The impact of language dominance on Russian-Hebrew bilingual children's narrative production

Microstructure, macrostructure, and Internal State Terms

Sveta Fichman[1,2] Joel Walters[1] Sharon Armon-Lotem[1] and Carmit Altman[1]
[1] Bar-Ilan University | [2] Talpiot College of Education

The study explores the effect of language dominance on microstructure, macrostructure, and Internal State Terms (ISTs) in narratives of Russian-Hebrew bilingual children and examines within-language and cross-language associations between narrative elements in two dominance groups. Narratives were collected from 38 Russian-Hebrew bilingual children aged 5;5–6;7 using the LITMUS-MAIN retelling procedures. The children were divided into L1-dominant ($N=19$) and L2-dominant ($N=19$) bilinguals based on performance on proficiency tests in L1/Russian and L2/Hebrew. The narratives were coded for microstructure measures: number of different words (NDW), total number of tokens (TNT), number of C-units (CUs), and Mean Length of C-unit (MLCU); and for macrostructure measures: Story Structure and Story Complexity. Ratios of IST tokens and types were calculated per C-unit. Children produced significantly higher NDW, TNT, and MLCU in L2/Hebrew than in L1/Russian. Scores on macrostructure measures and ratios of total ISTs were similar across languages. L1-dominant bilinguals demonstrated similarity between L1 and L2 for microstructure and IST types, whereas L2-dominant bilinguals produced more IST types in L2/Hebrew and had relatively few significant cross-language correlations. Findings for language dominance and cross-language differences are discussed for those narrative features which emerged as sensitive to these effects.

Keywords: narrative, bilingual children, dominance, macrostructure, microstructure, Internal State Terms

https://doi.org/10.1075/lab.20036.sve | Published online: 31 May 2022
Linguistic Approaches to Bilingualism 12:4 (2022), pp. 509–539. ISSN 1879-9264 | E-ISSN 1879-9272
© John Benjamins Publishing Company

1. Introduction

Language dominance is a unique construct in bilingualism which affects a range of domains, including narrative production. Narrative production is multidimensional, tapping into processes of lexical retrieval, utterance formulation, communication of ideas and use of Internal State Terms (ISTs). These aspects of narrative are traditionally divided into microstructure and macrostructure abilities. To date, few studies have examined the effect of language dominance on multiple narrative features in bilingual children. The main aim of the present study is to explore effects of language dominance and cross-language differences for three aspects of narrative production – microstructure, macrostructure, and ISTs. Microstructure reflects narrative productivity at the utterance level (e.g., mean length of utterance) and the word level (e.g., total number of words) (Gagarina et al., 2016; Uccelli & Páez, 2007). Macrostructure relates to the organization of events in a narrative using structural elements, such as goals and attempts (Heilmann et al., 2010). ISTs reflect lexical knowledge and thus may be associated with microstructure, but they also assist in conveying macrostructure knowledge (Burns et al., 2012; Symons et al., 2005). ISTs are primarily encoded in perceptual, motivational, metalinguistic, and metacognitive verbs, and physiological and emotional adjectives. A secondary aim of the present study is to better understand within- and cross-language relationships among microstructure, macrostructure, and ISTs in bilingual children's narrative production. To achieve these aims, narratives were elicited from 5–6-year-old Russian-Hebrew bilingual children. Since the majority of these children acquire their home language first, Russian will be referred to as 'L1', while the societal language Hebrew, acquired later in educational settings, will be referred to as 'L2'. The *Multilingual Assessment Instrument for Narratives* (*LITMUS-MAIN*, Gagarina et al., 2012), developed in the framework of COST Action IS0804 "Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment" (www.bi-sli.org), was chosen to address these questions, since it allows parallel assessment of both languages, targeting microstructure and macrostructure elements as well as ISTs. Narratives were collected using a retelling procedure with two stories in accordance with the MAIN procedures.

2. Background

2.1 Microstructure across languages in bilinguals' narratives

Narrative microstructure includes lexical, morphosyntactic and syntactic knowledge, often operationalized in terms of story length or narrative 'productivity' (e.g., total number of words and utterances), lexical diversity (e.g., word types), and mean length of utterance (e.g., Justice et al., 2010). Microstructure elements are by nature language specific, and a wide range of studies has shown differences in microstructure across bilingual children's two languages: Spanish-English (Lucero, 2015; Pearson, 2002; Silliman et al., 2002; Squires et al., 2014), English-Hebrew (Altman et al., 2016; Iluz-Cohen & Walters, 2012), Russian-Norwegian (Rodina, 2017), and Polish-English bilinguals (Otwinowska et al., 2018). Comparison of microstructure across languages provides insight into how bilingual children process narratives at the word and utterance levels involving expressive skills (Squires et al., 2014). The impact of language dominance on narrative microstructure is crucial for understanding narrative skills (Rodina, 2017). If language dominance directly impacts narrative microstructure, improving proficiency may enhance microstructure skills.

Research on cross-linguistic differences characterizing narrative microstructure is complicated by the use of different measures to assess microstructure and by the presence/absence of language proficiency scores as a factor explaining performance. Lexical measures include Number of Different Words (NDW); grammatical measures include Mean Length of C-unit (MLCU), and sometimes grammatical complexity (Kunnari et al., 2016; Lucero, 2015; Pearson, 2002; Rezzonico et al., 2016). Studies have also included mental and metalinguistic verbs, adverbs, and elaborated noun phrases (Squires et al., 2014). Some research has reported that only certain measures of microstructure differed across languages, such as lexical diversity (Simon-Cereijido & Gutiérrez-Clellen, 2009), NDW (Lucero, 2015), MLCU (Otwinowska et al., 2020), number of utterances (Otwinowska et al., 2018), while other measures remained stable across languages. Still others found that all measures of linguistic productivity were similar across the two languages (Fiestas & Peña, 2004; Rezzonico et al., 2016). Overall, studies agree that word-based, language-specific features of narrative microstructure potentially impact cross-linguistic comparisons but are needed, since they should shed light on bilingual children's patterns of language use across languages.

The impact of language proficiency on narrative microstructure in both languages of bilinguals has also been reported (Gutiérrez-Clellen, 2002; Kapalková et al., 2016; Montanari, 2004; Silliman et al., 2002). Montanari (2004) examined three 5-year-old Spanish-English bilinguals and classified their proficiency in each

language according to four levels: proficient, functional, intermediate, and limited. Via qualitative analyses, Montanari showed that in the dominant language (Spanish), children exploited a wider range of linguistic forms (in the verbal/aspectual system) and a greater variety of temporal and causal expressions. It was concluded that lack of proficiency might impede bilingual children's abilities to create a well-structured narrative in terms of microstructure. Bitetti et al. (2020) examined the effect of language dominance on narrative microstructure in both languages of 200 Spanish-English bilinguals by operationalizing dominance as relative language proficiency. Microstructure was assessed using measures such as NDW, MLCU, and a Subordination Index. L1-dominant children showed higher microstructure scores in L1/Spanish than their balanced bilingual peers, and balanced bilinguals had higher microstructure scores in L2/English than L1-dominant bilinguals. However, Gutiérrez-Clellen (2002) examined 196 Spanish-English bilingual children's narrative performance and showed that typically developing children who are fluent in both languages may still show different levels of narrative skills in L1 and L2.

The present study examines the effect of dominance and cross-linguistic differences for the following microstructure measures: NDW, TNT, number of CUs, and MLCU.

2.2 Macrostructure across languages in bilingual narratives

The use of macrostructure elements reflects an individual's ability to organize story information and order events (Stein & Glenn, 1979; Trabasso & Rodkin, 1994). Most models of macrostructure utilize Story Grammar (SG) categories to map events onto cognitive representations in order to organize information about characters and events in a hierarchical episodic structure (Westby, 2005). SG approaches operationalize macrostructure in different ways, sometimes measuring both SG components and episodic structure (e.g., ENNI; Schneider et al., 2006; NSS; Heilmann et al., 2010). In the MAIN story stimuli and scripts (Gagarina et al. 2012), each story begins with a Setting, followed by three episodes, each consisting of a Goal, an Attempt, and an Outcome. These elements are incorporated in the MAIN model as a Story Structure score, and its unique contribution involves characters' internal states as part of this structure: an IST as an Initiating Event and an IST as a Reaction. Identification of these internal states contributes to narrative coherence.

Research on children's narrative skills using a SG model have reported a developmental hierarchy in the use of macrostructure elements in storytelling (Stein & Glenn, 1979). Initiating Events, Attempts and Outcomes are produced earlier by children, while Goals are among the later developing SG elements

(Khan et al., 2016). Another aspect of the SG hierarchy which has received less attention is episode complexity, defined as the inclusion and combination of elements in a GAO (Goal-Attempt-Outcome) sequence. Story Complexity (SC) is assessed by scoring each episode for the sequence of SG components used by the child (Gagarina et al., 2012). SC is based on Westby's (2005) binary decision tree, where episodes are classified into categories of structural complexity: (i) sequences with no Goal statement (e.g., only Attempt and Outcome), (ii) incomplete episodes, which include a Goal but lack a complete Goal-Attempt-Outcome structure, and (iii) complete episodes, which include the three main SG components (Altman et al., 2016; Gagarina, 2016; Kunnari et al., 2016; Maviş et al., 2016; Roch et al., 2016). Research has reported developmental changes in SC (Gagarina, 2016; Kapalková et al., 2016), but no group differences for children with typical language development or with language impairment (Altman et al., 2016), for monolinguals and bilinguals (Kapalková et al., 2016), or for simultaneous and sequential bilinguals (Gagarina, 2016). In these studies, preschool children produced Goals, but the majority of their episodes were incomplete. Cross-linguistic similarity has been reported for SC (Bohnacker, 2016; Gagarina, 2016; Kapalková et al., 2016), but one study showed greater complexity in L1 than in L2 (Roch et al., 2016 for Italian-English bilinguals).

Overall, macrostructure abilities have been argued to be invariant across the two languages of bilingual children. Cross-linguistic similarity has been shown for Spanish-English bilingual children (Fiestas & Peña, 2004; Gutiérrez-Clellen, 2002; Montanari, 2004) using the wordless picture book *Frog, Where Are You?* (Mayer, 1969) and for English-Hebrew bilinguals using *Jungle Book* and *Goldilocks and the Three Bears* (Iluz-Cohen & Walters, 2012). Similarly, studies using MAIN data collection and coding procedures demonstrated invariance of macrostructure across languages for English-Hebrew (Altman et al. 2016), Russian-German (Gagarina et al., 2016), Polish-English (Otwinowska et al., 2018) and Finnish-Swedish (Kunnari et al., 2016) bilinguals. However, several studies have shown cross-linguistic differences in macrostructure. One study used *Frog, Where Are You?* with Cantonese-English bilinguals (Rezzonico et al., 2016); another used a picture book with Spanish-English bilinguals (Uccelli & Páez, 2007), and two studies used the MAIN procedures, one with Slovak-English bilingual children (Kalpaková et al., 2016) and one with German-Swedish bilinguals (Lindgren & Bohnacker, 2020). Rezzonico et al. (2016) interpreted their results as cultural differences influencing the way stories are told. Kapalková et al. (2016) attributed cross-language differences to the impact of language proficiency on macrostructure. Lindgren and Bohnacker (2020) found macrostructure skills to be more developed in the majority language (Swedish).

The present study addresses whether macrostructure measured by Story Structure and Story Complexity is similar or different across bilingual children's languages and whether the potential difference is affected by language dominance.

2.3 Internal State Terms across languages in bilingual narratives

Use of ISTs involves the interpretation of the sequence of events and of the characters' intentions and feelings (Burns et al., 2012; Ukrainetz et al., 2005). Well-organized stories with a clear beginning, middle, and end may still lack coherence if they do not include evaluative devices such as ISTs (Berman & Katzenberger, 2004). Of special interest are perceptual terms which are crucial for conveying causal relations between events (Fichman et al., 2021) and motivational terms which reflect characters' goals (Fusté-Herrmann et al., 2006). Furthermore, children's first attempts at evaluative devices involve elaborating on what characters feel (Nicolopoulou et al., 2021). In support of the claim that ISTs assist in conveying crucial macrostructure elements, Berman & Slobin (1994) have shown that the difference between a temporally ordered list of actions (which characterizes narratives of very young children) and a well-structured narrative is the use of different types of ISTs, such as perceptual ('see') and emotional terms ('sad', 'happy'). ISTs are also grounded in lexical and semantic knowledge (Florit et al., 2011) and as such reflect microstructure knowledge.

There is no consensus regarding the cross-linguistic use of ISTs in bilingual children's narratives. Some studies have demonstrated higher frequencies of ISTs for L2 than for L1 (Altman et al., 2016 for English-Hebrew and Gagarina, 2016, for Russian-German bilinguals), others show higher frequencies for L1 (Silliman et al., 2002, for Spanish-English bilinguals), and still others found similar frequencies of ISTs for both languages (Kunnari et al., 2016, for Finnish-Swedish; Otwinowska et al., 2018, for Polish-English; and Squires et al., 2014, for Spanish-English bilinguals). Several of these studies used the MAIN coding procedure where ISTs are incorporated into macrostructure coding. In the current study, ISTs are coded as part of macrostructure, but are also classified into different types following Fusté-Herrmann et al. (2006) in order to explore cross-linguistic differences and similarities.

2.4 Overview of Russian and Hebrew structures

Russian, a Slavic language from the Indo-European family, and Hebrew, a Semitic language, show both similarities and differences in morpho-syntactic features which are critical when comparing microstructure and macrostructure across languages. Both languages have relatively free word order (Doron, 2000; Timberlake,

2004), and rich morphology. However, differences in morphological encoding may influence narrative microstructure. Hebrew overtly marks definiteness by a bound prefix *ha-*, while Russian has neither definite nor indefinite articles. Russian has a rich case system marked by suffixes on nouns (Gagarina, 2011). In Hebrew, only accusative case is marked overtly, by the free morpheme *et* preceding definite nouns in direct object position (Berman, 1978). Possession is conveyed by case inflection in Russian, but in Hebrew the free morpheme *shel* precedes the possessed noun. Likewise, some arguments expressed by cases in Russian (e.g., ptiča dala edu ptenčik**am** 'the bird gave food to baby birds') are assigned by prepositions in Hebrew (e.g., ha- cipor natna oxel **la-** gozalim 'the bird gave food to the baby birds'). These language-specific features are essential for examining *cross-language* relationships in microstructure, macrostructure and ISTs.

2.5 The present study

The current study examines the impact of language dominance on both *within-language* and *cross-language* relationships for microstructure, macrostructure and IST measures in bilingual children from two language dominance groups, L1-dominant (L1D) and L2-dominant (L2D). We investigated the combined effects of dominance and cross-linguistic differences on different aspects of narrative production. Three research questions are addressed:

1. To what extent does language dominance affect microstructure, macrostructure and the use of ISTs in the Russian and Hebrew narratives of Russian-Hebrew bilingual children?
2. To what extent do L1D and L2D bilingual children show cross-linguistic differences in microstructure, macrostructure and the use of ISTs?
3. To what extent are microstructure, macrostructure, and ISTs related *within* and *across* languages and to what extent do these relationships differ for L1D and L2D bilinguals?

It is predicted that bilinguals dominant in L1 (L1D) will perform better in L1 macrostructure and microstructure, and bilinguals dominant in L2 (L2D) will perform better in L2. Microstructure and ISTs are expected to show cross-linguistic differences (Iluz-Cohen & Walters, 2012), while similarity is expected for macrostructure across the two languages (Pearson, 2002). The SC measure is expected to yield differences across episodes, since children do not always produce similar categories of episodic complexity across the episodes of a narrative (Fichman et al., 2017). Strong within-language relationships and moderate cross-language relationships are expected for microstructure, macrostructure, and ISTs

(Squires et al., 2014). Dominance is expected to influence both within- and cross-language associations (Montanari, 2004).

3. Method

3.1 Participants

Data from 38 sequential Russian-Hebrew bilingual children aged 5;5–6;7 were collected. Children were considered Russian-Hebrew bilinguals if they communicated in both languages and were exposed to L1/Russian from birth, with Russian spoken at home as a mother tongue by at least one parent. All children were exposed to L2/Hebrew primarily during preschool. Age of onset of bilingualism (AoB) was determined in months based on parental reports of the age of exposure to L2/Hebrew. Only participants exposed to Hebrew for more than 12 months were included. All parents reported that only Russian or both Russian and Hebrew were currently spoken at home. Initially, data from 51 participants were collected. Based on parental reports, children who had hearing problems or whose parents expressed concern regarding their children's language development were excluded, which produced a final set of 38 participants. The study received ethics approval from the university Institutional Review Board and was approved by the Israeli Ministry of Education. All parents gave their written consent and all children gave their oral assent.

Language proficiency tests were administered in L1 and L2 prior to the narrative data collection. To assess proficiency in Russian, the *Russian Language Proficiency Test for Multilingual Children* (Gagarina et al., 2010) was administered. Hebrew proficiency was evaluated using the *Goralnik Screening Test for Hebrew* (Goralnik, 1995). The two proficiency tests included sub-tests targeting vocabulary, grammar, morpho-syntax, and discourse. Raw scores for both proficiency tests were calculated as z-scores based on local bilingual standards. Children whose performance on proficiency tests was below the cut-off point of –1.25SD in both languages were at risk for Developmental Language Disorder and were excluded from the study (Armon-Lotem & Meir, 2016 for Russian; Altman et al., 2021 for Hebrew).

Language dominance was determined as relative proficiency between the two languages (Paradis et al., 2003). Total proficiency test scores were first transformed to z-scores, and dominance was operationalized as an index of relative proficiency based on the difference between the two standardized proficiency scores. A child was considered dominant in a language, if he or she received a higher score in that language on the proficiency tests. Differences between the z-

scores in L1 and L2 ranged from 0.43 to 5.98. In addition to the difference of 0.43, three children had z-scores less than 1, with difference scores of 0.82, 0.86, and 0.97. Table 1 presents background information for all participants and the results of t-tests comparing the two groups for all measures.

Table 1. Background information

	L1D *N* = 19	**L2D** *N* = 19	*t*(36)
Age (months)	71.47 (68–79)	69.74 (65–76)	1.61, $p = .12$
AoB (months)	32.79 (0–54)	18.84 (0–48)	**2.52, $p = .02$**
LoE (months)	39.50 (12–69)	48.58 (15–73)	1.51, $p = .14$
Russian proficiency z-score	0.33 (−1.14/1.83)	−2.25 (−4.56/−0.22)	**7.06, $p < .001$**
Hebrew proficiency z-score	−1.74 (−3.67/1.31)	0.19 (−1.19/1.65)	**5.65, $p < .001$**

Note. AoB=Age of onset of Bilingualism, LoE=Length of Exposure

The analyses revealed significant differences between L1D and L2D bilinguals for AoB and for the two proficiency tests. The difference for AoB was expected, since it is known to affect dominance (Altman et al., 2018).

3.2 Materials

Narrative assessment

Narratives were elicited using the *Multilingual Assessment Instrument for Narratives* (Gagarina et al., 2012). The instrument is comprised of four parallel stories, each with a carefully designed six-picture sequence controlled for cognitive and linguistic complexity and parallel macro- and microstructure. The present study used a retelling procedure with two 6-picture sequences and scripts: 'Baby Birds' and 'Baby Goats' adapted to Russian and Hebrew. The Russian script includes 124 and 122 words for 'Baby Birds' and 'Baby Goats' respectively. The Hebrew script includes 126 and 131 words for 'Baby Birds' and 'Baby Goats' respectively.

Parental questionnaire

A parent questionnaire, adapted from the Bilingual Parental Questionnaire (BIPAQ) (Abutbul-Oz et al., 2012), elicited information on children's age, language exposure, language use, parents' concern regarding their children's language development.

3.3 Procedure

Narratives were elicited in a designated area in the children's Hebrew-speaking preschools located in the central region of Israel. All narratives were collected using a story retell procedure; the data collectors were native speakers of Russian and Hebrew. Story content (Baby Birds/Baby Goats) and order of presentation were counterbalanced across languages (L1/L2). The story script was first read as the child followed along looking at the pictures. The child was then asked to retell the story to a puppet. The puppet was not present when the experimenter read the story, but appeared when the child was retelling her story. During retelling, pictures were perceptually available to the child and the puppet. All narratives were audio recorded.

3.4 Transcription and coding

All narratives were transcribed by trained research assistants, who were native speakers of Russian and Hebrew using CHAT conventions (MacWhinney, 2000). The transcribers, all graduate students in Linguistics or Education, were instructed and supervised by the first author, a Russian-Hebrew bilingual. Segmentation of audio data into utterances was based on communication units (CU) (Loban, 1976). A CU consists of one main clause along with its dependent clause(s). Coordinated clauses are treated as separate CUs except in cases where the co-referential subject of the second clause is omitted (e.g., 'the cat came and saw one baby bird').

The coding for *microstructure* included TNT, Number of Different Words (NDW), number of C-units (CU), and Mean Length of Utterance (MLCU). The NDW measure involved counting different words, defined in MAIN as equivalent to lemmas or root forms (Gagarina et al., 2015). Morphologically related words (*tofes* 'catch', *tafas* 'caught', and *etfos* 'will catch') were scored as three tokens and one type. Since the division into utterances was based on the CU approach, mean length of a CU was calculated (MLCU) rather than MLU. Bound morphemes (e.g., Hebrew *ba-bait* 'in the house') were not counted as separate words.

The coding for *macrostructure* included Story Structure (SS) (a quantitative measure) and Story Complexity (SC) (a qualitative measure, reflecting the well-formedness of each episode). SS was calculated as the sum of SG categories including Setting, IST as an Initiating Event, Goal, Attempt, Outcome, and IST as a reaction. Based on the MAIN coding system, each of the three episodes consists of five elements. The Initiating Event is expected to be encoded with an IST ('the mother bird saw that baby birds were hungry'), followed by a Goal, Attempt, and Outcome. Finally, the character's Internal Response to the Outcome is expressed with an IST as well. Each element assigned one point to the overall result of SS. The additional two points were assigned to Setting (1 point for time and 1 point for place). The maximum SS score was 17.

The coding of the SC measure assigned each episode to one of five categories, all tapping into episodic structure. The five categories reflected the complexity of the episode as determined by inclusion of SG elements in that episode. If no SG elements were produced, or only an Attempt or an Outcome, the label 'none' was assigned. When a child produced an Attempt and an Outcome (AO) but no Goal for an episode, the 'AO' label was assigned to that episode. When a Goal (but not an Attempt or Outcome) was produced, the episode was assigned to the 'G' category. When a child produced a Goal and an Attempt (GA) or a Goal and an Outcome (GO) for an episode, the episode was assigned the category 'GA/GO'. When the sequence of all three main SG elements, i.e. Goal-Attempt-Outcome (GAO) was produced for an episode, it was assigned the category 'GAO'.

In addition to the MAIN coding procedure for ISTs, ISTs were classified into seven categories. Six categories were adopted from Fusté-Herrmann et al. (2006): perceptual (see, hear), motivational (want, decide), physiological (thirsty, hungry), linguistic (say, shout), emotional (happy, sad), and mental (think, know) (Altman et al., 2016; Fusté-Herrmann et al., 2006; Gagarina et al., 2012). An additional category was trait ('brave'). In previous studies, the calculation of ISTs was conducted as ratios per utterance (Fusté-Herrmann et al., 2006) or per total number of tokens (Gagarina, 2016). While both techniques enable normalizing the frequency, we adopted the approach in Fusté-Herrmann et al. (2006) which calculates ratios by dividing the total number of ISTs and the number of IST types by the number of CUs.

Reliability

After the transcriptions were checked and coded by a bilingual speaker (first author) according to the protocols of MAIN, twenty percent of transcripts were randomly chosen in each language and coded for inter-rater reliability of macrostructure coding. The reliability coders were a native speaker of Russian and a native speaker of Hebrew, both with Ph.D.s in Linguistics. Reliability was

assessed using intra-class correlation coefficients (ICC). ICC scores between 0.75 and 0.90 are indicative of good reliability, and ICC scores greater than 0.90 are indicative of excellent reliability (Koo & Li, 2016). In Russian, the ICC score was 0.90 and in Hebrew the ICC score was 0.93.

Reliability for coding of ISTs was also performed on 20% of the data in each language. In L1/Russian for total ISTs, the ICC score was 0.98, for perceptual ISTs 0.96, for motivational ISTs 0.98, for physiological and linguistic ISTs 1.00, and emotional, mental, and trait were not found in the sample randomly chosen for reliability. In L2/Hebrew for total ISTs, the ICC score was 0.96, for perceptual ISTs 0.99, for motivational ISTs 0.95, for emotional, linguistic, trait and physiological ISTs 1.00, and for mental, none were found in the sample randomly chosen for reliability.

3.5 Data analysis

Narrative microstructure measures, Story Structure, and ISTs were examined for the effects of Dominance group (L1D/L2D) as a between-subject variable and Language (L1/L2) as a within-subject variable using mixed-design Analyses of Variance (ANOVAs). Chi-square tests were used to analyze the relationship between language dominance and the percentage of children using categories in the Story Complexity measure (none/A/O, AO, G, GA/GO, GAO). Pearson correlation analyses were performed to explore relationships among microstructure, macrostructure, and ISTs, within and across languages.

4. Results

The results section first presents the analysis of microstructure for the effects of Dominance group (L1D/L2D) and Language (L1/L2), followed by the analyses of the two macrostructure measures, Story Structure (SS) and Story Complexity (SC). Next, percentage of ISTs and IST types are presented. Finally, relationships among macrostructure and microstructure performance and IST use are examined for L1 and L2.

Prior to data analysis and in order to examine the potential difference between the two stimulus narratives, t-tests were performed comparing the two stories (Baby Birds and Baby Goats) for all dependent measures: microstructure (NDW, TNT, number of CUs, and MLCU), macrostructure (SS, SC), and ISTs (percentage and types). The analyses showed a significant difference for MLCU in L1, $t(36) = 2.17$, $p = .04$, where children produced narratives with higher MLCU for the Baby Goats narrative (see section below) than for the Baby Birds narrative.

For this reason, MLCU was analyzed via a three-way ANOVA with Story as an independent variable. No other measures differed across the two stimulus narratives.

4.1 Effects of Dominance and Language on microstructure

In order to provide insight into how bilingual children express microstructure in their two languages and how language dominance impacts microstructure the following measures were analyzed: NDW, TNT, number of CUs, and MLCU. Table 2 shows the mean scores (SDs) for the microstructure features in the two dominance groups (L1D/L2D) in both languages.

2×2 ANOVAs performed on NDW, TNT, and number of CUs did not yield significant effects for Dominance. An interaction between Dominance and Language was significant only for NDW. Post-hoc analysis with Bonferroni corrections revealed that in the L2D group children used more different words in L2/Hebrew than in L1/Russian ($p = .01$), while in the L1D group the difference between languages was not significant ($p = .71$). For MLCU, where a significant difference resulted between the Baby Birds and Baby Goats stories in L1/Russian (higher MLCU for Baby Goats), a Story content variable was added. For MLCU, the analysis did not produce a significant effect for Dominance nor for the interaction between Dominance and Language. The effect of Language on the microstructure measures was not significant. The three-way interaction Dominance x Language x Story content was not significant.

In sum, L2D children produced higher scores for NDW in Hebrew, while no language differences were observed for TNT, CU, and MLCU.

4.2 Effects of Dominance and language on macrostructure

Results for macrostructure are presented first for Story Structure and then for Story Complexity.

Story structure

Table 3 presents the means (SDs) for the total SS score for the two groups. Children from both groups produced on average 7/17 SS categories.

A 2×2 ANOVA revealed that children in both groups included a similar number of SG elements and performed similarly in both languages. The interaction between Dominance group and Language was not significant.

Table 2. Mean scores (and SDs) for microstructure features and ANOVA results

	L1D		L2D							
	L1/Russian	L2/Hebrew	L1/Russian	L2/Hebrew	Dominance	Language	Story	Language x Story	Dominance x Language	Dominance x Language x Story
NDW	36.16 (11.41)	35.26 (10.56)	33.42 (9.16)	39.63 (9.63)	$F(1,36)=0.05$, $p=.82$	$F(1,36)=2.56$, $p=.12$	NA	NA	**$F(1,36)=4.58$ $p=.04$, $\eta^2=.11$**	NA
TNT	55.16 (21.56)	48.74 (19.43)	53.32 (12.78)	54.21 (16.32)	$F(1,36)=0.15$, $p=.71$	$F(1,36)=0.71$, $p=.41$	NA	NA	$F(1,36)=1.25$, $p=.27$	NA
CUs	10.42 (3.96)	10.37 (3.00)	10.84 (3.44)	10.53 (3.20)	$F(1,36)=0.10$, $p=.76$	$F(1,36)=0.10$, $p=.75$	NA	NA	$F(1,36)=0.05$, $p=.82$	NA
MLCU	5.64 (1.47)	4.80 (1.86)	5.08 (1.25)	5.23 (0.98)	$F(1,34)=0.08$, $p=.78$	$F(1,34)=1.60$, $p=.21$	$F(1,34)=0.20$, $p=.66$	$F(1,34)=3.38$, $p=.08$	$F(1,34)=0.26$, $p=.61$	$F(1,34)=0.18$, $p=.67$

Notes. NDW=Number of Different Words; TNT=Total Number of Tokens; CUs=number of C-units; MLCU=Mean Length of C-unit

Table 3. Means (SDs) for Story Structure score and results of ANOVA

L1D		L2D				
L1/ Russian	L2/ Hebrew	L1/ Russian	L2/ Hebrew	Dominance $F(1,36)$	Language $F(1,36)$	Dominance x Language $F(1,36)$
7.21 (1.84)	6.89 (2.13)	7.16 (2.17)	7.16 (2.14)	0.05, $p=.83$	0.12, $p=.73$	0.12, $p=.73$

Story complexity

SC, which reflects episodic complexity, was coded using five categories: no elements (or single Attempt or single Outcome); AO (combination of an Attempt and Outcome); G (only Goal without an Attempt/Outcome); GA/GO (Goal and Attempt or Goal and Outcome); and GAO (full GAO sequence). The percentage of children using each category of SC is presented in Figure 1 for each of the three episodes in each language by Dominance group.

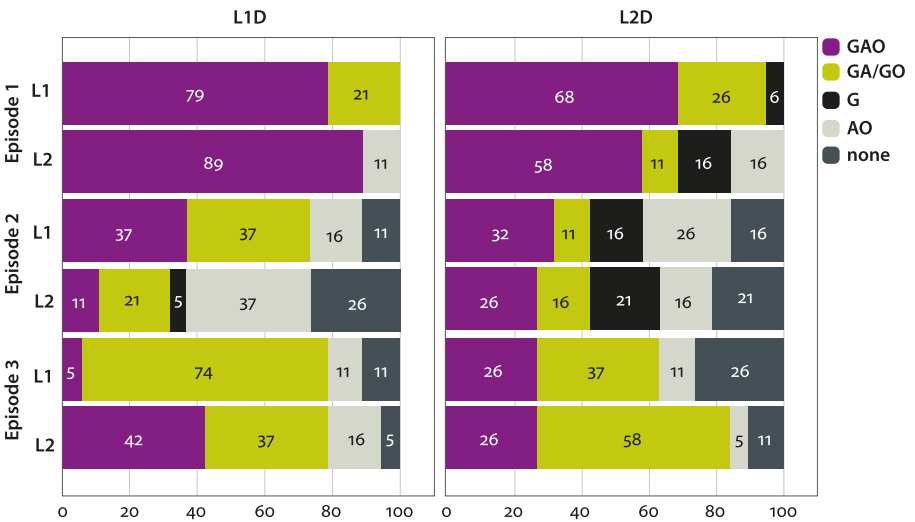

Figure 1. Percentage of children producing categories of SC in each episode in L1 and L2 (Total number of episodes = 228)

Notes: L1=Russian; L2=Hebrew; AO= Attempt and Outcome; G= Goal (without Attempt/Outcome); GA/GO= Goal and Attempt or Goal and Outcome; GAO=production of Goal-Attempt-Outcome

Figure 1 presents the SC measures for the L1D and L2D groups in each of the three episodes. Among the three episodes, episode 1 had the highest percentage of children who produced no elements or a single Attempt/Outcome. In episode

2, children who produced no elements or single Attempts/Outcomes were 13 in L1 and seven in L2. Finally, in episode 3, six children produced no elements or a single Attempt/Outcome in L1. A contrasting pattern emerged for the production of the full GAO sequence. No child produced the full GAO in episode 1, while in episode 2, five children used the complete GAO sequence in L1. In the third episode, 7 children produced the full GAO sequence in L1 and three children did so in L2. Overall, in L1 seven L1D and eight L2D children produced the full GAO. In L2, six L1D and six L2D children produced the full GAO.

Chi-square tests for the relationship between Dominance and production of SC categories showed no significant results for any of the three episodes in either L1 or L2.

Descriptively, approximately a third of all children produced Goals in their narratives producing categories 'GA/GO' or 'GAO'. The use of a Goal (without an Attempt or Outcome) was very infrequent and almost non-existent in the L1D group.

4.3 Effects of Dominance and language on Internal State Terms

The ratios of total ISTs (IST tokens) and IST types per CU are presented in Table 4.

Table 4. Mean ratios of ISTs (SDs) and IST types (SDs)

	L1D		L2D				
	L1/ Russian	L2/ Hebrew	L1/ Russian	L2/ Hebrew	Dominance $F(1,36)$	Language $F(1,36)$	Dominance x Language $F(1,36)$
IST tokens	0.46 (0.16)	0.48 (0.19)	0.49 (0.24)	0.48 (0.22)	0.10, $p=.76$	0.12, $p=.73$	0.13 $p=.72$
IST types	0.35 (0.16)	0.30 (0.10)	0.30 (0.15)	0.34 (0.15)	0.02, $p=.90$	0.01, $p=.92$	1.71, $p=.20$

Note. Ratios are calculated per CU

A 2×2 ANOVA for Dominance and Language performed on the ratio of IST tokens and types showed neither significant main effects nor significant interactions.

In order to further explore which IST types were most frequent, Figure 2 presents the distribution of IST categories, for L1 (left panel) and for L2 (right panel). In both languages perceptual ISTs (Russian: *uvidela*, Hebrew: *raata* 'saw') were the most frequent category, followed by motivational ISTs (Russian: *xotela*, Hebrew: *racta* 'want'), linguistic ISTs (Russian: *skazala*, Hebrew: *amra* 'said'),

and emotional ISTs (Russian: *ispugalsja*, Hebrew: *paxad*, 'was scared'). These categories were mainly used to convey macrostructure categories. All perceptual ISTs were signaling Initiating Events, motivational ISTs were signaling Goals, emotional ISTs were mainly used to convey Internal Responses. Linguistic ISTs were sometimes used to express Initiating Events.

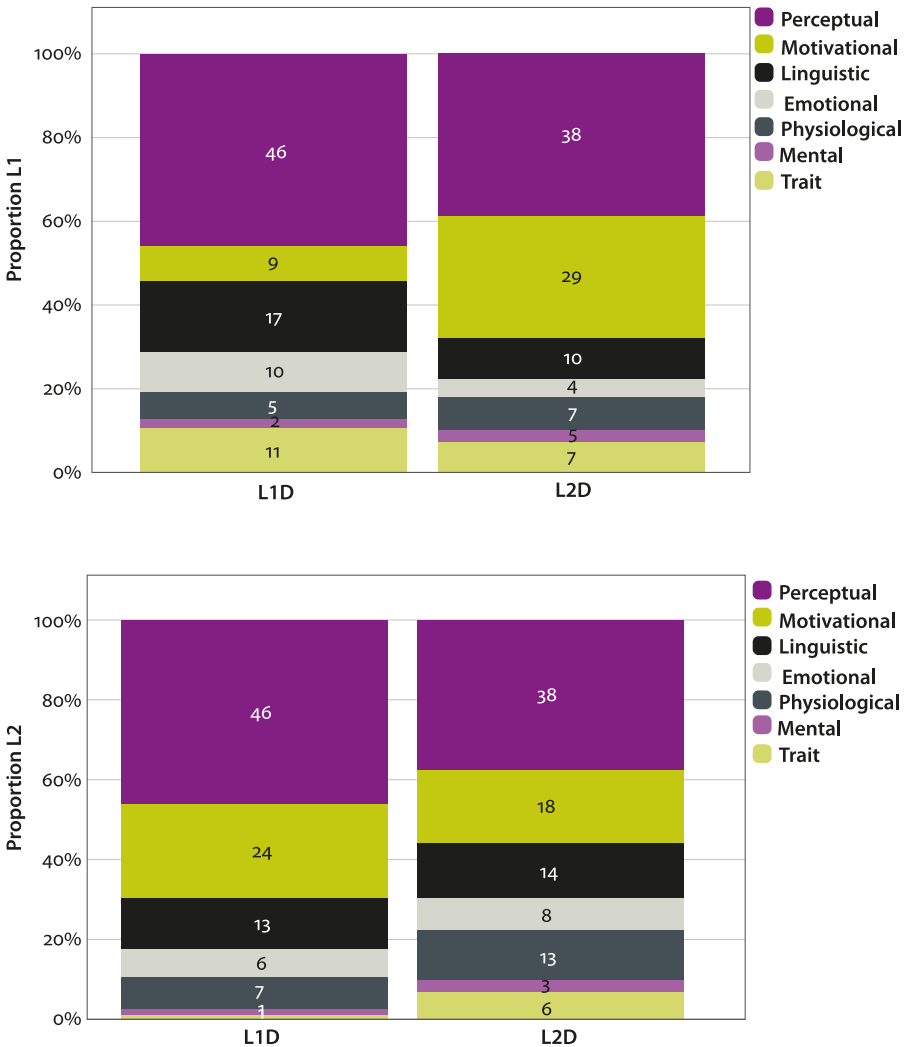

Figure 2. Percentages of IST types for L1D and L2D children in L1/Russian (left) and L2/Hebrew (right)

Mann-Whitney non-parametric tests examining the difference between L1D and L2D bilinguals for each IST category yielded no significant results in either language.

In sum, IST categories (primarily perceptual, motivational, and linguistic verbs) which assist in conveying macrostructure (Initiating Events and Goals) were found to be the most frequent for both dominance groups.

4.4 Within-language and cross-language relations of microstructure, macrostructure, and ISTs

The final set of analyses explored within-language and cross-language Pearson correlations among measures of microstructure (TNT, NDW, number of CUs, and MLCU), macrostructure (SS), and ISTs (frequency and type) for L1 and L2 separately for the L1D group (Table 5a) and L2D group (Table 5b) separately. Significant correlations are bolded.

Different patterns emerged for the L1D and L2D groups. For the L1D group, 14 significant cross-language correlations and 16 significant within-language correlations emerged. For the L2D group there were only four significant cross-language correlations and 20 within-language correlations. Both groups had significant *within-language* correlations, where microstructure measures correlated among themselves, the SS score correlated with several microstructure measures (NDW, TNT, and number of CU), and IST types correlated with SS and microstructure measures. However, significant *cross-language* correlations were found mainly in the L1D group, with significant correlations among microstructure, SS and IST types (upper right quadrant in Table 5a).

In sum, L1D children demonstrate significant associations among narrative measures both within and across languages, whereas the L2D children show mainly within-language associations.

5. Discussion

The goals of the present research were to examine the effects of language dominance on microstructure, macrostructure, and ISTs and to explore cross-linguistic differences in the narratives of L1 and L2 dominant Russian-Hebrew bilingual children aged 5–6. For microstructure, NDW, TNT, number of CUs, and MLCU were similar for the two language dominance groups, with higher scores in L2/Hebrew than in L1/Russian for NDW in the narratives of L2D bilinguals. For macrostructure, contrary to the predictions that dominance would affect macrostructure measures, the SS score yielded similar performance for the two

Table 5a. Within and cross-language Pearson correlations for microstructure measures, SS scores, and ISTs for *L1D* bilinguals

		L1/Russian							L2/Hebrew						
		SS	NDW	TNT	CU	MLCU	IST%	IST type	SS	NDW	TNT	CU	MLCU	IST%	IST type
L1/Russian	**SS**		**.51***	**.51***	**.62****	−.40	**.47***	.28	.42	**.61****	.43	**.52***	.10	−.01	−.00
	NDW			**.94****	**.87****	.17	.03	−.18	**.48***	**.75****	**.60****	**.48***	.29	.01	.12
	TNT				**.94****	.12	.04	−.20	**.53***	**.70****	**.60****	**.52***	.26	−.07	−.02
	CU					−.22	.09	−.11	.44	**.60****	**.48***	**.50***	.15	−.20	−.14
	MLCU						−.28	−.31	.15	.22	.27	−.03	.29	**.46***	.43
	IST%							**.84****	.14	.26	.23	.11	.04	−.03	−.22
	IST type								−.21	.00	.05	.24	.10	−.22	−.29
L2/Hebrew	SS									**.76****	**.64****	**.67****	.20	.04	−.09
	NDW										**.87****	**.74****	.38	.09	.19
	TNT											**.57****	**.69****	.09	.07
	CU												−.20	−.16	−.07
	MLCU													.19	.10
	IST%														**−.72****
	IST type														

Table 5b. Within and cross-language Pearson correlations for microstructure measures, SS scores, and ISTs for *L2D* bilinguals (L1/Russian on the left and L2/Hebrew on the right)

		L1/Russian							L2/Hebrew						
		SS	NDW	TNT	CU	MLCU	IST%	IST type	SS	NDW	TNT	CU	MLCU	IST%	IST type
L1/Russian	SS		**.50***	**.72****	.12	**.59****	**.69****	**.65****	−.21	.43	−.42	−.45	−.17	.01	.06
	NDW			**.75****	**.65****	.10	.26	.19	.40	.16	.04	.22	−.08	−.19	−.12
	TNT				**.61****	.42	.27	.33	.07	−.07	−.11	−.08	.09	−.12	−.1
	CU					−.43	−.19	−.22	.28	.26	.18	.44	−.29	−.39	−.30
	MLCU						**.50***	**.62****	−.29	−.36	−.38	**−.59****	.39	.27	.18
	IST%							**.87****	−.17	−.43	−.45	−.44	.15	.11	−.10
	IST type								−.11	**.50***	**−.51***	**−.51***	.09	.10	.00
L2/Hebrew	SS									**.59****	**.51***	**.51***	.05	.23	.23
	NDW										**.95****	**.82****	.25	.19	.04
	TNT											**.80****	.32	.17	.01
	CU												−.29	−.28	−.40
	MLCU													**.73****	**.62****
	IST%														**.86****
	IST type														

* $p < .05$ ** $p < .01$
***$p < .001$
Note:
NDW = Number of Different Words; TNT=Total Number of Tokens

dominance groups in both languages. The analyses of Story Complexity at the episodic level revealed that children in both dominance groups omitted SG elements or included only Attempt or Outcome for the first episode and produced narratives with higher complexity (inclusion of a Goal and one or more other elements) in the second and the third episodes in both languages. For ISTs, dominance effects were not significant. Finally, cross-linguistic associations among microstructure measures, SS score, ratio of total ISTs, and ratio of IST types showed different manifestations for the two dominance groups. We discuss the effects of dominance and cross-linguistic differences first for microstructure, then for macrostructure, next for ISTs and finally for the interrelationship of different narrative features.

5.1 Microstructure abilities across languages and dominance groups

Microstructure narrative abilities are widely reported to differ across the languages of bilingual children (Pearson, 2002; Rodina, 2017; Squires et al., 2014). But the claim that microstructure abilities differ across languages is not without controversy (Cereijido & Gutiérrez-Clellen, 2009; Rezzonico et al., 2016). The impact of language dominance on narrative microstructure is understudied. The results that exist suggest that bilinguals' language proficiency is not a unitary construct, and it may not map directly onto narrative performance (Bitetti et al., 2020; Cereijido & Gutiérrez-Clellen 2009; Kapalková et al., 2016; Montanari, 2004). Similarity across languages for TNT, number of CUs, and MLCU found in the present research corroborates results reported by Rodina (2017) for Russian-Norwegian bilinguals whose CU scores were invariant across languages.

Microstructure is not always operationalized in the same way. The definition of a word can differ across languages which directly affects measures of NDW, TNT, and MLCU. These measures tap into production at the sentence level, and thus a cross-linguistic typological difference, mainly related to differences in the use of function words, cannot be discounted. Russian is more inflected than Hebrew; thus fewer function words might have been used to retell narratives in Russian than in Hebrew. These typological differences influence cross-linguistic comparisons of microstructure and teasing apart the influence of typological features and the effect of dominance is challenging. The current study adopted a language specific approach to counting words based on orthography. Thus, Hebrew differed from Russian, resulting in comparable performance on most microstructure measures (TNT, CU, and MLCU). Effects of morpho-syntactic features on sentence-level microstructure measures have been addressed as a methodological challenge in the literature on bilingual production (e.g., Bernardini & Schlyter, 2004; Otwinowska et al., 2020; Yip & Matthews, 2006).

Three generalizations regarding microstructure emerge from the present results. First, cross-linguistic differences were found only for NDW, the only measure affected by language dominance. Second, the significant differences in the scores on standardized proficiency tests between L1D and L2D children did not lead to significant differences in narrative microstructure scores for most measures. The lack of a dominance effect at the sentence level may indicate that narrative microstructure performance is tapping into different skills than those targeted in proficiency assessment. Finally, cross-linguistic analyses of microstructure should take into account language-specific typological features.

5.2 Macrostructure across languages and dominance groups

For macrostructure, children produced on average fewer than half of the target story structure elements, with a striking similarity across the two dominance groups and across languages. This aligns with previous studies of bilingual children using MAIN procedures (Bohnacker, 2016; Rodina, 2017) where Attempts and Outcomes were the most frequent elements produced by Russian-Hebrew bilingual children.

Examination of Story Complexity at the episodic level (GAO) also revealed similarity across groups and languages. Attempts and/or Outcomes were the most frequent elements across all three episodes, and Goals (with or without other elements) were the least frequent, emerging mainly in episodes 2 and 3. Attempts which are triggered by Goals are expected to be mentioned by a child telling a story (Trabasso et al., 1989), but Goals were rarely produced by the children, as found previously with Russian-Hebrew bilinguals (Fichman et al., 2017). One difficulty in assessing narrative macrostructure is to identify cues signalling that the speaker has realized the character's goal. Overt expression of goal information usually comes in the form of motivational verbs, such as 'want,' 'decide'. Goals can also be expressed via prepositional and case constructions or converbs. Goals may also be covertly expressed using other ISTs. For example, the use of an IST as an Initiating Event (i.e., 'the dog saw that the Baby Birds were in danger') can be used instead of an explicit Goal statement.

Story Complexity reflects internal episodic structure and may be argued to be less dependent on linguistic knowledge. Differences in the use of full GAO sequences were observed across the stories' three episodes. Children in both dominance groups used more complete GAO structures in Episode 2. Goals, a category infrequently used by preschool children, did appear in the central episode of the narrative. These findings confirm that children are sensitive to internal episodic structure and that SC undergoes developmental changes; however, the impact of dominance was not found in children of the same age.

The current results support claims for invariant macrostructure abilities. Even though the L1D group showed significantly better performance in L1 on proficiency tests, they showed nearly balanced performance for narrative macrostructure. This suggests that children may rely on cognitive and language-universal resources, rather than language-specific skills when relating events by making use of SG elements. One reason for well-developed macrostructure skills could be exposure to storytelling: in preschool in L2/Hebrew, or at home in Russian, as telling stories at home is part of home language development (Schik et al., 2017). Another reason for enhanced macrostructure in L2 for L1D children is transfer from the stronger (here L1/Russian) language to the 'weaker' language (Squires et al., 2014). Transfer of narrative skills and its direction is examined in research manipulating bilingual intervention and monitoring children's gains in both languages (e.g., Petersen et al., 2016).

Several studies on bilingual macrostructure have shown differences across languages and offer a culture-specific explanation for such differences (Fiestas & Peña, 2004; Rezzonico et al., 2016). Story telling traditions vary across cultures and impact the development of narrative organization in children. For example, Spanish-speaking families may rely on oral narration by caregivers, while other cultures may show greater reliance on books and pictures (Schick et al., 2017; Sparks & Reese, 2013). Bilingual children exposed to Russian and Hebrew in Israel are culturally similar, especially regarding exposure to stories. In addition, the story stimuli here were developed to be suitable for children exposed to different cultures (Gagarina et al., 2012). Thus, it is not surprising that the same structural elements appeared in stories from both languages.

To summarize, in line with previous research, our bilingual children produced fewer than half of the MAIN narrative macrostructure elements and did so similarly across their two languages. Language dominance, a factor rarely examined for macrostructure abilities, did not impact narrative macrostructure.

5.3 Internal State Terms across languages and dominance groups

The ratio of IST tokens and types did not differ as a function of dominance or language. This finding is in line with the literature showing cross-linguistic similarity of ISTs (Lucero, 2015; Otwinowska et al., 2018). Fusté-Herrmann et al. (2006) claimed that the ability to produce different ISTs to convey different functions leads to a more advanced level of macrostructure. In the current results, 10% of ISTs were emotional terms, and these were all used to convey a Reaction, a story structure category rarely produced by preschool children (Shapiro & Hudson, 1991). In contrast, the most frequent IST categories, perceptual (Russian *uvidela*, Hebrew *raata*, 'see') and motivational (Russian *xotela*, Hebrew: racta, 'wanted')

terms, were used by children to convey other macrostructure elements (Initiating Events, Goals). ISTs reflecting Reaction were mainly produced in the final episode as a way to end the story (Russian: *vse kozljonki byli rady potomu chto spaslis*, 'all baby goats were happy because they were safe'). Failure to use ISTs more frequently as Reactions in other episodes should not necessarily be attributed to a difficulty in producing ISTs. Rather, it may be due to an undeveloped skill in talking about a character's emotions. This skill is said to develop in later school years (Berman & Slobin, 1994; Burns et al., 2012). For example, Bohnacker (2016) reported that younger 5-year-old children rarely used ISTs as an Initiating Event, but 6–7-year-old children did produce twice as many ISTs as younger children.

Evidence for the effect of dominance was seen in the narratives of L1D children who used more trait terms (Russian *xitryj*, Hebrew *rasha*, 'mean') and more emotional terms (Russian *grustnyj*, Hebrew *acuv*, 'sad') in L1/Russian than in L2/Hebrew. They also used more physiological terms (Russian *golodnyj*, Hebrew *raev*, 'hungry') in L2/Hebrew. L2D children show similar frequencies for most IST categories across L1 and L2, but more mental terms (Russian *dumal*, Hebrew *xašav*, 'thought') in L2/Hebrew than in L1/Russian. Although these trends were not statistically significant, they are similar to the results reported by Altman et al. (2016) for English-Hebrew bilingual children, who used significantly more mental ISTs in L2 than in L1 narratives. The authors explain this finding by more frequent exposure to mental state terms in children's input in preschool.

In sum, ISTs, which combine microstructure linguistic knowledge and macrostructure cognitive knowledge, show subtle sensitivity to the effects of dominance in the types of ISTs used.

5.4 Relationships among microstructure, macrostructure, and ISTs across languages and dominance groups

In a further effort to explore the effects of language dominance, *cross-language* and *within-language* correlations were examined for a range of narrative measures. Only a few studies have addressed the interrelationship of narrative measures across and within languages (Bitetti et al., 2020; Cereijido & Gutiérrez-Clellen, 2009; Lucero, 2015). Bitetti et al. (2020) examined these relationships in bilingual Spanish-English preschool children and did not find *cross-language* relationships nor effects for dominance. They did show, however, *within-language* associations between microstructure and macrostructure measures for both languages. Similarly, Lucero (2015) reported *within-language* associations between vocabulary and macrostructure for Spanish and English but no *cross-language* correlations. Both studies used the Narrative Story Scheme

(Heilmann et al., 2010) to assess macrostructure. This measure includes SG components as well as other narrative elements, such as character development, the use of metalinguistic and metacognitive verbs, and cohesion. Since different components comprise macrostructure assessment in MAIN (Gagarina et al., 2012) and in the Narrative Story Scheme, direct comparison is not possible.

The lack of cross-linguistic associations may be attributed to the impact of dominance. Unlike Bitetti et al. (2020), who did not find an effect of dominance on *cross-linguistic* relationships, the current research did find *cross-language* associations for L1D bilinguals who had stronger proficiency in L1/Russian but demonstrate similarity across most narrative measures. Cross-linguistic similarity in the L1D group was manifested in microstructure, macrostructure, ISTs, as well as in the correlation between L1 and L2 measures. Thus, the effect of dominance is subtly revealed in the presence of associations across languages for L1D bilinguals (but not for L2D bilinguals). Dominance in L1/Russian might have allowed the L1D children to transfer narrative knowledge to a somewhat greater extent than what was observed for L2D children. These results are in line with narrative intervention studies, where transfer is observed from L1 (often the stronger language) to L2 (often the weaker language). Cross-linguistic differences found for the L2D bilinguals were partially manifested in microstructure. Finally, the results in the present study show that high proficiency in L1/Russian in the L1D group did not hinder acquisition of L2/Hebrew narrative skills and may have even had an enhancing effect on narrative skills in both languages.

6. Conclusions

By and large dominance was found to have a limited effect on narrative performance among Russian-Hebrew bilingual children. The results are consistent with research showing that bilingual children with limited language proficiency in one of their languages can use different expressive devices to convey principal narrative elements. Use of ISTs to convey Story Structure elements and production of a variety of Story Grammar elements in the second and third episode show that children narrated coherent narratives rather than simply describing elements in each episode separately. Performance on a narrative retelling task differs from standardized proficiency test performance. It reflects a more encompassing picture of bilingual performance, since it involves linguistic, cognitive, and expressive levels of processing, all of which are interrelated and reflected in narrative microstructure, macrostructure, and the use if ISTs.

Beyond the need for research with larger numbers of participants in each dominance group, future research should examine the interaction of quantitative

and qualitative exposure measures and dominance. Furthermore, additional microstructure measures should be included, such as temporality, complex syntactic structures, referentiality, and errors. The effect of dominance on these narrative features will certainly provide rich information on bilingual narrative development.

Funding

Research funded by Israel Science Foundation (1716/19) to Carmit Altman.

References

Abutbul-Oz, H., Armon-Lotem, S., & Walters, J. (2012). *Bilingual Parents Questionnaire (BIPAQ)*. Bar-Ilan University.

Altman, C., Harel, E., Meir, N., Iluz-Cohen, P., Walters, J., & Armon-Lotem, S. (2021). Using a monolingual screening test for assessing bilingual children. *Clinical Linguistics & Phonetics*, https://doi.org/10.1080/02699206.2021.2000644

Altman, C., Armon-Lotem, S., Fichman, S., & Walters, J. (2016). Macrostructure, microstructure, and mental state terms in the narratives of English-Hebrew bilingual preschool children with and without specific language impairment. *Applied Psycholinguistics*, *37*(1), 165–193. https://doi.org/10.1017/S0142716415000466

Altman, C., Goldstein, T. & Armon-Lotem, S. (2018). Vocabulary, metalinguistic awareness and language dominance among bilingual preschool children. *Frontiers in Psychology*, *9*, 1953. https://doi.org/10.3389/fpsyg.2018.01953

Armon-Lotem, S., & Meir, N. (2016). Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in bilingual children: evidence from Russian and Hebrew. *International Journal of Language & Communication Disorders*, *51*(6), 715–731. https://doi.org/10.1111/1460-6984.12242

Berman, R.A. (1978). Early verbs: Comments on how and why a child uses his first words. *International Journal of Psycholinguistics*, *5*, 21–39.

Berman, R.A., & Katzenberger, I. (2004). Form and function in introducing narrative and expository texts: A developmental perspective. *Discourse Processes*, *38*(1), 57–94. https://doi.org/10.1207/s15326950dp3801_3

Berman, R.A., & Slobin, D.I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Lawrence Erlbaum.

Bernardini, P., & Schlyter, S. (2004). Growing syntactic structure and code-mixing in the weaker language: The Ivy Hypothesis. *Bilingualism: Language and Cognition*, *7*(1), 49–69. https://doi.org/10.1017/S1366728904001270

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*(1), 1–28. 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

Burns, F.A., de Villiers, P.A., Pearson, B.Z., & Champion, T.B. (2012). Dialect-neutral indices of narrative cohesion and evaluation. *Language, Speech, and Hearing Services in Schools*, *43*(2), 132–152. https://doi.org/10.1044/0161-1461(2011/10-0101)

Doron, E. (2000). Word order in Hebrew. In J. Lecarme, J. Lowenstamm, & U. Shlonsky (Eds.), *Studies in Afroasiatic Grammar* (pp. 41–56). Holland Academic Graphics. https://doi.org/10.1075/cilt.202.03dor

Fichman, S., Altman, C., Voloskovich, A., Armon-Lotem, S., & Walters, J. (2017). Story grammar elements and causal relations in the narratives of Russian-Hebrew bilingual children with SLI and typical language development. *Journal of Communication Disorders*, *69*(1), 72–93. https://doi.org/10.1016/j.jcomdis.2017.08.001

Fichman, S., Armon-Lotem, S., Walters, J., & Altman, C. (2021). Story grammar elements, mental state terms and clausal distance in the expression of enabling relations in narratives of bilingual preschool children. *Discourse Processes*. https://doi.org/10.1080/0163853X.2021.1972391

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

Florit, E., Roch, M., & Levorato, M.C. (2011). Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills. *Discourse Processes*, *48*(2), 119–138. https://doi.org/10.1080/0163853X.2010.494244

Fusté-Herrmann, B., Silliman, E.R., Bahr, R.H., Fasnacht, K.S., & Federico, J.E. (2006). Mental state verb production in the oral narratives of English- and Spanish-speaking preadolescents: An exploratory study of lexical diversity and depth. *Learning Disabilities Research & Practice*, *21*(1), 44–60. https://doi.org/10.1111/j.1540-5826.2006.00206.x

Gagarina, N. (2011). Acquisition and loss of L1 in a Russian-German bilingual child: A case study. In S. Cejtlin (Ed.), *Monolingual and bilingual path to language* (pp. 137–163). Jazyki slavjanskoj kul'tury.

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

Gagarina, N., Klassert, A., & Topaj, N. (2010). Russian language proficiency test for multilingual children. *ZAS Papers in Linguistics*, *54*. ZAS.

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*. ZAS.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2015). Assessment of narrative abilities in bilingual children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 243–276). Multilingual Matters. https://doi.org/10.21832/9781783093137-011

Gagarina, N., Klop, D., Tsimpli, I.M., & Walters, J. (2016). Narrative abilities in bilingual children. *Applied Psycholinguistics*, *37*(1), 11–17. https://doi.org/10.1017/S0142716415000399

Goralnik, E. (1995). *Language screening test for Hebrew-speaking preschool children*. Gai Agencies.

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

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)

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

Justice, L. M., Bowles, R., Pence, K., & Gosse, C. (2010). A scalable tool for assessing children's language abilities within a narrative context: The NAP (Narrative Assessment Protocol). *Early Childhood Research Quarterly*, *25*(2), 218–234. https://doi.org/10.1016/j.ecresq.2009.11.002

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

Khan, K. S., Gugiu, M. R., Justice, L. M., Bowles, R. P., Skibbe, L. E., & Piasta, S. B. (2016). Age-related progressions in story structure in young children's narratives. *Journal of Speech, Language, and Hearing Research*, *59*(6), 1395–1408. https://doi.org/10.1044/2016_JSLHR-L-15-0275

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

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

Lindgren, J., & Bohnacker, U. (2020). 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

Lucero, A. (2015). Cross-linguistic lexical, grammatical, and discourse performance on oral narrative retells among young Spanish speakers. *Child Development*, *86*(5), 1419–1433. https://doi.org/10.1111/cdev.12387

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. Erlbaum.

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

Mayer, M. (1969). *Frog, Where Are You?* Dial Press.

Montanari, S. (2004). The development of narrative competence in the L1 and L2 of Spanish-English bilingual children. *International Journal of Bilingualism*, *8*(4), 449–497. https://doi.org/10.1177/13670069040080040301

Nicolopoulou, A., Ilgaz, H., & Shiro, M. (2021). "And they had a big, big, very long fight:" The development of evaluative language in preschoolers' oral fictional stories told in a peer-group context. *Journal of Child Language*, 1–30. https://doi.org/10.1017/S0305000921000209

Otwinowska, A., Mieszkowska, K., Białecka-Pikul, M., Opacki, M. & Haman, E. (2018). Retelling a model story improves the narratives of Polish-English bilingual children. *International Journal of Bilingual Education and Bilingualism*, *23*(9), 1083–1107. https://doi.org/10.1080/13670050.2018.1434124

Otwinowska, A., Opacki, M., Mieszkowska, K., Białecka-Pikul, M., Wodniecka, Z., & Haman, E. (2020). Polish-English bilingual children overuse referential markers: MLCU inflation in Polish-language narratives. *First Language*, https://doi.org/10.1177/0142723720933769

Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). French-English Bilingual Children With SLI. *Journal of Speech, Language, and Hearing Research*, *46*(1), 113–127. https://doi.org/10.1044/1092-4388(2003/009)

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* (135–174). Multilingual Matters. https://doi.org/10.21832/9781853595721-008

Petersen, D.B., Thompsen, B., Guiberson, M.M., & Spencer, T.D. (2016). Cross-linguistic interactions from second language to first language as the result of individualized narrative language intervention with children with and without language impairment. *Applied Psycholinguistics*, *37*(3), 703–724. https://doi.org/10.1017/S0142716415000211

Rezzonico, S., Goldberg, A., Mak, K.K.Y., Yap, S., Milburn, T., Belletti, A., & Girolametto, L. (2016). Narratives in two languages: Storytelling of bilingual Cantonese-English preschoolers. *Journal of Speech, Language, and Hearing Research*, *59*(3), 521–532. https://doi.org/10.1044/2015_JSLHR-L-15-0052

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

Schick, A.R., Melzi, G., & Obregón, J. (2017). The bidirectional nature of narrative scaffolding: Latino caregivers' elaboration while creating stories from a picture book. *First Language*, *37*(3), 301–316. https://doi.org/10.1177/0142723716689692

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.

Shapiro, L.R., & Hudson, J.A. (1991). Tell me a make-believe story: Coherence and cohesion in young children's picture-elicited narratives. *Developmental Psychology*, *27*(6), 960–974. https://doi.org/10.1037/0012-1649.27.6.960

Silliman, E.R., Bahr, R.H., Brea, M.R., Hnath-Chisolm, T., & Mahecha, N.R. (2002). Spanish and English proficiency in the linguistic encoding of mental states in narrative retellings. *Linguistics and Education*, *13*(2), 199–234. https://doi.org/10.1016/S0898-5898(01)00062-6

Simon-Cereijido, G. & Gutiérrez-Clellen, V.F. (2009). A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains. *Applied Psycholinguistics*, *30*(2), 315–337. https://doi.org/10.1017/S0142716409090134

Sparks, A., & Reese, E. (2013). From reminiscing to reading: Home contributions to children's developing language and literacy in low-income families. *First Language*, *33*(1), 89–109. https://doi.org/10.1177/0142723711433583

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 and 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.), *New directions in discourse processing* (Vol. 2 pp. 53–120). Ablex.

Symons, D. K., Peterson, C. C., Slaughter, V., Roche, J., & Doyle, E. (2005). Theory of mind and mental state discourse during book reading and story-telling tasks. *British Journal of Developmental Psychology*, *23*(1), 81–102. https://doi.org/10.1348/026151004X21080

Timberlake, A. (2004). *A reference grammar of Russian*. Cambridge University Press.

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

Trabasso, T., Van den Broek, P., & Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, *12*(1), 1–25. https://doi.org/10.1080/01638538909544717

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)

Ukrainetz, T. A., Justice, L. M., Kaderavek, J. N., Eisenberg, S. L., Gillam, R. B., & Harm, H. M. (2005). The development of expressive elaboration in fictional narratives. *Journal of Speech, Language, and Hearing Research*, *48*(6), 1363–1377. https://doi.org/10.1044/1092-4388(2005/095)

Westby, C. E. (2005). Assessing and facilitating text comprehension problems. In H. Catts & A. Kamhi (Eds.), *Language and reading disabilities* (pp. 157–232). Allyn & Bacon.

Yip, V., & Matthews, S. (2006). Assessing language dominance in bilingual acquisition: A case for mean length utterance differentials. *Language Assessment Quarterly: An International Journal*, *3*(2), 97–116. https://doi.org/10.1207/s15434311laq0302_2

Address for correspondence

Sveta Fichman
Departments of English Literature and Linguistics
Bar-Ilan University
Ramat Gan, 52900
Israel

fichmans@talpiot.ac.il
https://orcid.org/0000-0002-4946-4506

Co-author information

Joel Walters
Departments of English Literature and Linguistics
Bar-Ilan University
waltej49@gmail.com

Sharon Armon-Lotem
Departments of English Literature and Linguistics
Gonda Multidisciplinary Brain Research Center
Bar-Ilan University
sharon.armonlotem@biu.ac.il

Carmit Altman
Gonda Multidisciplinary Brain Research Center
Faculty of Education
Bar-Ilan University
Carmit.Altman@biu.ac.il

Publication history

Date received: 16 May 2020
Date accepted: 14 March 2022
Published online: 31 May 2022