dimensions, such as productivity, lexical diversity, and syntactic complexity. Over the last few decades, a variety of measures have been proposed to assess the quality of narratives at the microstructure level. Productivity refers to the amount of language (words or utterances) produced in a narrative sample (Leadholm & Miller, 1992). Measures of productivity include the total number of communication units (C-units; Loban, 1976), the total number of terminable units (T-units; Hunt, 1966) and the total number of words (TNW; Hunt, 1970). Lexical diversity pertains to the number of different words (NDW) used in a sample (Yu, 2009). The more varied the vocabulary that a child produces, the greater their lexical diversity. Lexical diversity has commonly been measured in terms of NDW and type-token ratio (TTR), which is calculated as the total number of unique words (types) divided by the total number of words (tokens). The NDW has been treated as a measure of productivity (e.g., Justice et al., 2006; Muñoz et al., 2003) as well as of lexical diversity (e.g., Heilmann et al., 2010; Watkins et al., 1995), depending on the method of its calculation (full-length text vs. limited number of words/utterances). Previous research has shown that both NDW and TTR are affected by the length of the language sample (for a review, see, Malvern et al., 2004), leading researchers to propose more sophisticated measures that take text length into account, such as measure D (Malvern &

Richards, 1997) and moving-average type-token ratio (MATTR; Covington & McFall, 2010). Both MATTR and D perform better than TTR (e.g., Fergadiotis et al., 2015 for MATTR; Jarvis, 2002; Malvern et al., 2004 for D).

Measures of productivity and lexical diversity have often been used in the context of narrative development to assess children's vocabulary abilities (Fergadiotis et al., 2015; Heilmann et al., 2010;). As for children aged from four to 10 years, their productivity in terms of TNW and their lexical diversity in terms of NDW increase (Mäkinen et al., 2014; Westerveld et al., 2004). Justice et al. (2006) found a linear increase in NDW and TNW among English children aged five to 10 years, yet a gradual decrease in both measures beyond 10 years.

Narrative analysis also provides information about growth in syntax, which is generally analyzed in terms of syntactic complexity (see Frizelle et al., 2018). Syntactic complexity refers to the range of syntactic structures and the degree of sophistication of those structures in language production (Ortega, 2003). It can be defined as a multidimensional construct depending on the sources of complexity that can be measured (Norris & Ortega, 2009). Subconstructs of syntactic complexity include overall complexity, complexity via subordination, and subclausal complexity. Overall syntactic complexity is quantified in terms of global or generic metrics of language complexity, among which the most extensively used are the mean length of C-unit (MLCU; Heilmann et al., 2010; Loban, 1976) and the mean length of utterance measured in words (MLUw; Bishop & Donlan, 2005; Frizelle et al., 2018). In addition to these global measures of length, more sophisticated measures can be used to analyze syntactic complexity at the subordinate or subclause level, for example, clausal density (CD), calculated as the total number of main and subordinate clauses divided by the total number of C- or T-units (Gutierrez-Clellen & Hofsteter, 1994; Mäkinen et al., 2014). The CD is affected by variations in the number of subordinate clauses, so researchers may also calculate the mean length of clause (MLC) as a more specific measure of complexity on the subclausal level. The complexity captured by MLC is narrower than that captured by global length-based measures (Bedeković et al., 2021; Norris & Ortega, 2009): it calculates the use of fine-grained syntactic structures or low-frequency structures (Nippold, 2016) such as nominalizations, pre- or postmodification via nonfinite verbs, prepositional phrases, and adverbs.

The use of complex syntax emerges already in preschool-age children (Diessel & Tomasello, 2000; Hržica & Ordulj, 2013), and continues to develop, albeit more slowly, through childhood, adolescence, and early adulthood (Nippold, 2016). For example, Heilmann et al. (2010) noticed that English children aged five to seven years (Mage = 6;0) rarely used complex syntax, while Mäkinen et al. (2014) found that seven- and eight-year-old Finnish children produced syntactically more complex narratives than four-year-olds, as measured in terms of MLCU and CD. Similar results were found for other languages and age groups (Gutierrez-Clellen & Hofstetter, 1994; Justice et al., 2006; To et al., 2014). On the other hand, Frizelle et al. (2018) found no differences in MLUw

or CD between English children aged seven or 12 years. As Loban (1976) and Scott (1988) noticed, late school-age children use more non-finite verbs than subordinate clauses, and they also use other low-frequency syntactic structures in order to express their ideas more concisely. This further implies that measures of sentence length do not always reflect more advanced syntactic development (Nippold, 2016). For this reason, Norris & Ortega (2009) promote MLC as a way to capture more subtle changes in syntactic development. However, we are unaware of studies on narrative development using this measure.

The Relationship Between Microstructure and Macrostructure

The two levels of narrative analysis, micro- and macrostructure, are often treated as distinct aspects of underlying narrative competence. It has been proposed that macrostructure is the underlying structure story (Berman & Slobin, 1994; Hughes et al., 1997), which is somehow independent of the child's linguistic development (e.g., Kupersmitt & Berman, 2001; Westby et al., 1989). In other words, it has been suggested that narrative skills rely on general cognitive abilities (e.g., Trabasso & Nickels 1992; van den Broek, 1997) or that they are guided by sociolinguistic processes (e.g., Eaton et al., 1999; Peterson & McCabe, 2004). Consequently, the results of the two levels of narrative analysis are usually reported separately. However, some studies have examined the relationship between micro- and macrostructure, suggesting that individual variations in lexical and syntactic content co-occur with variations in story structure (e.g., Justice et al., 2006; Owens, 1999). Some have focused on specific linguistic forms, for example, connectives, verb tense, relative clauses, or specific lexical elements (e.g., Berman & Slobin, 1994), while others have taken into consideration general measures of microstructure, namely measures of productivity, lexical diversity, syntactic complexity, and the use of cohesive devices (e.g., Bishop & Donlan, 2005). Still other studies have differentiated between story structure and story or episodic complexity, claiming that these two components might be related to microstructural elements in different ways. Story complexity seems to be more universal and less dependent on children's linguistic skills, in contrast to general story structure, which is more language-dependent (see, Bohnacker, 2016; Gagarina et al., 2012; Gagarina, Klop, Kunnari, et al., 2019; Gagarina et al., 2015).

The Relationship Between Macrostructure and Lexical Diversity

Several studies have analyzed the relationship between macrostructure and lexical diversity, while also including another aspect of microstructure: productivity. Heilmann et al. (2010), who treated NDW as a measure of lexical diversity and TNW as a measure of productivity, found that both NDW and TNW were correlated, and were significant predictors of children's macrostructure.

Macrostructure was measured using the narrative scoring scheme (NSS), a metric that includes the basic components of the story grammar (introduction, conflict, and conclusion), the use of metacognitive and metalinguistic verbs, and cohesive elements according to Halliday and Hasan (1976). Terry et al. (2013) also found macrostructure to correlate with lexical diversity, as measured in terms of NDW, and with productivity, as measured in terms of TNW and the number of C-units. They measured macrostructure using the NSS, proposed by Heilmann et al. (2010), and high-point analysis (Peterson & McCabe, 1983). Mäkinen et al. (2014) found NDW to be a significant predictor of the amount of relevant information used in narratives (event content), although they used NDW as a measure of productivity rather than lexical diversity. The narratives in that study were divided into units of information that together constituted the overall plot and were components of a coherent story.

The Relationship Between Macrostructure and Syntactic Complexity

Studies of the relationship between macrostructure and syntactic complexity have been less conclusive than studies of the relationship between macrostructure and lexical diversity. Bishop and Donlan (2015) showed a close relationship between MLU and measures of content (main ideas and the use of cognitive terms), and Terry et al. (2013) obtained similar results for the MLCU. However, Heilmann et al. (2010) and McFadden and Gilliam (1996) found no connection between macrostructure and syntactic complexity, which the respective studies measured in terms of mean length of C-units in words or the mean length of T-units in morphemes. Other studies using measures of syntactic complexity other than those mentioned above have also provided evidence for a relationship between syntactic complexity and story structure. Liles et al. (1995), who measured episode structure and cohesiveness (achieved through subordinate structures), showed that these two aspects are related. Similarly, Allen et al. (1994) found that higher syntactic complexity (measured as the number of complex sentences) led to more globally organized narratives by six- and eight-year-olds.

Age-Related Differences in the Relationship Between Microstructure and Macrostructure

Few studies have focused on age-related differences in the relationship between micro- and macrostructure, even though many studies have suggested that such differences likely exist (e.g., Berman, 2009). For example, Berman (2009) found that relationships between linguistic and macrostructure variables change with age (see also Aksu-Koç & Aktan-Erciyes, 2018). In fact, Berman and Slobin (1994) first suggested this idea in their analysis of narrative development by children speaking English, German, Hebrew and Spanish. Three- or four-year-olds produce narratives in fragmented descriptive sequences without explicit causal

or temporal marking. By the end of preschool age, children relate events in terms of causality and temporality, and GAO sequences begin to occur. As children grow older, their narratives become more advanced in both macrostructure and syntactic complexity, and the relationship between the two increases (e.g., Bishop & Donlan, 2005; Heilmann et al., 2010). For example, Bishop and Donlan (2005) found a correlation between the use of complex syntax (MLU) and macrostructure in the narratives of children aged seven to nine. Heilmann et al. (2010) found lexical diversity to be the most robust predictor of macrostructure skills among five- to seven-year-olds and proposed, based on the results of Bishop and Donlan (2005), that the most robust predictor among seven- to nine-year-olds may be syntax. Unfortunately, many previous studies on the relationship between micro- and macrostructure did not examine age-related differences, even if they included children of different ages (e.g., Mäkinen et al., 2014).

Furthermore, previous studies on the relationship between micro- and macrostructure have differed substantially in their methods. For example, studies have used different elicitation material (Frog, where are you?, Mayer, 1969, in Heilmann et al., 2010, but Cat Story in Mäkinen et al.,2014) and measures of macrostructure. Heilmann et al. (2010) used NSS, Terry et al. (2013) used NSS and high-point analysis, and Mäkinen et al. (2014) focused on the amount of relevant information used in the story. Therefore, it is difficult to compare the results of those studies, preventing us from understanding the role of age in the relationship between micro- and macrostructure.

Aims and Hypotheses

Previous research has indicated a relationship between story structure and lexical and syntactic skills (e.g., Berman & Slobin, 1994; Fernández, 2003; Mäkinen et al., 2014). There have been some suggestions that the role of these microstructure skills in building the narrative macrostructure changes with age (e.g., Heilmann et al., 2010; Mäkinen et al., 2014). However, previous studies have used different measures of the two levels of story structure, as well as different procedures of data elicitation, analysis, and evaluation. In addition, previous studies did not directly compare children of different ages, preventing the identification of possible developmental differences in the relationship between micro- and macrostructure. Thus, studies on potential age-related differences in this relationship are difficult to compare with one another and are inconclusive.

In the present study, we examined the narrative abilities of six- and eight-year-old monolingual speakers of Croatian using the same elicitation material and procedure for all participants. This study had two aims: (a) to test for age-related differences in narrative abilities at the macro- and microstructural levels, and (b) to examine which aspects of microstructure might explain individual differences in narrative macrostructure at different ages. Based on previous research, we hypothesized that:

1. Eight-year-olds would achieve a higher MAIN story structure score than six-year-olds, as well as higher lexical diversity measured by D and greater syntactic complexity in terms of the mean length of clause and CD.
2. Different aspects of microstructure would be positively related to the story structure at different ages. Of all microstructure aspects, only the lexical diversity score would contribute significantly to explaining individual differences in the story structure of six-year-olds, whereas syntactic complexity scores would contribute significantly to the overall story structure of eight-year-olds.

Method

Participants

We collected oral narrative samples from 93 monolingual children speaking Croatian. Children were recruited from kindergartens and one elementary school in Zagreb, Croatia. All children were tested using two standardized language tests, the Croatian version of the Peabody Picture Vocabulary Test (PPVT-III-HR; Dunn et al., 2010) and the Croatian version of the Test for the Reception of Grammar (TROG-2: HR; Bishop et al., 2013). Only narratives from typically developing children (without language delays) were eligible for inclusion in this study, so all children who scored two SDs below their age group average were excluded from the sample. In addition, two children were excluded during the data analysis phase of the study because the analysis indicated the presence of multivariate outliers in their results. Thus, the final sample in the present study consisted of 89 children.

The children were divided into two groups based on their age. Average chronological age was six years and three months in the younger group (n = 37), and eight years and five months in the older group (n = 52). Each group was balanced for gender (see Table 1 for more detailed information on age and gender). The two groups were also balanced in terms of the educational level of their parents: in both groups, approximately 40% of mothers had a secondary education and 60% had a tertiary education, while nearly equal proportions of fathers had a secondary or tertiary education. The study was approved by the Croatian Ministry of Science and Education, as well as by the participating institutions. The children agreed orally to participate in the activity, and parents signed written informed consent.

Instruments

Narrative samples were collected using the MAIN (Gagarina et al., 2012; Gagarina, Klop, Kunnari, et al., 2019), adapted for Croatian (Hržica & Kuvač

Table 1. *Demographic Characteristics of Participants (N = 89)*

		Chronological age			Gender	
Age group	*n*	*M*	*SD*	Age range	Male	Female
Six-year-old children	37	6;3	0;6	5;0–7;0	17	20
Eight-year-old children	52	8;5	0;4	7;8–9;0	29	23

Kraljević, 2012). The MAIN is part of the "Language Impairment Testing in Multilingual Settings" (LITMUS) battery, developed within COST Action IS0804 "Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment." It is suitable for assessing narrative skills in children aged three to 10 years. The instrument contains a total of four stories designed to assess comprehension and production of narratives using standardized procedures. In the present study, only narrative production was assessed. To adhere to the original MAIN procedure, the children were given one of the two stories for the storytelling task: half of the participants were presented with the Baby Goats story, and the other half with the Baby Birds story. Both stories consist of six colored pictures based on a multidimensional model of story organization. They are matched in terms of the number of main characters and GAO sequences. Also, the stories in the MAIN are controlled for linguistic and cognitive complexity, correspondence in micro- and macrostructure, as well as cultural appropriateness and robustness (Gagarina et al., 2012; Gagarina, Klop, Kunnari, et al., 2019). Several studies have shown no significant difference in macrostructure between Baby Goats and Baby Birds (e.g., Bohnacker, 2016; Lindgren 2018), although some work suggests that such differences may exist (e.g., Gagarina, Bohnacker, & Lindgren, 2019; Lindgren, 2019).

Procedure

Children were tested individually in a quiet room, with only the participant and the interviewer present. The study procedure differed from the procedure described in the MAIN manual (Gagarina et al., 2012; Gagarina, Klop, Kunnari, et al., 2019) because the assessment was conducted on a computer screen instead of on paper, although it adhered to the currently recommended procedures for online testing available on the MAIN website (e.g., Hamdani et al., 2021). Effects of shared knowledge and joint attention during elicitation were controlled for, as suggested by van der Lely (1996) and in accordance with the standardized MAIN procedure. This was accomplished in two ways: the children were allowed to choose the story and only they could see the picture prompts as they told the story.

Each child selected a story from four differently colored squares presented on a 15.60 in. screen, under conditions designed to lead the child to believe that the interviewer did not know which story the child would choose. The child chose

by clicking on one square, with assistance from the interviewer if necessary. In reality, the act of selection was simulated because the interviewer had preselected the story. Next, a set of six pictures appeared on the screen, three in the upper half and three in the lower half. The child, but not the interviewer, could see the pictures on the computer screen, leading the child to believe that the interviewer was unfamiliar with the story. The child had to press a key on the keyboard to scroll through the picture sequence. With each key press, a pair of pictures was revealed, simulating the offline procedure. During the telling of the story, the child could simultaneously look at the first and second picture, then the third and fourth, and finally, the fifth and sixth. All pictures were the same in size. In contrast to the printed version of the MAIN and other online versions that have been reported (e.g., Hamdani et al., 2021), the children in the current study were unable to return to earlier pictures during the testing, but instead could view only the two active pictures. The same procedure was used to test both age groups.

The narratives were audio-recorded and then transcribed and coded using the Codes for Human Analysis of Transcripts (CHAT) system and the Computerized Language Analysis of Transcripts (CLAN) program. The CHAT and CLAN form parts of the Child Language Data Exchange System (CHILDES; MacWhinney, 2000). Transcription and coding were done by native monolingual Croatian speakers trained in transcription and coding for language research. Repetitions, revisions, fillers, codeswitching, nonwords, onomatopoeia, and hesitations were coded using special markers and excluded from further analysis. All language samples successfully passed the CHECK option in the CLAN program. Seventeen randomly selected samples (18%) were retranscribed independently by two transcribers and word-level agreement with the corresponding original was at least 89%, suggesting a high degree of reliability.

Narrative Measures and Data Analysis

Macrostructure

Children's narrative abilities at the macrostructural level were assessed using the MAIN scoring procedure (Gagarina et al., 2012; Gagarina, Klop, Kunnari, et al., 2019). The MAIN macrostructure consists of two principal elements: setting and three episodes. Each episode includes five structural components: internal state as an initiating event, goal, attempt, outcome, and internal state as a reaction. All scoring was performed by the same investigator with extensive experience in testing with the MAIN protocol. Children were given points for every macrostructural component they expressed in their narrative. A child could receive 2 points for the production of setting (one for time, the other for place), and 1 point for the production of each of the other story elements (a total of 5 points in each episode). The points were summed to obtain an overall story

structure (macrostructure) score, which could be a maximum of 17. Another rater independently scored the macrostructure in 20 randomly selected narratives (21.5%), and a two-way mixed-effects model based on single ratings and absolute agreement was used to assess inter-rater reliability. The intraclass correlation coefficient showed excellent reliability, with ICC = .98 (95% CI [.95, .99]).

The two picture prompts in the present study were expected to be comparable at the macrostructural level (Bohnacker, 2016; Lindgren, 2018). Although previous studies showed mixed results (e.g., Gagarina, Bohnacker, & Lindgren, 2019; Lindgren, 2019), no significant difference in story structure scores was found between the two stories in the present study, $t(34) = -1.43$, $p = .161$ in the six-year-old group, and $t(50) = -0.73$, $p = .471$ in the eight-year-old group. Therefore, we averaged the two scores within each age group.

Microstructure

Several language measures of microstructure were applied to assess lexical diversity and syntactic complexity.

Lexical diversity. We analyzed lexical diversity using the measure D (Malvern et al., 2004). D was selected from among various measures of lexical diversity because previous research on Croatian speakers showed that this measure can predict receptive lexicon score (Hržica & Roch, 2020) and differentiate between children showing typical language development or developmental language disorder (Hržica et al., 2019). D is based on mathematical modelling of the decreasing TTR curve and it is calculated automatically in CLAN (MacWhinney, 2000). The program calculates the TTR curve of the real sample and then compares it to an ideal curve that the model generates based on the probabilities of the highest and lowest TTRs for the sample. Thus, D provides valid assessment of lexical diversity even when language samples differ in length (Jarvis, 2002; McCarthy & Jarvis, 2010). A higher score indicates higher lexical diversity.

Syntactic complexity. Syntactic complexity was assessed in terms of CD and the MLC. CD has been found to capture age-related differences in complexity via subordination (e.g., Heilmann et al., 2010; Liles et al., 1995), and it is calculated manually by extracting all the main and subordinate clauses from a child's narrative and dividing that number by the total number of C-units. Thus, CD indicates the extent to which C-units within the particular narrative sample contain subordinate clauses. For example, a score of 2.0 means that C-units contain an average of two clauses (one main clause and one subordinate clause), while a score of 1.0 means that all C-units in a sample are simple, one-clause constructions (Scott & Stokes, 1995). The MLC was used to assess complexity on the subclausal level via phrasal elaboration, for example, via addition of adjectives, adverbs, nominalizations, nonfinite clauses, and prepositional phrases (Norris & Ortega, 2009). The MLC was calculated semi-automatically as follows: first, all the subordinate clauses were extracted manually, then CLAN was used to calculate the total number of

words in the language sample, and finally, the MLC was calculated as the mean number of words per clause (main or subordinate). Higher CD or MLC indicates greater syntactic complexity. An independent rater was asked to manually extract the subordinate clauses in 20 randomly selected samples (21.5%), and a two-way mixed-effects model based on single ratings and absolute agreement was used to assess inter-rater reliability. The intraclass correlation coefficient showed excellent reliability, with ICC = .99 (95% CI [.97, 1.00]).

Data Analysis

All analyses were performed in SPSS version 23.0. Before running the main analyses, we screened our data for outliers and tested the normality of the distributions. This was done separately within each age group. To check for univariate outliers, children's individual scores on all variables were standardized via transformation into z scores. Following recommendations from the literature (e.g., Tabachnick & Fidell, 2013), a z score of +/-3.29 ($p < .001$) was chosen as the cut-off point for detecting outliers. All scores on all measured variables were within this range, meaning that no univariate outliers were detected. To identify multivariate outliers, we followed the procedure described in Tabachnick and Fidell (2013). The Mahalanobis distance was used as a statistical indicator of multivariate outliers. It is the distance of each case from the centroid of the remaining cases, where the centroid is the point created at the intersection of the means of all the variables. It can be evaluated using the χ^2 distribution with degrees of freedom equal to the number of explanatory variables (Tabachnick & Fidell, 2013). $\chi^2(3) = 16.27$ ($p < .001$) was used as the cut-off value for multivariate outlier detection. Two cases with a Mahalanobis distance greater than this value were identified as multivariate outliers and excluded from the sample because their scores might distort the regression analysis.

Only the CD distribution showed significant skew, which was in the same direction in both age groups. Since parametric tests are robust to minor deviations from a normal distribution, we decided to use the t test to assess the significance of age-related differences in story structure score, lexical diversity, and syntactic complexity. In addition, we performed the Mann-Whitney U test to check the robustness of the t test results obtained for CD. The results did not differ, therefore, we report only the results of the t test. To examine the relationship between two aspects of microstructure (lexical diversity and syntactic complexity) and macrostructure of narratives, we performed Pearson's correlation and linear regression analyses for the two age groups separately because we hypothesized that the pattern of relationships might differ with age. All assumptions for the linear regression analysis were met: the relationships between all variables were linear, there was no multicollinearity among the explanatory variables (all variance inflation factors were less than 2, and the tolerance values were higher than 0.2), the residuals followed a normal distribution, and were homoscedastic.

Since both age groups were small, which can lead to biased estimates, we used the bootstrap bias-corrected and accelerated procedure with 10,000 samples to obtain 95% confidence intervals for correlation and regression coefficients.

Results

Age-Related Differences in Story Structure, Lexical Diversity and Syntactic Complexity

Descriptive statistics for scores on overall story structure (macrostructure), lexical diversity, and syntactic complexity for the two age groups are presented in Table 2. Mean and median values are presented for all measures, although only the CD distribution showed a significant skew. The story structure score for six-year-olds indicated relatively poor overall macrostructure in the narratives. Children at this age produced around 31% of the scored macrostructural components (M = 5.3, SD = 1.93, maximum possible score = 17). In contrast, eight-year-old children scored higher on narrative macrostructure. They produced around 47% of scored macrostructural components (M = 8.0, SD = 2.08, maximum possible score = 17). The t test confirmed that this difference was significant with a large effect size, $t(87) = -6.20$, $p < .001$, $d = 1.34$, 95% CI [0.87, 1.81]. Eight-year-olds also showed significantly higher lexical diversity ($M = 18.1$, $SD = 4.83$) than six-year-olds ($M = 13.4$, $SD = 5.74$), $t(87) = -4.14$, $p < .001$, $d = 0.88$, 95% CI [0.43, 1.33]. Although both age groups produced a fair number of subordinate clauses in their narratives, eight-year-olds had a significantly higher mean CD score ($M = 1.2$, $SD = 0.13$) than six-year-olds (M = 1.1, SD = 0.09), Welch's $t(86.67) = -4.07$, $p < .001$, $d = 0.85$, 95% CI [0.42, 1.28]. The effect size for both lexical diversity and CD was greater than d = 0.80, indicating a large effect size (Cohen, 1988). On the other hand, we found no significant difference in the MLC, which was around five words in both age groups, $t(87) = -1.46$, $p = .147$, $d = 0.31$, 95% CI [-0.12, 0.74].

Table 2. *Mean and Range of Scores for Macrostructure and Aspects of Microstructure in Six-Year-Olds (n = 37) and Eight-Year-Olds (n = 52)*

	Six-year-olds					Eight-year-olds				
	M	*SD*	*Mdn*	*IQR*	Range	M	SD	Mdn	IQR	Range
Macro	5.3	1.9	5.0	3.0	1–9	8.0	2.1	8.0	3.5	4–13
D	13.41	5.74	12.61	9.32	4.00–27.08	18.06	4.83	18.34	5.90	8.05–30.42
MLC	5.00	0.70	5.07	0.88	3.00–6.67	5.22	0.67	5.20	0.88	3.75–6.70
CD	1.07	0.09	1.04	0.12	1.00–1.33	1.17	0.13	1.13	0.17	1.00–1.50

Note. Macro = overall story structure score (max = 17 points); D = measure of lexical diversity; MLC = mean length of clause; CD = clausal density; IQR = interquartile range.

The Relationship Between Microstructure and Macrostructure

To examine the relationships between different aspects of microstructure and macrostructure in children's narratives, we calculated their correlations and performed regression analysis on the two age groups. Table 3 presents Pearson's correlations and corresponding 95% CIs for relationships in the six-year-old group. Lexical diversity showed a significant positive correlation with overall story structure score, $r(35) = .57$, $p < .001$, and with CD, $r(35) = .33$, $p = .046$. Other correlations were not significant, although the correlation between CD and macrostructure approached significance, $r(35) = .31$, $p = .064$. The magnitude of this correlation was similar to that between lexical diversity and CD, so its lack of significance may reflect low statistical power due to the small sample and correspondingly large standard error.

Next, linear regression was conducted to determine which lexical and syntactic aspects of microstructure might explain the individual differences in overall story structure score. All variables together accounted for 29% of the variance in narrative macrostructure ($R = .59$, adj. $R^2 = .29$, $p = .002$; see Table 4). Only lexical diversity showed a significant contribution in explaining the variance in macrostructure ($\beta = .50$, $p = .003$).

We carried out the same types of analyses in the group of eight-year-old children. Macrostructure showed significant positive correlations with lexical diversity, $r(50) = .43$, $p = .002$, CD, $r(50) = .37$, $p = .006$, and the MLC, $r(50) = .29$, $p = .040$ (see Table 5). However, 95% CIs obtained for the correlation with the MLC contained zero, indicating a nonsignificant relationship, so this result should be interpreted with caution. Lexical diversity also showed a significant positive correlation with MLC, $r(50) = .32$, $p = .021$. The correlation between lexical diversity and CD approached the significance level, $r(50) = .27$, $p = .052$ and 95% CIs obtained by bootstrapping did not contain zero, indicating a significant relationship. Nevertheless, these results should be interpreted with caution due to the small sample size and large standard error.

Linear regression analysis showed that all variables together accounted for 24% of the variance in narrative macrostructure ($R = .53$, adj. $R^2 = .24$, $p = .001$;

Table 3. *Correlations Between Variables in the Group of Six-Year-Old Children (n = 37)*

Variable	Macro	D	MLC
Macro	-		
D	.57*** [.30, .78]	-	
MLC	.18 [−.19, .51]	.20 [−.15, .53]	-
CD	.31 [.00, .55]	.33* [.09, .60]	−.12 [−.40, .15]

Note. Macro = overall story structure score; D = measure of lexical diversity; MLC = mean length of clause; CD = clausal density. 95% CIs obtained by bootstrapping are shown in brackets.
*** $p < .001$, * $p < .05$.

Table 4. *Summary of Regression Analysis in the Six-Year-Old Group (n = 37)*

Variable	β	*b* (*SE*)	*b* 95% CI[a]	Model summary
D	.50**	0.17 (0.05)	0.09, 0.25	$R = .59$
MLC	.10	0.27 (0.41)	-0.50, 0.87	adj. $R^2 = .29$
CD	.15	3.32 (3.28)	-3.16, 8.47	$F(3, 33) = 5.97$**

Note. D = measure of lexical diversity; MLC = mean length of clause; CD = clausal density.
[a] 95% CIs were estimated using the bootstrap BCa method with 10,000 samples.
** $p < .01$.

Table 5. *Correlations Between Variables in the Group of Eight-Year-Old Children (n = 52)*

Variable	Macro	D	MLC
Macro	-		
D	.43** [.20, .62]	-	
MLC	.29* [-.05, .59]	.32* [.03, .57]	-
CD	.37** [.10, .61]	.27 [.08, .46]	.07 [−.18, .31]

Note. Macro = overall story structure score; D = measure of lexical diversity; MLC = mean length of clause; CD = clausal density. 95% CIs obtained by bootstrapping are shown in brackets.
** $p < .01$; * $p < .05$.

Table 6. *Summary of Hierarchical Regression Analysis in the Eight-Year-Old Group (n = 52)*

Variable	β	*b* (*SE*)	*b* 95% CI[a]	Model summary
D	.29*	0.13 (0.06)	0.03, 0.25	$R = .53$
MLC	.17	0.54 (0.40)	−0.40, 1.47	adj. $R^2 = .24$
CD	.28*	4.39 (1.98)	0.24, 8.81	$F(3, 48) = 6.22$**

Note. D = measure of lexical diversity; MLC = mean length of clause; CD = clausal density.
[a] 95% CIs were estimated using the bootstrap BCa method with 10,000 samples.
** p < .01; * p < .05.

see Table 6). Both lexical diversity (β = .29, $p = .034$) and CD (β = .28, $p = .031$) contributed significantly to explaining variance in macrostructure.

Discussion

In investigating Croatian monolingual children aged six years ($M = 6;3$) and eight years ($M = 8;5$), this study had two aims: (a) to examine age-related differences in macrostructure production and lexical and syntactic abilities, and (b) to test the relationship between micro- and macrostructure in their narratives. Eight-year-olds in this study achieved a higher average story structure score than six-year-olds,

which is in line with Mäkinen et al. (2014), who also reported age-related differences in children's story structure. The overall story structure of eight-year-old Finnish children was more advanced than that of six-year-olds. However, our results differ from those of Lindgren (2019), who found in a longitudinal study that Swedish children produced around 43% of the scored macrostructural components at age 5;10 and about 47% at age 7;4, with no significant difference between the two age groups. Similar results were obtained by Blom and Boerma (2016) in another longitudinal study, which examined narratives of Dutch-speaking children from ages 5;9 to 6;9. Our six-year-old children achieved somewhat lower results on overall macrostructure than Swedish children of a similar age, while our eight-year-olds achieved the same results as seven-year-old Swedish children. However, comparisons across these studies should be made with caution, since Blom and Boerma (2016) and Lindgren (2019) examined macrostructure development in the same group of children over time, whereas the present study compared two different groups of children. Thus, we cannot exclude that at least some of the differences between our two age groups may reflect generational differences unrelated to developmental changes in narrative acquisition.

The range of the results on lexical diversity was similar in both groups of children, but the average score was significantly higher for eight-year-olds than for six-year-olds. The findings of our study regarding age-related differences in lexical abilities are consistent with previous studies that also reported improvements in children's lexical abilities, as measured based on NDW (Justice et al., 2006; Mäkinen et al., 2014; Westerveld et al., 2004). Our study extends the evidence in support of age-related differences by using measure D, which has shown greater validity than NDW in measuring lexical diversity in language samples of different length (Jarvis, 2002; McCarthy & Jarvis, 2010). To the best of our knowledge, D has not been used so far in research on children's narratives.

The syntactic complexity of narratives in terms of CD was low in both age groups in our study, but higher for eight-year-olds. The difference between the two age groups was significant with a large effect size. The eight-year-olds produced about 17% of the C-units that contained subordinate clauses, and the six-year-olds only 7%. Similar results were obtained by other studies that assessed syntactic complexity in children's narratives (Gutierrez-Clellen & Hofstetter, 1994; Heilman et al., 2010; Justice et al., 2006; Mäkinen et al., 2014; but see Frizelle et al., 2018). This is not so surprising given the fact that syntax still develops in school-age children and even later (Nippold, 2016). Nevertheless, alternative explanations of the observed age-related differences are possible. First, the elicited stories in the present study were relatively short, and although the MAIN controls for nonlinearity of action, there is a limited number of situations in which subordinate clauses are needed. Second, languages vary in the number and types of clause combining they prefer (Berman & Slobin, 1994; Kuiken et al., 2019), and subordinate clauses may be less frequent in Croatian than in other languages (but see Mäkinen et al., 2014). For example, the frequency of

subordinate clauses in English contributes to greater syntactic complexity than Dutch (Hannay & Mackenzie, 1996). Unfortunately, there is no previous research comparing Croatian with other languages, but Mamula and Trtanj (2018), who elicited narratives using Frog, where are you? (Mayer, 1969), showed that about 10% of C-units produced by eight- or ten-year-olds contain subordinate clauses. Since previous studies using the MAIN did not measure CD, it is difficult to determine whether the low CD in the current study results from the characteristics of the elicitation material or features of the language.

The two age groups in the current study did not differ significantly in syntactic complexity as measured in terms of the MLC. Children of both ages produced clauses that were an average of five words long. To the best of our knowledge, no research on children's narratives has used the MLC as a measure of syntactic complexity. Since the MLC differs from other length-based measures (e.g., mean length of utterance, mean length of C-unit), it is not possible to compare the present results to previous studies. The MLC indicates increases in clause length through pre- or post-modification involving infinitives, gerunds, nominalizations, and prepositional phrases. Other length-based measures indicate increases involving pre- or post-modification as well as subordination. It must be noted that due to its morphological complexity, Croatian has fewer function words: cases are often used instead of prepositions to signal relationships between words (e.g., čaša vod-e "glass of water"), it is a pro-drop language (e.g., živi-m "I live"), it does not have articles, and aspect and actionality are expressed through derivation. Expression of syntactic relations by bound morphemes may prevent length-based measures of syntactic complexity from detecting new syntactic relationships and therefore differences between the two age groups.

Several studies have shown that the two levels of narrative structure are interconnected and that language aspects at the lower level explain story structure at the higher level of discourse organization (Bishop & Donlan, 2005; Fernández, 2013; Heilmann et al., 2010; Mäkinen et al., 2014). Further, some have suggested that the contribution of different microstructure aspects to narrative macrostructure changes with age (e.g., Heilmann et al., 2010). Therefore, the present study aimed to examine whether the relationship between different language aspects and story structure differs with age. Our results showed that only lexical diversity made a significant contribution to explaining individual differences in the story structure of six-year-old children, whereas both lexical diversity and CD made significant contributions in eight-year-olds. Our results indicate that in older children, lexical abilities remain relevant for the organization of narratives, and that syntactic abilities become important as well. CD appeared to be more relevant than length-based MLC for explaining individual differences in the story structure of eight-year-old children.

A novel finding of the present study is the importance of vocabulary for explaining macrostructure skills in six-year-old children, contrasted with the importance of both vocabulary and syntax for explaining the corresponding skills

in eight-year-old children. However, this finding has been suggested previously. Heilmann et al. (2010) pointed out that preschool children and young school children produce narratives as sequences of events, and only in older school children do narratives become more complex (e.g., hierarchically organized; Berman, 1988). For such narratives, older children must use complex syntax, but before they can do so, they must rely on their vocabulary skills when organizing the story (Heilmann et al., 2010). The results with the two age groups of children in the present study point to an interesting developmental trend: children of different ages may rely on different sets of language skills for building macrostructure (vocabulary vs. vocabulary and syntax). We might expect that for even older children, a new set of skills plays the most important role, for example, discourse-building skills. This possibility emerges from previous studies on the development of narrative abilities that have shown lower, rather than higher, productivity or even syntactic complexity in older than in younger children. For example, Justice et al. (2006) reported a decrease in the number of C-units, number of words, mean length of T-unit and number of subordinate clauses in children aged 11 and 12, and similar results were reported in other studies (e.g., Kuvač Kraljević et al., 2016; Mäkinen et al., 2014; Trtanj, 2015). Justice et al. (2006) attributed this result to testing, skills, or motivation. However, an alternative explanation might be that the improvement of discourse skills in older children allows them to produce complex narratives despite the shorter language samples. This implies that measures focused on coherence and cohesion may better predict variance in the macrostructure of narratives produced by older children.

Like any study, the present work has some limitations. First is the small sample in both age groups, which reflects the desire to avoid excessive pretesting of children, the need to keep the amount of transcription manageable, and other constraints. According to some suggestions in the literature (e.g., Tabachnick & Fidell, 2013), samples for regression analysis should be approximately two to three times larger than in our study. We have employed the bootstrap bias-corrected and accelerated the procedure with 10,000 samples to obtain more accurate 95% confidence intervals for correlation and regression coefficients. However, this does not alter the fact that the current sample may have been too small to provide adequate power to detect some significant differences. As a result, some of our nonsignificant results may be the consequence of inadequate statistical power, and so should be interpreted with caution. These findings should be tested further in larger and more representative samples (both groups of children in the current study were from a convenience sample and from only one region of the country). We hypothesized that the pattern of relationships between different language abilities and story organization at the macrostructural level may differ at different developmental stages. Therefore, we conducted analyses separately for the two age groups. Our results indeed suggest age-related differences in the relationship between macrostructure and different aspects of microstructure, but future studies with larger samples of children should further examine the moderating effects of

age on the relationship between micro- and macrostructure. Because the present study was cross-sectional, we can make conclusions only about age-related differences and not about developmental changes. Future studies should employ longitudinal designs to gain better insight into how the relationship between lower-level language skills and higher-level discourse organization changes over time. Cross-linguistic differences between Croatian and other languages studied so far could make it difficult to compare our findings directly with those in the literature. Some previous studies have opted to apply lexical diversity measures to "lemmatized" narrative samples (e.g., Mäkinen et al., 2014), in which inflectional forms of a word have been reduced to a common base form or lemma, but some did not (e.g., Bishop & Donlan, 2014). When explaining the results of Mäkinen et al. (2014), Aksu-Koç and Aktan Erciyes (2018) stated that in agglutinating languages in which derivational morphology carries both syntactic and semantic load, lexical diversity should be treated as a measure that also reflects syntactic complexity. This claim could be expanded to fusional languages with high paradigmatic richness (Dressler, 2003). If the same lemma has many morphological types, measures of lexical diversity may also reflect the (morpho)syntactic abilities of the speaker, as opposed to languages with low paradigmatic richness.

In conclusion, the current study adds to the literature on the relationship between microstructure aspects and the macrostructure of narratives. Contrary to previous work, this study compared the abilities of children at different ages and examined whether the relationship between micro- and macrostructure differs with age. Our results suggest that children at different developmental stages rely on different microstructure skills when organizing narratives. Six-year-old children appear to rely primarily on lexical diversity, whereas eight-year-old children appear to rely on both lexical diversity and syntactic complexity. Since vocabulary develops earlier, it is not surprising that younger children lean primarily on their lexical abilities when organizing their narratives. As children get older, they start to use other language abilities in their discourse as well.

Acknowledgments

This work was supported by the Croatian Science Foundation under the project entitled Multilevel approach to spoken discourse in language development (UIP-2017-05-6603), by Europen Commision Seventh Framework Programme under the project Language Dominance of Bilingual Speakers Perceived as Balanced (LADOBI, PISCRTJQ4Y, Piscopia Marie Curie Action (COFUND)), by the Arts and Humanities Research Council under the project entitled Feast and Famine Project: Confronting Overabundance and Defectivity in Language (AH/T002859/1) and by the COST Action under the project NexusLinguarum – European network for Web-centred linguistic data science (CA18209). Sara Košutar was supported by the project Young Researchers' Career Development project – Training of New Doctoral Students. Any opinions, findings, conclusions, or recommendations presented in this manuscript are those of the author(s) and do not necessarily reflect the views of the Croatian Science Foundation.

Conflict of Interest Disclosure

The authors declared no potential conflicts of interest.

Funding

Croatian Science Foundation under the project Multilevel approach to spoken discourse in language development (UIP-2017-05-6603), Europen Commission Seventh Framework Programme under the project Language Dominance of Bilingual Speakers Perceived as Balanced (LADOBI, PISCRTJQ4Y, Piscopia Marie Curie Action (COFUND), Framework 7 (FP7)), Arts and Humanities Research Council under the project Feast and Famine Project: Confronting Overabundance and Defectivity in Language (AH/T002859/1), COST Action under the project NexusLinguarum – European network for Web-centred linguistic data science (CA18209).

Research Ethics Statement

Children were included in this study in the course of the two funded research projects. The Ethics Committee of the Faculty of Education and Rehabilitation of the University of Zagreb gave a positive statement about the ethical aspect of the projects. Ministry of Science and Education approved the research. Guidelines outlined in the Declaration of Helsinki and guidelines of the American Psychological Association were taken into consideration when planning and conducting the research.

References

Aksu-Koç, A., & Aktan Erciyes, A. (2018). Narrative discourse: Developmental perspectives. In A. Bar-On & D. Ravid (Eds.), *Handbook of communication disorders* (pp. 329–356). De Gruyter Mouton. https://doi.org/10.1515/9781614514909-017

Allen, M., Kertoy, M., Sherblom, J., & Pettit, J. (1994). Children's narrative productions: A comparison of personal event and fictional stories. *Applied Psycholinguistics, 15*(2), 149–176. https://doi.org/10.1017/S0142716400005300

Bedeković, M., Hržica, G., & Kramarić, M. (2021). Analiza sintaktičke složenosti dječjeg pripovjednog diskursa [Analysis of syntactic complexity of children's narrative discourse]. *Fluminensia, 33*(2), 417–443. https://doi.org/10.31820/f.33.2.8

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). Setting the narrative scene: How children begin to tell a story. In K. E. Nelson, A. Aksu-Koç, & C. E. Johnson (Eds.), *Children's language: Developing narrative and discourse competence* (Vol. 10, pp. 1–30). Lawrence Erlbaum Associates Publishers.

Berman, R. A. (2009). Trends in research on narrative development. In S. Foster Cohen (Ed.), *Language acquisition* (pp. 294–318). Palgrave Macmillan. https://doi.org/10.1057/9780230240780_13

Berman, R., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study.* Lawrence Erlbaum. https://doi.org/10.1017/S0305000900009016

Bishop, D., & Donlan, C. (2005). The role of syntax in encoding and recall of pictorial narratives: Evidence from specific language impairment. *British Journal of Developmental Psychology, 23*(1), 25–46. https://doi.org/10.1348/026151004X20685

Bishop, D., Kuvač Kraljević, J., Hržica, G., Kovačević, M., & Kologranić Belić, L. (2013). *TROG-2:HR Test razumijevanja gramatike* [TROG-2:HR Test for reception of grammar]. Naklada Slap.

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

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

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

Capone Singleton, N., & Shulman, B. B. (2014). *Language development.*

foundations, processes, and clinical applications. Jones & Bartlett Learning.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Covington, M. A., & McFall, J. E. (2010). Cutting the gordian knot: The moving–average type–token ratio (MATTR). *Journal of Quantitative Linguistics, 17*(2), 94–100. https://doi.org/10.1080/09296171003643098

Diessel, H., & Tomasello, M. (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguistics, 11*(1-2), 131–151. https://doi.org/10.1515/cogl.2001.006

Dressler, W. (2003). Degrees of grammatical productivity in inflectional morphology. *Italian Journal of Linguistics, 15*(1), 31–62.

Dunn, L. M., Dunn, L. M., Kovačević, M., Padovan, N., Hržica, G., Kuvač Kraljević, J., Mustapić, M., Dobravac, G., & Palmović, M. (2010). *PPVT-III-HR Peabody slikovni test rječnika* [PPVT-III-HR Peabody picture vocabulary test]. Naklada Slap.

Eaton, J. H., Collis, G. M., & Lewis, V. A. (1999). Evaluative explanations in children's narratives of a video sequence without dialogue. *Journal of Child Language, 26*(3), 699–720. https://doi.org/10.1017/S0305000999003967

Fergadiotis, G., Wright, H. H., & Green, S. B. (2015). Psychometric evaluation of lexical diversity indices: Assessing length effects. *Journal of Speech Language and Hearing Research, 58*(3), 840–852. https://doi.org/10.1044/2015_JSLHR-L-14-0280

Fernández, C. (2013). Mindful storytellers: Emerging pragmatics and theory of mind development. *First Language, 33*(1), 20–46. https://doi.org/10.1177/0142723711422633

Frizelle, P., Thompson, P. A., McDonalnd, D., & Bishop, D. V. M. (2018). Growth in syntactic complexity between four years and adulthood: Evidence from a narrative task. *Journal of Child Language, 45*(5), 1174–1197. https://doi.org/10.1017/S0305000918000144

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*, 1–155. https://doi.org/10.21248/zaspil.56.2019.414

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*, 1–21. https://doi.org/10.21248/zaspil.63.2019.516

Gagarina, N., Klop, D., Tsimpli, I., & Walters, J. (2015). Narrative abilities in bilingual children. *Applied Psycholinguistics, 37*(1), 11–17. https://doi.org/10.1017/S0142716415000399

Gutierrez-Clellen, V. F., & Hofstetter, R. (1994). Syntactic complexity in Spanish narratives: A developmental study. *Journal of Speech & Hearing Research, 37*(3), 645–654. https://doi.org/10.1044/jshr.3703.645

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

Hamdani, S., Kan, R., Chan, A., & Gagarina, N. (2021, January 25–27). Summarizing experience: Identifying bilingual DLD using online testing [Paper presentation]. Workshop on LITMUS-MAIN: "Online elicitation of narrative texts: Summarizing experience and making plans", online.

Hannay, M., & Mackenzie, J. L. (1996). *Effective writing in English: A resource guide*. Martinus Nijhoff.

Heilmann, J., Miller, J. F., Nockerts, A., & Dunaway, C. (2010). Properties of the narrative scoring scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology, 19*(2), 154–166. https://doi.org/10.1044/1058-0360(2009/08-0024)

Hickmann, M. (2003). *Children's discourse: Person, space and time across languages.* Cambridge University Press. https://doi.org/10.1017/S0047404506270349

Hržica, G., Košutar, S., & Kramarić, M. (2019). Rječnička raznolikost pisanih tekstova osoba s razvojnim jezičnim poremećajem [Lexical diversity in written texts of persons with developmental language disorder]. *Hrvatska Revija za Rehabilitacijska Istraživanja, 55*(2), 14–30. https://doi.org/10.31299/hrri.55.2.2

Hržica, G., & Kuvač Kraljević, J. (2012). MAIN - hrvatska inačica: Višejezični instrument za ispitivanje pripovijedanja [MAIN - Croatian version: Multilingual Assessment Instrument for Narratives]. *ZAS Papers in Linguistics, 56*, 323–323.

Hržica, G., & Ordulj, A. (2013). Dvočlane glagolske konstrukcije u usvajanju hrvatskoga jezika [Word order in early language acquisition: The syntax of two-word utterances]. *Rasprave: Časopis Instituta za Hrvatski Jezik i Jezikoslovlje, 39*(2), 433–456.

Hržica, G., & Roch, M. (2020). Lexical diversity in bilingual speakers of Croatian and Italian. In S. Armon-Lotem & K. Grohmann (Eds.), *LITMUS in action: Crosscomparison studies across Europe* (pp. 41–71). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.29.04hrz

Hughes, D., McGillivray, L., & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessment.* Thinking Publications.

Hunt, K. W. (1966). Recent measures in syntactic development. *Elementary English, 43*(7), 732–739. https://www.jstor.org/stable/41386067

Hunt, K. W. (1970). Syntactic maturity in school children and adults. *Monographs of the Society for Research in Child Development, 35*(1), 1–67. https://doi.org/10.2307/1165818

Jarvis, S. (2002). Short texts, best–fitting curves and new measures of lexical diversity. *Language Testing, 19*(1), 57–84. https://doi.

org/10.1191/0265532202lt220oa

Justice, L. M., Bowles, R. P., Kaderavek, J. N., Kaderavek, T. A., Eisenberg, S. L., & Gillam, R. B. (2006). The Index of Narrative Microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech-Language Pathology, 15*(2), 177–191. https://doi.org/10.1044/1058-0360(2006/017)

Karmiloff-Smith, K., & Karmiloff-Smith, A. (2002). *Pathways to language: From fetus to adolescent*. First Harvard University Press. https://doi.org/10.1017/S0305000903225630

Kuiken, F., Vedder, I., Housen, A., & De Clercq, B. (2019). Variation in syntactic complexity: Introduction. *International Journal of Applied Linguistics, 29*(2), 161–170. https://doi.org/10.1111/ijal.12255

Kupersmitt, J., & Berman, R. A. (2001). Linguistic features of Spanish–Hebrew children's narratives. In L. Verhoeven & S. Strömqvist (Eds.), *Narrative development in a multilingual context* (pp. 277–371). John Benjamins.

Kuvač Kraljević, J., Bošnjak Botica, T., Vujnović Malivuk, K., & Pinjušić, P. (2016). Linguistic complexity in late language development: Reformulation of relative clauses. *Suvremena Lingvistika, 42*(81), 27–42.

Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 3–38). University of Washington Press.

Leadholm, B., & Miller, J. F. (1992). *Language sample analysis: The Wisconsin guide*. Department of Public Instruction.

Liles, B. Z., Duffy, R. J., Merritt, D. D., & Purcell, S. L. (1995). Measurement of narrative discourse ability in children with language disorders. *Journal of Speech and Hearing Research, 38*(2), 415–425. https://doi.org/10.1044/jshr.3802.415

Lindgren, J. (2018). *Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish children aged 4–6* [Doctoral dissertation, Acta Universitatis Upsaliensis]. DiVA – Academic Archive Online. http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-347102

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

Loban, W. (1976). *Language development: Kindergarten through grade twelve (Research Report No. 18)*. National Council of Teachers of English. https://files.eric.ed.gov/fulltext/ED128818.pdf

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk: Transcription format and programs* (3rd ed.). Lawrence Erlbaum Associates Publishers.

Mäkinen, L., Loukusa, S., Nieminen, L., P., Leinonen, E., & Kunnari, S. (2014). The development of narrative productivity, syntactic complexity, referential cohesion and event content in four- to eight-year-old Finnish children. *First*

Language, 34(1), 24–42. https://doi.org/10.1177/0142723713511000

Malvern, D., & Richards, B. (1997). A new measure of lexical diversity. In A. Ryan & A. Wray (Eds.), *Evolving models of language* (pp. 58–71). Multilingual Matters.

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Palgrave MacMillan. https://doi.org/10.1057/9780230511804

Mamula, M., & Trtanj, I. (2018). Vrste rečenica u uzorcima govornoga jezika djece mlađe školske dobi [Types of sentences in spoken language samples of younger school-age children]. *Život i Škola: Časopis za Teoriju i Praksu Odgoja i Obrazovanja, 64*(1), 171–185. https://doi.org/10.32903/zs.64.1.13

Mayer, M. (1969). *Frog, where are you?* Dial Press.

McCabe, A., & Bliss, L. S. (2003). *Patterns of narrative discourse.* Allyn & Bacon.

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd–D, and HD–D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods, 42*(2), 381–392. https://doi.org/10.3758/BRM.42.2.381

McFadden, U. T., & Gillam, R. B. (1996). An examination of the quality of narratives produced by children with language disorders. *Language, Speech, and Hearing Services in Schools, 27*(1), 48–56. https://doi.org/10.1044/0161-1461.2701.48

Muñoz, M. L., Gillam, R. B, Peña, E. D., & Gully-Faehnle A. (2003). Measures of language development in fictional narratives of Latino children. *Language, Speech, and Hearing Services in Schools, 34*(4), 332–342. https://doi.org/10.1044/0161-1461(2003/027)

Nicolopoulou A., & Richner, E. S. (2007). From actors to agents to persons: the development of character representation in young children's narratives. *Child Development, 78*(2), 412–429. https://doi.org/10.1111/j.1467-8624.2007.01006.x

Nippold, M. A. (2016). *Later language development: School-age children, adolescents, and Young Adults*. PRO-ED, Inc.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555–578. https://doi.org/10.1093/applin/amp044

Ortega, L. (2003). Syntactic complexity measures and their relationship to l2 proficiency: A research synthesis of college-level l2 writing. *Applied Linguistics, 24*(4), 492–518. https://doi.org/10.1093/applin/24.4.492

Owens, R. E. (1999). *Language disorders: A functional approach to assessment and intervention*. Allyn & Bacon.

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

Peterson, C., & McCabe, A. (1983). High Point Analysis. In C. Peterson & A.

McCabe (Eds.), *Developmental Psycholinguistics: Three ways of looking at a child's narrative* (pp. 29–47). Springer. https://doi.org/10.1007/978-1-4757-0608-6

Peterson, C., & McCabe, A. (1991). Linking children's connective use and narrative microstructure. In C. Peterson & A. McCabe (Eds.), *Developing narrative structure* (pp. 29–54). Lawrence Erlbaum Associates, Inc. https://doi.org/10.1177/002383099103400305

Peterson, C., & McCabe, A. (2004). Echoing our parents: Parental influences on children's narration. In M. W. Pratt & B. H. Fiese (Eds.), *Family stories and the life course: Across time and generations* (pp. 27–54). Lawrence Erlbaum Associates Publishers. https://doi.org/10.4324/9781410610300

Scott, C. (1988). Spoken and written syntax. In M. Nippold (Ed.), *Later language development: Ages 9 through 19* (pp. 49–95). College-Hill.

Scott, C. M., & Stokes, S. L. (1995). Measures of syntax in school-age children and adolescents. *Language, Speech, and Hearing Services in Schools, 26*(4), 309–319. https://doi.org/10.1044/0161-1461.2604.309

Soodla, P., & Kikas, E. (2010). Macrostructure in the narratives of Estonian children with typical development and language impairment. *Journal of Speech, Language, and Hearing Research, 53*(5), 1321–1333. https://10.1044/1092-4388(2010/08-0113)

Stein, N. L., & Albro, E. R. (1997). Building complexity and coherence: Children's use of goal-structured knowledge in telling stories. In M. G. W. Bamberg (Ed.), *Narrative development: Six approaches* (pp. 5–44). Lawrence Erlbaum Associates Publishers.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing: Advances in discourse processes* (Vol. 2, pp. 53–120). Ablex.

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

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Allyn and Bacon.

Terry, N. P., Mills, M. T., Bingham, G. E., Mansour, S., & Marencin, N. (2013). Oral narrative performance of African American prekindergartners who speak nonmainstream American English. *Language, Speech, and Hearing Services in Schools, 44*(3), 291–305. https://doi.org/10.1044/0161-1461(2013/12-0037)

To, C. K., Stokes, S. F., Hin-Tat, C., & T'sou, B. (2010) Narrative assessment of Cantonese-speaking children. *Journal of Speech, Language, and Hearing Research, 53*(3), 648–669. https://doi.org/10.1044/1092-4388(2009/08-0039)

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 cross-linguistic developmental study* (pp. 85–106). Psychology Press.

Trtanj, I. (2015). *Jezično označavanje likova u dječjem pripovjednom diskursu* [Linguistic character marking in children's narrative discourse] [Doctoral dissertation]. Josip Juraj Strossmayer University of Osijek Faculty of Humanities and Social Sciences]. FFOS-repozitorij. https://urn.nsk.hr/urn:nbn:hr:142:554959

van den Broek, P. (1997). Discovering the cement of the universe: The development of event comprehension from childhood to adulthood. In P. W. van den Broek, P. J. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and representation: Bridging fictional and actual events* (pp. 321–342). Routledge Taylor & Francis Group. https://doi.org/10.4324/9781315044934

van der Lely, H. K. J. (1996). Grammatical specific language impaired children: Evidence for modularity. In C. Koster & F. Wijnen (Eds.), *The Groningen assembly on language acquisition* (pp. 283–292). University of Groningen Press.

Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech, Language, and Hearing Disorders, 38*(6), 1349–1355. https://doi.org/10.1044/jshr.3806.1349

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

Westby, C., Van Dongen, R., & Maggart, Z. (1989). Assessing narrative competence. *Seminars in Speech and Language, 10*(1), 63–75. https://doi.org/10.1055/s-0028-1082490

Westerveld, M. F., Gillon, G. T., & Miller, J. F. (2004). Spoken language samples of New Zealand children in conversation and narration. *Advances in Speech Language Pathology, 6*(4), 195–208. https://doi.org/10.1080/14417040400010140

Yu, G. (2009). Lexical diversity in writing and speaking task performances. *Applied Linguistics, 31*(2), 236–259. https://doi.org/10.1093/applin/amp024