

ORIGINAL ARTICLE

Bilingual narrative intervention among Russian-Hebrew bilingual preschool children with developmental language disorder

Adaya Dadon Jorno¹ , Sveta Fichman² , Joel Walters²  and Carmit Altman^{1,3} 

¹Faculty of Education, Bar-Ilan University, Ramat Gan, Israel, ²Department of English Literature & Linguistics, Bar-Ilan University, Ramat Gan, Israel and ³Gonda Brain Science Research Center, Bar-Ilan University, Ramat Gan, Israel

Corresponding author: Carmit Altman; Email: carmit.altman@biu.ac.il

(Received 2 October 2023; revised 19 November 2024; accepted 11 January 2025)

Abstract

Changes in narrative skills among Russian-Hebrew bilingual preschoolers with and without developmental language disorder (DLD) were examined following Bilingual Narrative Intervention. Eight children with DLD and nine typically developing children participated in two six-session intervention blocks, first in the home language HL/Russian and then in the societal language SL/Hebrew. Intervention sessions involved retelling single-episode stories accompanied by icons/gestures, repetition, and peer interaction. Narrative skills were assessed at four-time points. Results showed that while typically developing children performed better overall, both groups followed similar trajectories. Children performed better in HL/Russian across all assessments. Macrostructure improved in both groups after HL/Russian intervention, particularly for “Feeling” and “Goal” elements. Bilingual typical language development children showed higher lexical diversity, with significant improvement following HL/Russian intervention. Children with bilingual DLD made more errors in HL/Russian, but similar error rates emerged between groups for SL/Hebrew. Earlier age of onset of bilingualism correlated with better macrostructure in SL/Hebrew but not in HL/Russian. Findings underscore the need for tailored intervention in both languages which considers clinical status and bilingual children’s background.

Keywords: Bilingualism; child language disorders; discourse; narrative

Introduction

The present paper brings together research from three areas in early childhood education: developmental language disorder (DLD), narrative development, and cross-linguistic transfer. DLD is diagnosed when a child shows below-age performance on a range of linguistic tasks while having intact hearing, no intellectual delay, no sensory impairment, or other neurological conditions

(Leonard, 2017). Bilingual children need to meet these exclusionary and inclusionary criteria in both their languages to be diagnosed with bilingual DLD (BiDLD) (Abutbul-Oz & Armon-Lotem, 2022; Armon-Lotem *et al.*, 2021). A coherent narrative involves macrostructure (including reference to characters, a problem, feelings, goal, attempt, outcome, and internal response) and microstructure elements (vocabulary and syntax), and the development of both components is crucial for successful storytelling (Gillam *et al.*, 2018; Pico *et al.*, 2021). DLD poses challenges to the acquisition of both macrostructure and microstructure narrative skills among monolingual and bilingual children (e.g., Boerma *et al.*, 2016; Cleave *et al.*, 2010; Thordardottir *et al.*, 2015; Méndez *et al.*, 2015). Investigating the challenges and implementing narrative intervention are of special importance for children with BiDLD, whose narrative development is weak in the two languages, and may benefit from cross-linguistic transfer when instruction in one language enhances the other language's narrative performance (Boerma *et al.*, 2016; Squires *et al.*, 2014). The literature review that follows summarizes findings on narrative skills in children with BiDLD and reviews narrative intervention studies in bilinguals.

Narrative skills in bilingual children with DLD

Children with DLD demonstrate weak expressive language skills, which are crucial for successful narrative performance (Bishop, 2017; Leonard, 2017). Most studies on narrative skills among monolingual children have shown weaker performance of children with DLD than their peers with typical language development (TLD) for both macrostructure and microstructure (Duinmeijer *et al.*, 2012; Fey *et al.*, 2004; Hayward & Schneider, 2000; Reilly *et al.*, 2004; Scott & Windsor, 2000). Studies with bilingual children have shown inconsistent results. For macrostructure, some studies have found that children with BiDLD produced fewer Story Grammar (SG) elements than peers with bilingual TLD (BiTLD) (Boerma *et al.*, 2016; Fichman *et al.*, 2017; Gagarina *et al.*, 2019; Hao *et al.*, 2018; Govindarajan & Paradis, 2019; Paradis *et al.*, 2013; Rezzonico *et al.*, 2015; Squires *et al.*, 2014) and used fewer causal relations in their narratives (Fichman *et al.*, 2017; Kupersmitt & Armon-Lotem, 2019). Other studies showed similarities between the groups (Altman *et al.*, 2016; Iluz-Cohen & Walters, 2012; Tsimpli *et al.*, 2016). For the results on microstructure, there is a greater consensus in research findings, where children with BiDLD exhibit poorer performance on morphosyntactic features, sentence complexity measures, and narrative length in comparison to children with BiTLD in both languages (Iluz-Cohen, & Walters, 2012; Paradis *et al.*, 2013; Rezzonico *et al.*, 2015; Squires *et al.*, 2014; Tsimpli *et al.*, 2016).

One possible reason for the lack of consensus in these findings can be the high variability of factors affecting narrative performance, such as exposure to both languages and proficiency in each language (Kapalková *et al.*, 2016). Bilinguals may benefit from the interaction of their two languages due to possible cross-linguistic transfer (Hipfner-Boucher *et al.*, 2014; Squires *et al.*, 2014), and this interaction has not been completely understood. In addition, the gap between microstructure and macrostructure performance, which is more evident in bilinguals and in children with DLD due to their different levels of proficiency, contributes to variability in the

results. Another uniquely bilingual measure that may affect language development in bilingual children is the Age of onset of Bilingualism (AoB), the age a child begins to receive exposure to the second language, which is often the societal language (Armon-Lotem et al., 2021; Golberg et al., 2008; Paradis et al., 2021; Unsworth, 2013). Later AoB implies longer periods of (mostly) monolingual development in the home, or heritage, language, whereas earlier AoB is associated with simultaneous early development of two systems (Paradis, 2023). While studies agree that exposure measures are important in bilingual typical narrative development, their impact among children with BiDLD is not clear and has not been sufficiently addressed in research on narrative development. The role of AoB in children with DLD is of special importance because these children often have a later onset of speech and slower development (Leonard, 2017). For example, Govindarajan and Paradis (2019) found that exposure factors, such as length of exposure and richness of the SL environment, predicted the narrative performance of children with BiTLD, but not that of children with BiDLD.

Narrative intervention in bilingual children with DLD

Children with BiDLD, who speak a home language different from the language of the instruction in preschool/school, usually receive minimal academic support in the home language (Zehler et al., 2003). In most cases, intervention is provided only in the societal language, due to the absence of bilingual professionals who can provide treatment in both languages (ASHA, 2010; Jordaan, 2008) as well as the common practice of providing treatment exclusively in the societal language for social integration and academic success (Capin et al., 2023; Daelman et al., 2023; Drysdale et al., 2015; Kay-Raining Bird et al., 2012). We know of four **bilingual narrative intervention** studies that have implemented intervention **in both languages** of children with TLD (Armon-Lotem et al., 2021; Lipner et al., 2021; Petersen et al., 2016; Spencer et al., 2019), two bilingual intervention studies with children with DLD (Lugo-Neris et al., 2015; Thordardottir et al., 2015), and two systematic reviews (Kk Nair et al., 2023; Pico et al., 2021). Bilingual intervention studies have reported advantages of bilingual over monolingual intervention (Gutierrez-Clellen et al., 2014; Gutierrez-Clellen et al., 2012; Thordardottir et al., 2015) and/or evidence of cross-linguistic transfer (e.g., Kambanaros et al., 2017; Lü et al., 2023). Thordardottir et al. (2015) examined 29 children from diverse backgrounds in three conditions: monolingual French, bilingual (where home language instruction was provided by parents), or no intervention. The intervention was delivered via narratives targeting vocabulary and syntactic skills. Children's vocabulary and syntax improved, but only the vocabulary gains were attributed to intervention. The study concluded that the bilingual treatment condition, when conducted through collaboration with parents, was not effective in creating a sufficiently intense bilingual context. Lugo-Neris et al. (2015) conducted 24 small-group intervention sessions with Spanish–English children with BiDLD in both languages targeting macrostructure (story grammar) and microstructure (vocabulary and syntax). Gains were reported for macrostructure and vocabulary in both languages.

Research has not sufficiently examined, which SG elements were most susceptible to improvement after the intervention. Intervention studies focusing on SG elements reported that children produced more complete stories including such elements as an Initiating Event, Attempt, and Consequence (Petersen & Spencer, 2016). Studies examining macrostructure skills of children with DLD at one point in time have suggested that Goals are acquired later and are more challenging elements to produce (Altman et al., 2016). Some studies suggested that children with BiDLD used fewer SG elements overall (Altman et al., 2016; Altman et al., 2024; Fichman et al., 2017). Goals and Internal Responses are less “concrete,” and it may be more challenging to increase their use as a result of intervention (Khan et al., 2016).

A unique feature of bilingual intervention delivered in two languages is the potential for cross-linguistic transfer of skills from one language to another (Ebert et al., 2014; Harvey et al., 2021; Isurin, 2005; Kk Nair et al., 2023; Petersen et al., 2016). Narrative knowledge plays a crucial role in school and enhancing this knowledge via transfer, in addition to the instruction in the school language, may have important educational benefits, especially for children with BiDLD. Methodologically, transfer has been argued for when there was an improvement in the home language in the experimental group following intervention in the societal language, but not in the control group (Petersen et al., 2016). Harvey et al. (2021) showed transfer because of improvement in the SL in the bilingual treatment group which was similar to the improvement in the “SL only” treatment group, despite having had half the intervention time. Lü et al. (2023) claimed transfer of definition skills when HL and SL scores significantly correlated. Armon-Lotem et al., (2021) and Lipner et al (2021) showed transfer in vocabulary knowledge by examining improvement in HL/SL after intervention in the other language using different set of words in each language (Armon-Lotem et al., 2021). Lipner et al. (2021) showed bidirectional transfer of lexical knowledge. Based on these studies, cross-linguistic transfer of narrative skills is possible, but empirical evidence is still limited, especially in children with DLD and in the context of narrative intervention focusing on macrostructure and microstructure skills.

In sum, only a handful of studies performed narrative intervention while focusing on macrostructure and microstructure, in both languages. Some studies suggested the possibility of cross-linguistic transfer of macrostructure skills. The present paper examined narrative performance *in both languages* following two blocks of intervention, first in the home language and then in the societal languages.

The current research

The study examined whether narrative macrostructure and microstructure skills develop in both languages of Russian-Hebrew preschool children with BiDLD and their peers with BiTLD, following a bilingual narrative intervention (BINARI). Russian was children’s Home Language (HL), and Hebrew was their Societal Language (SL). All children were clinically referred to receive treatment by a Speech-Language Pathologist (SLP). The study also addressed a possible cross-

linguistic transfer of macrostructure and microstructure skills. The following research questions were addressed.

1. **Group.** To what extent does narrative performance, evaluated using macrostructure and microstructure measures, differ for children with BiDLD and BiTLD at all stages of a bilingual narrative intervention procedure?
2. **Language.** Are there differences in macrostructure and microstructure in BiDLD and BiTLD children's narratives in HL/Russian vs. SL/Hebrew?
3. **Within vs. cross-language change.** To what extent are there changes in macrostructure and microstructure skills across four PM time points when the language of the intervention and the language of the PM match (within-language) and when the language of the intervention and the PM do not match (across-languages)?
4. **Exposure.** To what extent is the AoB related to macrostructure and microstructure performance of bilingual children with BiDLD and BiTLD?

Bilingual children with DLD are predicted to perform like children with BiTLD for macrostructure features and to show weaker performance for microstructure features (in particular for verbal productivity and morpho-syntax), across all stages of the intervention (e.g., Armon-Lotem et al., 2020; Iluz-Cohen, & Walters, 2012). Children may perform better in HL/Russian than in SL/Hebrew at all time points (Lipner et al., 2021). Bilingual narrative intervention is expected to result in changes within and across languages, such that a change across languages would be compatible with the presence of cross-language transfer (Armon-Lotem et al., 2021). Children exposed earlier to the SL may show better performance in the HL (Lipner et al., 2021).

Method

Participants

Seventeen Russian-Hebrew bilingual children at risk for DLD were included in the study. All children were referred by a professional in the child's educational setting for treatment to the clinic from which the children were recruited. Referrals were made because of weak performance in SL/Hebrew. Initially, 24 bilingual children were screened in the clinic; four were excluded since they could not carry on a basic conversation in HL/Russian, and at the end of the intervention, three children were excluded due to infrequent attendance. Parents provided written informed consent, and children expressed oral assent. They were informed that they could discontinue participation at any time. The study was approved by IRB and the office of the Chief Scientist of the Ministry of Education.

BiDLD criteria

Due to unavailability of bilingual SLPs in Israeli clinics and absence of bilingual standardized assessment tools, children are usually screened in clinics using monolingual standardized tests in SL/Hebrew. As a result, children are often over-diagnosed with DLD, which leads to incorrect placement of children in special education preschools (Abutbul-Oz & Armon-Lotem, 2022). In the current research,

Table 1. Age, AoB, and proficiency scores in children with developmental language disorder (DLD) and typical language development (TLD)

	DLD			TLD		
	<i>M (SD)</i>	<i>Min</i>	<i>Max</i>	<i>M (SD)</i>	<i>Min</i>	<i>Max</i>
Age	67.50 (5.48)	59	76	67.56 (4.95)	60	74
AoB (months)	31.50 (16.89)	0	48	36.00 (14.70)	0	48
Language Proficiency in HL*	−2.56 (0.78)	−3.66	−1.59	−0.13 (0.77)	−0.79	1.04
Language Proficiency in SL*	−3.81 (1.52)	−6.60	−1.90	−5.05 (2.22)	−8.80	1.84

Note: Russian (HL) proficiency was assessed by the *Russian Language Proficiency Test for Multilingual Children* (Gagarina *et al.*, 2010); Hebrew (SL) proficiency was assessed by the *Goralnik Screening Test for Hebrew* (Goralnik, 1995); AoB = Age of onset of Bilingualism.

*z-score based on local bilingual standards for Russian-Hebrew bilingual children.

children’s language abilities were screened in both languages using tests normalized for bilingual children in HL and SL (see Materials section). As a result of the screening and applying bilingual local standards (Altman *et al.*, 2021), eight children performed 1.25SDs below the standard in both languages; these children were classified as children with BiDLD, based on the definition of bilingual DLD (Armon-Lotem, 2014). Nine children scored 1.25SDs below the mean in SL/Hebrew and 1.25SDs above the mean in HL/Russian and were classified as children with BiTLD for the purpose of the present study. All 17 parents reported that they were concerned about their child’s language skills. Ten of 17 parents (6 DLD and 4 TLD) reported that their child had some kind of treatment in other clinics in SL/Hebrew.

Bilingualism criteria

Children were considered bilingual if they were exposed to HL/Russian from birth in a home where Russian was spoken by at least one caregiver and could carry on a conversation in both languages. The bilingual status was reported by parents (see Materials section) and verified by the Russian-speaking and the Hebrew-speaking experimenters. If a child could not hold a spontaneous conversation in HL or SL, she was excluded from the research. All children had at least 24 months of exposure to Hebrew, and all attended preschools (prior to school entry) where Hebrew was the language of instruction and the main language of social interaction.

The final group of participants included 11 boys and 6 girls; their ages ranged between 59 and 76 months ($M = 67.5$; $SD = 5.03$). Children’s AoB and proficiency information are presented in Table 1.

Materials

Background questionnaire

A parental questionnaire was adapted from the *Bilingual Parental Questionnaire* (Abutbul-Oz & Armon-Lotem, 2022), eliciting information about age, developmental milestones, exposure to HL and SL, language preferences, and speech-language treatment.

Language screening

To assess proficiency in Russian, *The Russian Language Proficiency Test for Multilingual Children* (Gagarina et al., 2019) was administered. It included measures of expressive (noun/verb naming, production of case and verb inflections) and receptive language (comprehension of grammatical constructions, nouns, and verbs). Hebrew proficiency was assessed using the *Goralnik Screening Test for Hebrew* (Goralnik, 1995). It included six sub-tests: vocabulary, sentence repetition, comprehension, oral expression, pronunciation, and storytelling. Local bilingual standards were applied for scoring (Altman et al., 2021; Armon-Lotem & Meir, 2016). Each test had been used with large samples of bilingual children prior to the establishment of the standards (Altman et al., 2021; Fichman et al., 2017; 2023).

Narrative instrument

Narrative intervention materials were based on Spencer et al. (2019) and were adapted culturally to the Israeli bilingual environment (Armon-Lotem et al., 2021; Lipner et al., 2021). Overall, 20 sets of pictures were used, eight sets for Progress Monitoring, (four in Russian and four in Hebrew) and 12 sets for the narrative intervention procedure (six in Russian and six in Hebrew), see the Procedure section. In all testing sessions and in all intervention sessions, different stories were used; however, all the stories were designed to be similar in terms of length, macrostructure, and language complexity. In terms of macrostructure, all stories were based on seven SG elements, appropriate for ages 5–6. The elements were character, problem, feeling, goal, action, ending, and internal response. Each set had a narrative script. The scripts were similar in terms of lexical and syntactic complexity, such that they included four subordinate clauses with temporal pronouns (e.g., “**After** Yael reached behind her coat, she found her hidden scarf”) and causal connectors (e.g., “Yael was frustrated **because** her red scarf was hidden”). All stories were based on themes relevant for the age of 5–6, such as addressing the difficulty of reaching a book on a shelf or looking for a lost thing.

Procedure

Research design

The study employed a single arm within-subject pre-post design, where the main goal was to monitor the progress in narrative skills in both languages of clinically referred bilingual children classified as BiDLD and BiTLD. The study did not use a randomized controlled trial design, primarily because this was not ethically appropriate in a setting where the participants were clinically referred. All children who are referred to the clinic receive treatment, and if we were to give a different language treatment to a group of children, they could not serve as a control group. Thus, the most appropriate research design was the single arm within-subject pre-post design, where children served as their own control, both within and across languages. Such a design is appropriate to implement in a kindergarten or a clinic (Green & Klecan-Aker, 2012).

Screening tests in HL and SL yielded two groups, with children performing below age-appropriate norms in both languages being assigned to the BiDLD group and

children performing below age-appropriate norms in one of the languages to the BiTLD group. This allowed us to address the first research question comparing macrostructure and microstructure performance of children from each of the two groups across four-time points.

The comparison across languages in both groups allowed us to examine differences between HL and SL in terms of narrative performance at four-time points, which was at the center of the second research question. Two blocks of intervention were conducted: the first block was conducted in HL, and the second block in SL. Narrative skills were tested before and after each block in both languages, allowing us to establish whether narrative performance changed in the language of the intervention (within-language change) or in both languages (potential cross-linguistic transfer). This design aimed to address the third research question.

The overall procedure spanned over 12–16 weeks and is presented in the schematic outline of Bilingual NARrative Intervention (BINARI).

Prior to the intervention sessions, two language screening sessions were conducted in HL and SL over two weeks, with a week gap between them. The program included four progress monitoring (PM) sessions in each language (eight PM sessions), which were conducted before intervention, after intervention in HL, after intervention in SL, and six weeks after intervention was over. There were 12 intervention sessions (six in HL and six in SL), such that the six intervention sessions (first in HL and then in SL) were conducted over three weeks in each language. Thus, each participant participated in 22 sessions which included two screening sessions, eight PM sessions, and 12 intervention sessions.

Progress monitoring (PM)

Participants told narratives using a retelling mode, one in each language (HL and SL) at each of the four-time points, as follows: PM1 was aimed to provide pre-intervention (baseline) performance; PM2 followed the intervention in HL/Russian and was conducted in HL and SL; PM3 was conducted in both languages following the intervention in SL/Hebrew; and in PM4 children told stories in each language six weeks after the last intervention session. The elicitation materials (story scripts and pictures) used for retelling in the PMs were single-episode stories, culturally adapted from *Puente de Cuentos* (Spencer *et al.*, 2017). The stimulus stories were based on Stein and Glenn's (1979) story grammar model. The HL/Russian stories had an average of 71 words, and the SL/Hebrew had an average of 68 words. The narratives in both languages were matched for plot structure, syntactic complexity, and mental state terms.

In each PM session, the child first looked at the pictures while the experimenter told the story. Next, the children were asked to retell the story with the pictures laid out in front of them. The experimenters were trained by the second and fourth authors and were instructed not to interfere with the child's storytelling. If the child hesitated while retelling the story, the child was encouraged to continue, but no verbal prompts were given. The PM sessions were conducted individually in the clinic.

Intervention

Intervention was conducted in small groups of 3–4 children in the clinic. It consisted of two blocks of six sessions each, one block in HL (Russian) and the other in SL (Hebrew) for a total of 12 sessions with 12 different stories, all with the same structural features as those used for the PM sessions. Intervention sessions were conducted twice a week, and each language block lasted three weeks. Each intervention session lasted 20–25 minutes. Sessions in each language were conducted by a native speaker of the language. A student of Speech-Language Pathology conducted the intervention in Russian, and a certified SLP conducted interventions in Hebrew (the first author). The intervention procedure consisted of five steps: 1) story modeling and introduction of icons for targeted elements; 2) pairing icons with gestures; 3) group retelling; 4) explicit focus on target features (for macrostructure: character, problem, feeling, goal, attempt, outcome, and internal response; for microstructure: productivity, accuracy, and complexity); and 5) individual retelling. For each intervention session, the experimenter first read the story, while the children were looking at the pictures. Then, the experimenter explained each of the seven narrative elements (SG elements): character, problem, feeling, goal, attempt, outcome, or internal response. The explanation involved defining each SG element, showing an icon for that element, and demonstrating a gesture associated with each icon (Spencer et al., 2019). The icons and gestures were used throughout all intervention sessions. During intervention, the children were asked to repeat parts of the story together with the experimenter, which yielded 4–6 opportunities for each child to participate and tell their part of the narrative during a given session. To provide further support, the examiner gave verbal encouragement, if necessary. The procedure encouraged group interaction and elaboration, and at the same time allowed individual participation of each child.

The intervention took place in a quiet area of the clinic. The project was conducted in 2022, and parts of the intervention overlapped with periods of isolation/lockdown due to COVID-19. During such periods, the experimenter completed the missing intervention sessions individually immediately following isolation (5 individuals in Hebrew and 9 individuals in Russian). Between PM2 and PM3 there was a national lockdown, and children had no exposure to SL/Hebrew at preschool.

Fidelity of intervention

Experimenters who administered the intervention were trained by two of the authors. A step-by-step protocol for each session was created for each language and piloted prior to conducting the study (Lipner et al., 2019). Each experimenter was given identical folders with printed materials for each intervention session. The experimenter first checked that all materials were present prior to an intervention session. After each intervention session, the experimenter made sure everything on the fidelity checklist was covered. The checklist included i) a list of all tasks needed to be performed and their order; and ii) a list of responses to be elicited from each child.

Coding

Narratives from all four PM sessions in both languages were transcribed and coded by independent research assistants trained to transcribe using CHAT conventions (MacWhinney, 2000) and who were ignorant of participants' clinical status. The division into utterances was based on C-units, where each C-unit consisted of one main clause or a main clause with a single subordinate clause (Hunt, 1970; Loban, 1976). Each transcribed narrative was then coded for macrostructure and microstructure elements.

Coding for **macrostructure** included examining each utterance for whether the child produced one of the SG elements: character, problem, feeling, goal, attempt, outcome, or internal response. The total macrostructure score was the sum of the responses rated on a three-point scale, as follows: 2 points for accurate production of an SG element, 1 point for a general description of the element, and no points if the element was not produced at all. This yielded a maximum score of 14 for each child for each PM story. A second coding system was used for one of the analyses, where the production of each element was assigned a binary coding, that is, 1 point if the child produced an element, and 0 points if the child did not mention the element.

Microstructure measures were: total number of word tokens (TW), number of different words (NDW), C-unit complexity (percent of complex C-Units), and C-unit accuracy (percent of errorless C-units). A complex utterance was defined as an utterance including two clauses connected via coordination or subordination.

Data analyses

All statistical analyses were conducted in R, version 4.0.5 (R Core Team, 2017). Descriptive statistics are reported for all measures first. To test which factors contributed to the macrostructure and the microstructure measures, linear mixed models (LMM) analyses were performed. LMM analyses were used because of the repeated design of the study, where each participant was tested in four PMs and in both languages. LMM is recommended for clinical intervention studies, since inclusion of random effects allows for control for individual variation (Wiley & Rapp, 2019). Moreover, there is a consensus that there is high variability in the language abilities of children with DLD, and including the random factor partially addresses this challenge. Analysis of each measure began with the null model, which included only random effects, while each subsequent model tested fixed factors and their interactions. Likelihood ratio tests were used to compare two consecutive models by calculating likelihoods for two models, using maximum likelihood estimation, and then statistically comparing those likelihoods which establish significance of each fixed factor (Brown, 2021). A model including a factor of interest (e.g., Group) was compared with a model lacking that factor (i.e., only with the random factor). All models included a by-Participant random factor.

For **macrostructure**, two analyses were performed. First, to test factors predicting the total macrostructure score, we ran an LMM analysis with the fixed factors Group (BiDLD/BiTLD), Language (HL/SL), PM (1/2/3/4), AoB (a continuous measure), and their interactions. The dependent variable was total macrostructure score with the maximum value of 14. Second, to test the probability

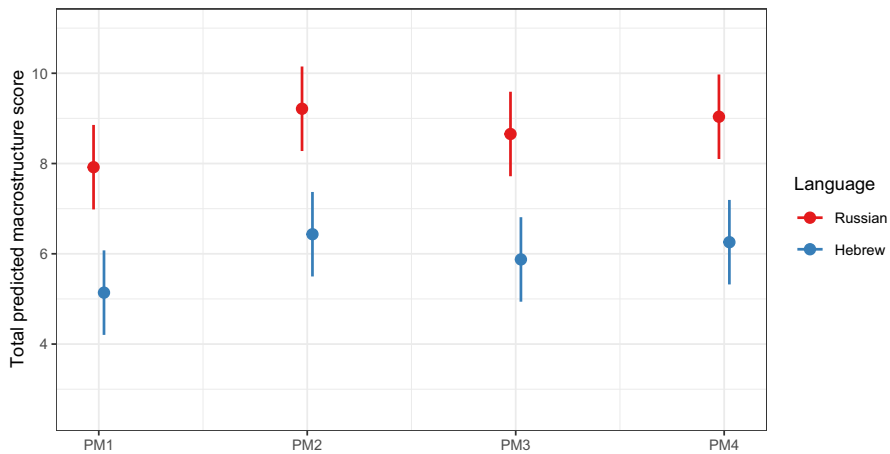


Figure 1. Predicted total macrostructure score across four progress monitorings in HL/Russian and SL/Hebrew.

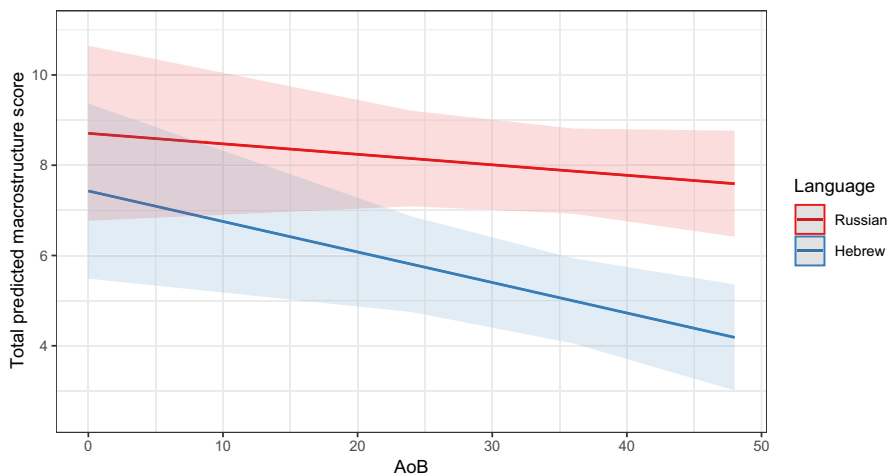


Figure 2. Predicted total macrostructure score for HL/Russian and SL/Hebrew as a function of Age of onset of Bilingualism.

of producing each macrostructure element (character, problem, feeling, goal, action, ending, and internal response), a generalized linear mixed model (GLMM) with binomial distribution was used, where each element had a binary coding (1-produced, 0-not produced).

For **microstructure**, separate LMM models were built for the following measures: TW, NDW, C-units, Complexity (percent of complex C-units), and Accuracy (percent of errorless C-units). In each analysis, the fixed factors were Group (BiDLD/BiTLD), Language (HL/SL), PM (1/2/3/4), AoB (a continuous measure), and their interactions.

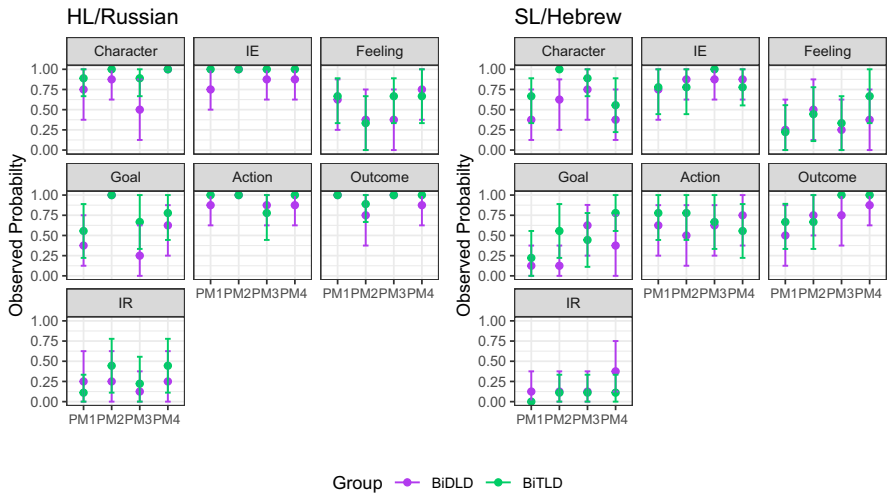


Figure 3. Predicted probabilities of producing each SG element in HL/Russian and SL/Hebrew across four PMs. *Note:* IE = Initiating Event; IR = Internal Response; PM = Progress Monitoring.

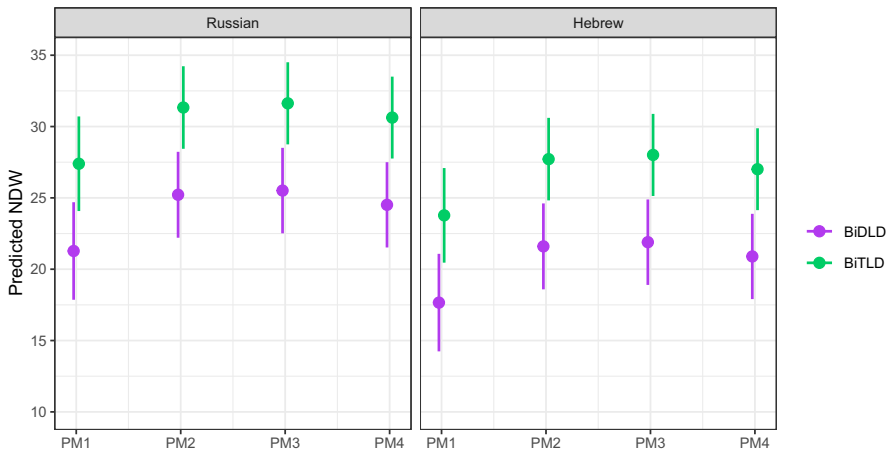


Figure 4. Predicted number of different words scores by Group and Language across four progress monitorings.

In the analyses, a main effect for Group would indicate a difference in performance between children with BiDLD and BiTLD. A main effect for Language would reveal differences between HL and SL, collapsed across Groups for all four PMs. An effect for PM would indicate changes in both languages across the four-time points. Since the first block of intervention was conducted only in HL and the second block only in SL, improvement in a language that was not the target of the intervention would be interpreted as evidence for transfer. The interaction of Language*PM would show that the change is stronger in one language than the other.

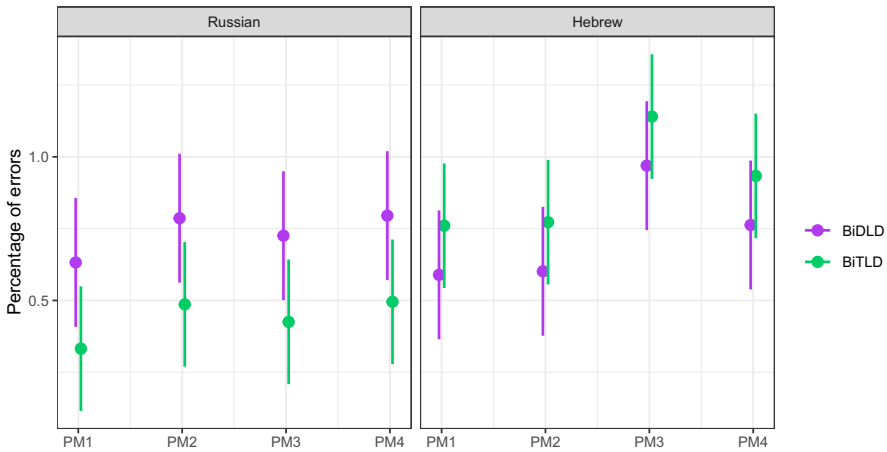


Figure 5. Predicted percentage of errors by Group and Language across four progress monitorings.

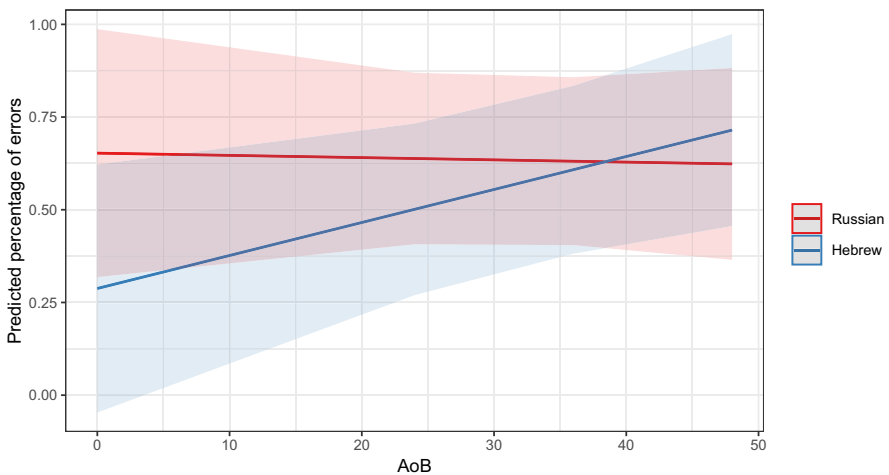


Figure 6. Predicted percentage of errors for HL/Russian and SL/Hebrew as a function of Age of onset of Bilingualism.

LMM analyses were run using the *lmer* function from the *lmerTest* package (Kuznetsova et al., 2017). Likelihood-ratio tests determined the significance of each factor by comparing two models using AIC, BIC, logLik, deviance, Chi-square, and *p*-values. Post-hoc tests were run using contrasts from the *emmeans* package (Lenth, 2019). Predicted values were plotted using the *ggplot* function from the *ggplot2* package (Wickham, 2016) or the *plot_model* function from the *sjPlot* package (Lüdtke, 2021). When plotting predicted values or predicted probabilities with two categorical predictors (Figures 1–4), dots represent the mean predicted scores/probabilities, and error bars represent a range of values around the mean. Figures 5 and 6 display

Table 2. Total macrostructure score (Means and SDs) for each of four PMs for children with bilingual developmental language disorder (BiDLD) and bilingual typical language development (BiTLD)

	HL/Russian		SL/Hebrew	
	BiDLD	BiTLD	BiDLD	BiTLD
PM1	7.67 (2.66)	8.44 (1.66)	4.25 (2.25)	5.77 (3.19)
PM2	8.00 (3.46)	10.00 (2.06)	5.00 (3.16)	6.55 (2.65)
PM3	7.00 (2.20)	9.55 (2.69)	5.62 (2.19)	6.66 (2.00)
PM4	8.00 (1.69)	9.88 (1.96)	5.75 (2.65)	6.55 (3.32)

Note: PM = Progress Monitoring.

predicted values (lines) and the confidence intervals (colored areas) across the continuous variable (AoB) plotted on the X-axis.

Results

Findings for total macrostructure scores for the two groups (BiDLD and BiTLD) in both languages for each of the four PMs are presented first, followed by the performance on each macrostructure element. Next, microstructure measures of productivity (TW, NDW, and number of C-units), complexity, and accuracy for each language are presented. Findings for the impact of AoB and transfer effects are integrated into each analysis.

Total macrostructure score

Table 2 displays the total macrostructure scores for each of the four PMs in the two languages for children with BiDLD and BiTLD.

Table 2 shows that children with BiDLD had lower scores than children with BiTLD at all four PMs and that the performance of all children was better in HL/Russian than in SL/Hebrew. To test whether there were significant differences in the performance of children with BiDLD and children with BiTLD across PMs in HL/Russian and SL/Hebrew, LMM models with the following fixed factors were tested in this order: Group, Language, PM, AoB, and their interactions. The analyses included by-Participant random intercepts. The following factors emerged as significant: Language, $\chi^2 = 52.16$, $p < .001$; PM, $\chi^2 = 8.69$, $p = .03$, and the Language*AoB interaction, $\chi^2 = 7.62$, $p = .02$. Table A1 in the Appendix includes the results of Likelihood Ratio tests. The final optimal model included the fixed factors of Language, PM, Language*AoB, and the random factor. The variance explained by the fixed factors was 35%, and another 20% was explained by the random factor. For the Language factor, HL/Russian had higher macrostructure scores than SL/Hebrew. For the difference between PMs, reverse Helmert coding was used for pairwise comparisons, since this coding allows comparison of each PM level with the mean of the previous level(s). The analysis revealed that the mean for PM2 was higher and was statistically significant compared to PM1 ($p = .03$),

but PM3 did not differ statistically from the mean for either PM1 and PM2 ($p = .87$), and PM4 did not differ significantly from the mean of all previous PMs ($p = .37$). Cohen's d was calculated to evaluate the effect size of the change; it was 0.5 indicating a medium effect size. This was found for both languages. Figure 1 plots the predicted total macrostructure score for the four PMs in HL/Russian and SL/Hebrew.

As shown in Figure 1, children's scores in both groups were higher in PM2 than in PM1, and scores were higher in HL/Russian than in SL/Hebrew across all PMs. In addition, the scores for HL/Russian (red) were higher than the scores for SL/Hebrew (blue) across all four PMs.

To interpret the Language by AoB interaction, the predicted macrostructure scores for Russian and Hebrew were plotted as a function of AoB, collapsed across the four PMs (Figure 5).

As seen in Figure 5, the total predicted macrostructure score decreased sharply for SL/Hebrew as a function of AoB, but not for HL/Russian, where the decrease was negligible. In other words, later AoB was associated with lower predicted macrostructure score in SL/Hebrew, whereas but not related to predicted macrostructure score in HL/Russian.

Story grammar elements

To test the probabilities of producing each SG element, a GLMM analysis with binomial distribution was conducted. The analyses included random by-Participant intercepts and random slopes by macrostructure element. All fixed factors in the order tested and the results of likelihood ratio tests are presented in Table A2 in the Appendix.

The following fixed factors were significant: Group, $\chi^2 = 5.47$, $p = .02$, Language, $\chi^2 = 54.34$, $p < .001$, PM, $\chi^2 = 15.96$, $p = .001$, and SG Element, $\chi^2 = 35.73$, $p < .001$. For Group, there were lower probabilities for children with BiDLD to produce elements. The Language and the PM effects confirmed the finding reported for the total macrostructure score. For the effect of Element, Feelings, Goals, and IR were the two elements with the lowest probability to be produced ($p < .001$). The final optimal model explained 42% of variance by the fixed factors and 16% by the random factors. Like the results reported above for the total macrostructure score, HL/Russian performance was significantly better than SL/Hebrew at all four PMs, and the BiTLD group performed better than the BiDLD group in both languages.

To explore the effect of SG elements across languages for the two groups, we plotted the probabilities for using each element (Figure 2).

Macrostructure analysis showed BiDLD-BiTLD similarity for the total score. However, group differences emerged for individual SG elements. Performance was better in HL/Russian than in SL/Hebrew for both groups. Children in both the BiDLD and BiTLD groups significantly improved in both languages following intervention in HL/Russian (at PM2) but with no subsequent increase in scores at PM3 and PM4. In addition, there was an increase in macrostructure elements in SL/Hebrew, in particular for Feeling and Goal, which were very low initially and showed better performance across the four PMs. An AoB effect emerged in

Table 3. Descriptive statistics (means and standard deviations) for the productivity measures in HL/Russian and SL/Hebrew for the two groups (bilingual developmental language disorder/bilingual typical language development [BiDLD/BiTLD]) across four PMs

		HL/Russian		SL/Hebrew	
		BiDLD	BiTLD	BiDLD	BiTLD
TW	PM1	37.75 (15.60)	42.00 (18.28)	30.75 (18.57)	48.22 (21.07)
	PM2	37.88 (10.30)	49.00 (7.18)	37.88 (14.55)	44.00 (12.75)
	PM3	35.63 (13.68)	53.78 (13.02)	42.38 (7.78)	44.56 (13.58)
	PM4	37.88 (11.87)	47.89 (10.55)	34.50 (10.83)	47.44 (23.43)
NDW	PM1	21.63 (7.07)	26.67 (9.26)	18.38 (8.55)	23.56 (5.79)
	PM2	25.75 (8.14)	31.89 (4.99)	23.38 (7.78)	25.11 (7.67)
	PM3	23.25 (6.63)	33.89 (6.95)	24.00 (5.66)	25.89 (5.99)
	PM4	24.12 (5.33)	30.11 (4.94)	21.63 (6.72)	27.22 (7.10)
C-units	PM1	9.38 (4.24)	9.22 (3.63)	6.63 (3.89)	8.78 (2.82)
	PM2	7.38 (1.85)	8.89 (1.17)	7.75 (3.41)	8.78 (1.20)
	PM3	7.63 (2.00)	9.33 (3.50)	9.50 (2.51)	8.11 (2.15)
	PM4	8.63 (2.20)	10.00 (2.60)	7.13 (2.03)	9.22 (4.12)

SL/Hebrew in both groups, where the younger the onset of bilingualism, the higher the predicted total macrostructure score in Hebrew.

Microstructure

Narrative microstructure analysis included five measures, three for productivity (Total Words, Number of Different Words, and number of C-units), one measure of Complexity (percent of complex C-units), and one measure of Accuracy (percent of accurate C-units).

Productivity

Table 3 summarizes the descriptive statistics (means and standard deviations) for the productivity measures in HL/Russian and SL/Hebrew for the two groups (BiDLD/BiTLD) across four PMs.

To test which factors contributed to productivity measures, LMM analyses were applied to each measure with the following fixed factors: Group, Language, PM, and AoB, as well as interactions among these factors. The analyses included by-Participant random intercepts and by-Story random slopes. For all three measures, the models with random slopes failed to converge, and only the random intercepts were tested. Table A3 in the Appendix shows all the fixed factors, in the order they were tested and the results of the likelihood ratio tests.

For TW, Group, $\chi^2 = 6.26$, $p = .01$, and AoB, $\chi^2 = 4.33$, $p = .04$, were significant. The variance explained by the fixed factors in the final model was 18%

Table 4. Means and SDs for Complexity (ratio of complex C-units) and Accuracy (ratio of C-units containing errors)

		DLD		TLD	
		HL/Russian	SL/Hebrew	HL/Russian	SL/Hebrew
Complexity	PM1	0.38 (0.74)	0.63 (0.52)	1.11 (1.27)	2.11 (1.87)
	PM2	1.13 (1.25)	1.13 (0.99)	1.67 (0.50)	2.22 (1.64)
	PM3	0.38 (0.52)	1.75 (1.17)	2.00 (1.87)	2.78 (1.56)
	PM4	1.13 (1.36)	1.25 (1.28)	1.78 (1.09)	1.56 (1.13)
Accuracy	PM1	6.13 (3.48)	4.00 (2.14)	2.56 (2.19)	6.44 (4.64)
	PM2	5.87 (4.29)	4.88 (3.52)	4.56 (2.22)	6.67 (4.39)
	PM3	5.25 (1.58)	7.00 (2.78)	3.78 (2.05)	10.44 (6.37)
	PM4	7.00 (3.21)	5.38 (1.30)	4.89 (3.18)	6.78 (3.03)

Note: PM = Progress Monitoring.

with another 10% explained by the random factor. For Group, children with BiDLD produced shorter narratives than children with BiTLD at all PMs. Later AoB was associated with lower TW. This finding held for both languages, since the Language*AoB interaction was not significant.

For NDW, Group, $\chi^2 = 5.85$, $p = .02$, Language, $\chi^2 = 14.70$, $p < .001$, PM, $\chi^2 = 10.64$, $p = .01$, AoB, $\chi^2 = 6.39$, $p = .01$, and the Language*AoB interaction, $\chi^2 = 3.76$, $p = .05$, were significant. The final model including these factors explained 54% of the variance, 35% by the fixed factors, and 19% by the random factor. The Language effect showed that greater lexical diversity (NDW) was found in the HL/Russian narratives than in SL/Hebrew narratives. For the difference across PMs in terms of NDW, we used a reversed Helmert coding which revealed that the mean of PM2 was higher and differed statistically from PM1 ($p = .01$); the difference between PM3 and the mean of PM1 and PM2 missed significance ($p = .06$); PM4 did not differ significantly from the mean of all previous PMs ($p = .65$). Cohen's d for the PM1-PM2 difference was 1.04, which represents a large effect size. This was predicted for both languages. For Group, the narratives of children with BiTLD were more lexically diverse than those of children with BiDLD. AoB was also a significant predictor of lexical diversity. Figure 3 graphically displays the differences across the four PMs and the effects of Language and Group on the predicted NDW scores.

Figure 3 shows greater NDW in PM2 compared to PM1 for both BiDLD and BiTLD groups and in both languages (HL/SL). Both groups improved from PM1 to PM2 in both languages, even though the BiDLD group did not perform as well as the BiTLD group, and HL/Russian showed higher predicted NDW values than SL/Hebrew. Improvement from PM1 to PM2 was similar in both languages for the two groups, which is supported by a lack of significant interactions between the fixed factors.

For C-units, none of the fixed factors were significant.

Complexity and accuracy

Table 4 displays the Means and SDs for Complexity (ratio of complex C-units) and Accuracy (ratio of C-units containing errors).

To test the effects of the fixed factors on Complexity and Accuracy, LMM analyses were performed. Table A4 in the Appendix presents the fixed factors in the order they were entered into the models. For Complexity, none of the fixed factors was significant. For Accuracy, Language, $\chi^2 = 12.01$, $p = .001$, Group*Language, $\chi^2 = 14.88$, $p = .001$, Language*PM, $\chi^2 = 14.07$, $p = .03$, and Language*AoB, $\chi^2 = 6.82$, $p = .03$ were significant. The best-fit model for Accuracy included these factors, which explained 27% of the variance, and the random factors explained an additional 14% of variance. Children performed better (fewer errors) in HL/Russian than in SL/Hebrew; however, a significant Group*Language interaction showed that the gap between children with BiDLD and their peers with BiTLD was wider in Russian than in Hebrew, as seen in Figure 4.

Figure 4 shows that children with BiDLD (purple) had more errors than children with BiTLD (green). A post-hoc analysis with Tukey corrections revealed that in HL/Russian, children with BiDLD had a significantly higher percent of errors than children with BiTLD ($p = .03$), but the group difference was not significant in SL/Hebrew ($p = .20$). A post-hoc analysis for the Language*PM interaction, applying Tukey corrections, revealed that in Hebrew, children in both groups produced more errors in PM3 compared to PM1 ($p = .01$) and more errors compared to PM2 ($p = .01$), also visible in Figure 4; in Russian, there were no significant differences across PMs.

Figure 6 displays the interaction of Language*AoB.

Like the results for NDW, the curve reflecting SL/Hebrew (blue) shows that the predicted percentage of errors in C-units rises as a function of AoB. In other words, the later the AoB, the higher the percentage of errors in Hebrew. In contrast, in HL/Russian, later AoB is not associated with a higher percentage of C-units with errors, as reflected by the relatively flat curve across AoB time intervals.

Discussion

The primary interest of this study was to examine changes in macrostructure and microstructure skills associated with Bilingual NARrative Intervention (BINARI) at four time points in BiDLD and BiTLD Russian-Hebrew preschool children's narrative performance in both their languages. For **macrostructure**, the total score as well as specific macrostructure elements (especially Feeling and Goal) showed improvement for both groups after the first block of intervention in HL/Russian, but remained stable, and even decreased somewhat after the second block of intervention in SL/Hebrew. All children performed better in HL than in SL across all four time points. Earlier AoB was found to be related to better macrostructure performance only in SL/Hebrew and not in HL/Russian. Analyses conducted at the level of individual SG elements showed that children with BiDLD performed significantly lower than children with BiTLD. Probability plots for the optimal model revealed that group differences were prominent for most elements, particularly for Character and Goal (Figure 2). For **microstructure**, children with

BiTLD produced longer narratives (more TW) and had higher lexical diversity (greater NDW) than children with BiDLD. Of the five microstructure measures examined (TW, NDW, C-units, Complexity, and Accuracy), only lexical diversity (NDW) showed significant improvement in both languages following the intervention in HL/Russian. For TW, the effect of AoB (but not the Language*AoB interaction) was significant. Children produced higher lexical diversity (NDW) in HL than in SL. For Accuracy, children with BiDLD had a significantly higher percentage of errors in HL/Russian than children with BiTLD, but the difference was not significant in SL/Hebrew. In SL/Hebrew, children had more errors in PM3 compared to PM1 and PM2.

Bilingual narrative performance reflects an interplay of child-internal factors, such as relative proficiency in the two languages, clinical status of the child (BiDLD or BiTLD), and child-external factors, such as the order of intervention (HL first/SL first) and exposure. This multitude of factors creates a challenge in understanding the mechanism of narrative development in HL and SL. The present research addressed some of these factors, in particular the child's clinical status (BiDLD vs. BiTLD), changes in narrative within and across languages, and AoB, offering preliminary answers related to the framework of intervention in both languages among bilingual preschool children.

Clinical status and development of narrative skills

Macrostructure

Findings showed that whereas children with BiDLD showed lower narrative performance, their rate of improvement was similar to children with BiTLD, as evidenced by the increase in the total macrostructure score between PM1 and PM2. Bilingual intervention, using a variety of procedures involving icons/gestures, multiple repetitions, peer interaction, and multi-modal processing, created the environment for narrative skills to develop. The Group differences here support previous studies, which have shown lower macrostructure scores for children with BiDLD than for children with BiTLD (e.g., Boerma et al., 2016 for Dutch bilingual children; Paradis et al., 2013 for varied HLs with L2/English). The two groups do, however, show similarity in the trajectory of their improvement, as evidenced by the fact that there was no interaction between Group and PM.

The bilingual children in the present study were recruited from among children who were clinically referred, largely based on their weak performance in the SL/Hebrew. Evidence-based practices call for evaluation in both languages of bilinguals (ASHA, 2019). Following screening and division into two groups, children with BiDLD were found to perform worse than those with BiTLD, but despite this difference, they showed a parallel trajectory across the four PM time points. Examining a range of macrostructure and microstructure abilities was found to be important in order to go beyond previous research which dichotomized BiDLD-BiTLD differences as similar in macrostructure but different in microstructure. The present study documents group differences at the level of SG elements and lexical diversity (NDW). Bilingual assessment and intervention enabled us to show that children with BiDLD are expected to progress along the same trajectory as

children with BiTLD as well as to identify particular features of macrostructure and microstructure that can be targeted in future interventions.

Macrostructure is grounded in the use of specific elements, all of which contribute to story's coherence. Examining narrative performance at the level of SG elements was crucial to have a more detailed understanding of children's performance. For the total macrostructure score, the LMM analysis explained 35% of the variance by the fixed factors and 20% by the random factor (55% overall). We then conducted a GLMM analysis with the production of each SG element as the response variable, and it explained 42% of the variance by the fixed factors and 16% by the random factor (58% overall). In the LMM, the fixed factors were Language, PM, and Language*AoB, and in the GLMM, they were Group, Language, PM, and Element. Thus, the GLMM analysis was able to explain a greater amount of variance by the predictors than the LMM. This is reflected in the higher variance explained by the fixed factors in the GLMM. It is possible that some of the variance attributed to individual variation in the LMM (i.e., the random factor) is explained in the GLMM by the fixed factor. This is seen in the lower variance explained by the random factor in GLMM. The need to consider individual SG elements was also evident in Figure 2 displaying predicted probabilities of each element by Group and Language.

Microstructure

The narratives of children with BiDLD contained fewer words, were less lexically diverse, and contained fewer complex sentences than the narratives of the BiTLD group. Moreover, a significant Group by Language interaction emerged for morphosyntactic accuracy, where the gap between children with BiDLD and their peers with BiTLD was wider in Russian than in Hebrew. Children with BiDLD had more morphosyntactic errors than children with BiTLD in HL/Russian, but the difference between the groups was not significant in SL/Hebrew. This finding is not surprising, since there is a general consensus that children with DLD have weaker microstructure abilities (Altman *et al.*, 2016; Iluz-Cohen & Walters, 2012; Rezzonico *et al.*, 2015; Squires *et al.*, 2014; Tsimpli *et al.*, 2016). With respect to accuracy, the group difference was due to a lower percentage of errors in the HL/Russian for children with BiDLD, resulting in higher HL proficiency. Both groups were clinically referred because of their weak performance in SL/Hebrew, and only the BiDLD group showed low proficiency scores in both languages, which shows that at the age of 5–6, performance in HL is critical for distinguishing typical development and DLD.

Two results are in need of explanation. First, the total macrostructure score did not improve after the second block of intervention in SL/Hebrew. One possible explanation for this finding is that the children may have needed more intervention sessions in SL/Hebrew (more than in HL/Russian) in order to demonstrate improvement. Since the children were familiar with storytelling in HL/Russian from home, six intervention sessions may have been sufficient to show improvement, given their stronger proficiency in that language. In SL/Hebrew, however, more sessions may have been needed due to lower vocabulary levels and less exposure (Armon-Lotem *et al.*, 2021). The second finding requiring explanation was that in

SL/Hebrew, children had more errors in PM3 than in the previous PMs. During the intervention, in particular between PM2 and PM3, the COVID-19 pandemic led to a national lockdown leaving children at home and not exposed to Hebrew in preschool. Since the children came from HL/Russian dominant homes, the only exposure to SL/Hebrew they received during this period was from the current research involving narrative intervention. Bao et al. (2020) reported that kindergarten children lost 67% of their literacy abilities during lockdown. Thus, the increase in errors may be a result of the lockdown. Finally, the unexpected result for TW, where the effect of AoB (but not the Language**AoB* interaction) was significant, indicated that later exposure **to both languages** was related to lower TW.

Enhanced performance for both macrostructure and microstructure raises a question about the connection between macrostructure and microstructure abilities. Macrostructure elements provide scaffolding for story coherence, while lexis and morphosyntax are essential to produce a well-structured story. The question becomes: How can weak microstructure abilities found among children with DLD be bootstrapped onto relatively intact macrostructure? Two features of the intervention implemented in the present study facilitated this process: repetition and variation. Each child produced 5–6 repetitions of a story in each session. In addition, the intervention involved elaborating content about characters and their feelings, which were modeled by the experimenter as well as other children in the group. For description and elaboration of story elements, the need to diversify vocabulary arose subsequently, pointing to an interplay between macrostructure story elements and vocabulary.

Narrative development across languages

The findings documenting cross-language transfer for both BiDLD and BiTLD children contrast with those reported by Petersen et al. (2016) who found cross-linguistic transfer only for children with TLD. The Petersen et al. (2016) study delivered intervention only in SL/English and reported SL-to-HL transfer. The current findings are more in line with the few studies claiming that when intervention begins with the HL, it creates optimal conditions for transfer, especially for children with BiDLD (Lugo-Neris et al., 2015; Thordardottir, 2010). Transfer of macrostructure skills after the first block of intervention in HL can be explained by evidence that story structure is shared across languages from similar cultural backgrounds (Berman & Slobin, 1994; Squires et al., 2014), and both Russian and Hebrew storytelling has been shown to follow similar patterns (Altman et al., 2016).

Improvement following intervention resulted in lexical diversity (NDW) for each of the four monitoring points, showing transfer from HL to SL. In other words, following intervention in HL/Russian, we find higher lexical diversity in SL/Hebrew. The intervention implicitly trained for lexical diversity, teaching children to talk about narrative elements they had less experience with, e.g., characters' feelings, goals, and internal responses. For this, they needed to use new and different words. Theoretically, the high degree of imitation/repetition and variation engendered by the intervention procedures is grounded in Walters (2005) model of bilingual production, which draws widely from psychology (e.g. William James' 1890

‘consistent ends by variable means’; Clark’s 1987 ‘principle of contrast’) and linguistics (e.g., Tannen’s 1989 five functions of repetition in conversation: fluency, to facilitate comprehension, as connective links), some of which equally apply to narrative.

Establishing a cross-linguistic transfer effect is a challenge because it is difficult to disentangle the effects of time and maturation and is especially challenging due to the absence of a control condition. However, the modest results in the current study regarding transfer may have important educational implications, since in the present study following intervention in HL there was improvement in both languages (Restrepo *et al.*, 2013). Treatment is usually conducted in the SL—because of the SLPs’ training and because of the lack of an explicitly bilingual policy in treating BiDLD (Abutbul-Oz & Armon-Lotem, 2022). The current intervention procedure offers a uniquely bilingual perspective on the assessment and treatment of children with BiDLD by inviting educators and SLPs to integrate both languages. The optimal balance and order of the two languages are yet to be researched; the current results suggest that beginning the intervention with the HL may lead to positive changes in SL.

Age of onset of Bilingualism and narrative development

The analyses conducted in the current study found three significant interactions involving Language and AoB (for the total macrostructure score, for lexical diversity (NDW), and for morphosyntactic errors (Accuracy)). For all three interactions, AoB showed differences in SL and not in HL. Specifically, the effect was negative for the total macrostructure score and NDW (later AoB was related to lower scores on these measures in SL) and was positive for Accuracy (later AoB was associated with fewer errors). Thus, in all three analyses, later AoB was associated with enhanced performance in macrostructure and microstructure skills. These findings for the impact of AoB on the SL conform to the adage in other areas of language acquisition that ‘earlier is better.’ One finding, however, contrasts with the described trends; a significant and positive relationship was found between AoB and the total number of words in both languages. This indicates that later AoB was associated with fewer total words in HL and SL. One possible explanation is that the productivity measure of TW does not necessarily reflect the proficiency, and NDW is a better measure of narrative microstructure.

The other part of the interaction documented here, i.e. that AoB had almost no effect on the HL/home language, implies that narrative macrostructure skills as well as lexical diversity and morphosyntactic accuracy in HL are acquired relatively early, remain stable, and do not interact with (lack of) exposure to the SL. This result challenges studies showing mixed findings for the effects of AoB on HL vocabulary, morphosyntax, and syntax. Those studies focusing on morphosyntax generally show positive effects of AoB (Albirini, 2018; Meir *et al.*, 2017; Montrul, 2008, 2016; Soto-Corominas, 2021), whereas studies of SL syntax give a mixed picture: no effects of AoB or negative effects (Roesch & Chondrogianni, 2016), depending on the syntactic structure examined and the population tested (Chiat *et al.*, 2013; Meir *et al.*, 2017; Kaltsa *et al.*, 2020). Thus, at least for Russian-Hebrew bilingual children

in Israel, early AoB (exposure to the SL) results in better macrostructure and microstructure performance without negatively interacting with HL abilities.

Limitations and future research

The design of the bilingual intervention in the present study included a block of HL intervention sessions followed by a block of SL sessions. This design was adopted since previous studies showed benefits of beginning in the child's HL. In order to evaluate whether this design is indeed beneficial for bilingual children, future research should compare two intervention orders, HL first-SL second vs. SL first-HL second, and employ a multiple-baseline design. A more focused *suggestion based on the findings for SG elements would be to construct intervention sessions targeting Feelings and Goals*, found to be those most vulnerable in the stories of children with BiDLD.

One obvious limitation of the present study was that it could not use a control group, as explained above. For that reason, we employed a within-subject design and focused on changes in narrative performance across four points prior to, during, and following intervention. The small sample size ($n = 17$) is another limitation, resulting in reduced statistical power.

Finally, the present study used single-episode stories for both intervention and progress monitoring. Short stories are designed to make narration easier for young children and for those with DLD, since they constrain the number of opportunities to convey macrostructure elements (each story contains seven possible elements). Future studies should use longer stories in order to elicit greater variability.

Conclusions and implications

The results of this study suggest that intervention in both languages can be effective for children with BiDLD as well as those with BiTLD. Children with BiDLD perform poorer than children defined as BiTLD, but their improvement from the intervention followed the same trajectory as children without impairment. In addition to improvement for both groups, the bilingual intervention showed benefits in both languages as well as transfer of narrative skills from HL to SL. Speculating about what else may have contributed to the beneficial impact of the intervention, the social scaffolding that the intervention procedure offered is one candidate. Children participated in small groups, with a great deal of interaction among peers as well as with the adult research assistant. This procedure created a socio-pragmatic context involving turn-taking and becoming part of a supportive peer group. Furthermore, the repetitive use of icons and gestures for the targeted elements that accompanied the story facilitated learning across multiple modalities, viz. auditory, visual, and kinesthetic. These multisensory channels may be appropriate for children with BiDLD as they have been shown to be for children with dyslexia and for adolescents at risk for foreign language learning.

Funding statement. This research was funded by an ISF grant 1716/19; PIs: Altman & Walters.

References

- Abutbul-Oz, H., & Armon-Lotem, S. (2022). Parent questionnaires in screening for developmental language disorder among bilingual children in speech and language clinics. *Frontiers in Education*, 7, 323.
- Albirini, A. (2018). The role of age of exposure to English in the development of Arabic as a heritage language in the United States. *Language Acquisition*, 25(2), 178–196.
- 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.
- Altman, C., Fichman, S., Perry, N., Osher, P., & Walters, J. (2024). Internal state terms in the narratives of bilingual children with developmental language disorder: the role of microstructure and macrostructure. *Language, Speech, and Hearing Services in Schools*, 55(4), 1039–1053.
- 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*, 36(12), 1132–1152.
- American Speech-Language-Hearing Association (2010). Bilingual Service Delivery (Practice Portal). Available from www.asha.org/Practice-Portal/Professional-Issues/Bilingual-Service-Delivery/.
- American Speech-Language-Hearing Association. (2019). Bilingual service delivery. <http://www.asha.org/Practice-Portal/Professional-Issues/Bilingual-Service-Delivery/>.
- Armon-Lotem, S. (2014). Between L2 and SLI: Inflections and prepositions in the Hebrew of bilingual children with TLD and monolingual children with SLI. *Journal of Child Language*, 41(1), 3–33.
- 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.
- Armon-Lotem, S., Restrepo, M. A., Lipner, M., Ahituv-Shlomo, P., & Altman, C. (2020). Vocabulary gains in bilingual narrative intervention. *Language, Speech, and Hearing Services in Schools*, 52(1), 436–448.
- Armon-Lotem, S., Restrepo, M. A., Lipner, M., Ahituv-Shlomo, P., & Altman, C. (2021). Vocabulary gains in bilingual narrative intervention. *Language, speech, and hearing services in schools*, 52(1), 436–448.
- Armon-Lotem, S., Rose, K., & Altman, C. (2021). The development of English as a heritage language: the role of chronological age and age of onset of bilingualism. *First Language*, 41(1), 67–89.
- Bao, X., Qu, H., Zhang, R., & Hogan, T. P. (2020). Modeling reading ability gain in kindergarten children during COVID-19 school closures. *International Journal of Environmental Research and Public Health*, 17(17), 6371.
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bishop, D. V. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International journal of language & communication disorders*, 52(6), 671–680.
- Boerma, T., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2016). Narrative abilities of monolingual and bilingual children with and without language impairment: Implications for clinical practice. *International Journal of Language & Communication Disorders*, 51(6), 626–638.
- Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, 4(1), 1–19.
- Capin, P., Vaughn, S., Gillam, S. L., Fall, A. M., Roberts, G., Israelsen-Augenstein, M., ... & Gillam, R. B. (2023). Evaluating the efficacy of a narrative language intervention for bilingual students. *American Journal of Speech-Language Pathology*, 32(6), 2999–3020.
- Chiat, S., Armon-Lotem, S., Marinis, T., Polišenská, K., Roy, P., Seeff-Gabriel, B., & Gathercole, V. M. (2013). Assessment of language abilities in sequential bilingual children: The potential of sentence imitation tasks. In V. C. M. Gathercole (Ed.), *Issues in the assessment of bilinguals* (pp. 56–89). Bristol: Multilingual Matters.
- Clark, E.V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Mahwah, NJ: Lawrence Erlbaum Associates.

- Cleave, P. L., Girolametto, L. E., Chen, X., & Johnson, C. J. (2010). Narrative abilities in monolingual and dual language learning children with specific language impairment. *Journal of communication disorders*, 43(6), 511–522.
- Daelman, J., Alighieri, C., Van Lierde, K., Simon, E., Altinkamis, F., Baudonck, N., & D'haeseleer, E. (2023). Effect of a group-based narrative language training in typically developing Turkish-Dutch children. *Folia Phoniatrica et Logopaedica*, 76, 192–205.
- Drysdale, H., van der Meer, L. & Kagohara, D. (2015). Children with autism spectrum disorder from bilingual families: A systematic review. *Review Journal of Autism and Developmental Disorders*, 2, 26–38.
- Duinmeijer, I., de Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language & Communication Disorders*, 47(5), 542–555.
- Ebert, K. D., Kohnert, K., Pham, G., Disher, J. R., & Payesteh, B. (2014). Three treatments for Bilingual children with primary language impairment: examining cross-linguistic and cross-domain effects. *Journal of Speech, Language, and Hearing Research*, 57(1), 172–186.
- Fey M. E., Catts H. W., Proctor-Williams K., Tomblin B. J., & Zhang X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research*, 47, 1301–1318.
- Fichman, S., & Altman, C. (2023). Disfluencies in the narratives of Russian-Hebrew bilingual children with and without Developmental Language Disorder (DLD). *Clinical Linguistics & Phonetics*, 38, 1–25.
- 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, 72–93.
- Gagarina, N. V., Klassert, A., & Topaj, N. (2010). Russian language proficiency test for multilingual children. *ZAS Papers in Linguistics*, 54, 54–54.
- Gagarina, N. V., Klassert, A., & Topaj, N. (2019). Russian language proficiency test for multilingual children. *ZAS Papers in Linguistics*, 54, 54–54.
- Gillam, S. L., Olszewski, A., Squires, K., Wolfe, K., Slocum, T., & Gillam, R. B. (2018). Improving narrative production in children with language disorders: An early-stage efficacy study of a narrative intervention program. *Language, Speech, and Hearing Services in Schools*, 49(2), 197–212.
- Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics*, 29(1), 41–65.
- Goralnik, E. (1995). *Goralnik screening test for Hebrew*. Even Yehuda: Matan.
- Govindarajan, K., & Paradis, J. (2019). Narrative abilities of bilingual children with and without developmental language disorder (SLI): Differentiation and the role of age and input factors. *Journal of Communication Disorders*, 77, 1–16.
- Green, L. B., & Klecan-Aker, J. S. (2012). Teaching story grammar components to increase oral narrative ability: A group intervention study. *Child Language Teaching and Therapy*, 28(3), 263–276.
- Gutierrez-Ciellen, V., Simon-Cerejido, G., & Restrepo, M. A. (2014). *Improving the vocabulary and oral language skills of bilingual Latino preschoolers: An intervention for speech-language pathologists*. Plural Publishing.
- Gutierrez-Ciellen, V., Simon-Cerejido, G., & Sweet, M. (2012). Predictors of second language acquisition in Latino children with specific language impairment. *American Journal of Speech-Language Pathology*, 21(1), 64–77.
- Hao, Y., Sheng, L., Zhang, Y., Jiang, F., De Villiers, J., Lee, W., & Liu, X. L. (2018). A narrative evaluation of Mandarin-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*, 61(2), 345–359.
- Harvey, H., Allaway, H., & Jones, S. (2021). The effectiveness of therapies for dual language children with developmental language disorder: A systematic review of interventional studies. *International Journal of Bilingual Education and Bilingualism*, 24(7), 1043–1064.
- Hayward, D., & Schneider, P. (2000). Effectiveness of teaching story grammar knowledge to pre-school children with language impairment. An exploratory study. *Child Language Teaching and Therapy*, 16(3), 255–284.
- Hipfner-Boucher, K., Lam, K., & Chen, X. (2014). The effects of bilingual education on the English language and literacy outcomes of Chinese-speaking children. *Written Language & Literacy*, 17(1), 116–138.

- Hunt, K. (1970). Syntactic maturity in school children and adults. *Monographs of the Society for Research in Child Development*, 53, 134.
- 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.
- Isurin, L. (2005). Cross linguistic transfer in word order: Evidence from HL forgetting and SL acquisition. In *Proceedings of the 4th International Symposium on Bilingualism* (pp. 1115–1130). Cascadilla Press.
- James, W. (1890). *The principles of psychology*. New York: Dover.
- Jordaan, H. (2008). Clinical intervention for bilingual children: An international survey. *Folia Phoniatrica et Logopaedica*, 60(2), 97–105.
- Kaltsa, M., Prentza, A., Papadopoulou, D., & Tsimpli, I. M. (2020). Language external and language internal factors in the acquisition of gender: the case of Albanian-Greek and English-Greek bilingual children. *International Journal of Bilingual Education and Bilingualism*, 23(8), 981–1002.
- Kambanaros, M., Michaelides, M., & Grohmann, K. K. (2017). Cross-linguistic transfer effects after phonologically based cognate therapy in a case of multilingual specific language impairment (SLI). *International Journal of Language & Communication Disorders*, 52(3), 270–284.
- 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.
- Kay-Raining Bird, E., Lamond, E., & Holden, J. (2012). Survey of bilingualism in autism spectrum disorders. *International Journal of Language & Communication Disorders*, 47, 52–64.
- 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.
- Kk Nair, V., Clark, G. T., Siyambalapitiya, S., & Reuterskiöld, C. (2023). Language intervention in bilingual children with developmental language disorder: A systematic review. *International Journal of Language & Communication Disorders*, 58(2), 576–600.
- Kupersmitt, J. R., & Armon-Lotem, S. (2019). The linguistic expression of causal relations in picture-based narratives: A comparative study of bilingual and monolingual children with TLD and DLD. *First Language*, 39(3), 319–343.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1–26.
- Leith, R. (2019). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.3.3.
- Leonard, L. B. (2017). *Children with specific language impairment*. MIT Press.
- Lipner, M., Armon-Lotem, S., Restrepo, A. M., & Altman, C. (2019). Effect of bilingual narrative intervention (BINARI) on narrative skills in the home language of bilingual preschool children. *American Speech Hearing Association Conference*.
- Lipner, M., Armon-Lotem, S., Walters, J., & Altman, C. (2021). Crosslinguistic influence (CLI) of lexical breadth and depth in the vocabulary of bilingual kindergarten children - a bilingual intervention study. *Frontiers in Psychology*, 12, 671928.
- Loban, W. (1976). *Language development: Kindergarten through Grade 12*. Urbana, IL: National Council of Teachers of English.
- Lü, C., Pace, A. E., & Ke, S. (2023). Bidirectional transfer of definition skills and expressive vocabulary knowledge in Chinese-English dual language learners. *International Journal of Bilingual Education and Bilingualism*, 26(2), 201–215.
- Lüdecke D. (2021). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.8.10.
- Lugo-Neris, M. J., Bedore, L. M., & Pena, E. D. (2015). Dual language intervention for bilinguals at risk for language impairment. *Seminars in Speech and Language*, 36(2), 133–142.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. transcription format and programs*. Psychology Press.
- Meir, N., Walters, J., & Armon-Lotem, S. (2017). Bi-directional cross-linguistic influence in bilingual Russian-Hebrew speaking children. *Linguistic Approaches to Bilingualism*, 7(5), 514–553.
- Méndez, L. I., Crais, E. R., Castro, D. C., & Kainz, K. (2015). A culturally and linguistically responsive vocabulary approach for young Latino dual language learners. *Journal of Speech, Language, and Hearing Research*, 58(1), 93–106.

- Montrul, S. (2008). Second language acquisition welcomes the heritage language learner: opportunities of a new field. *Second Language Research*, 24(4), 487–506.
- Montrul, S. (2016). Age of onset of bilingualism effects and availability of input in first language attrition. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 141–161). American Psychological Association.
- Paradis, J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. *Journal of Child Language*, 50(4), 793–817.
- Paradis, J., Schneider, P., & Sorenson Duncan, T. (2013). Discriminating children with language impairment among English language learners from diverse first language backgrounds. *Journal of Speech, Language and Hearing Research*, 56, 971–981.
- Paradis, J., Soto-Corominas, A., Daskalaki, E., Chen, X., & Gottardo, A. (2021). Morphosyntactic development in first generation Arabic—English children: The effect of cognitive, age, and input factors over time and across Languages. *Languages*, 6(1), 51.
- Petersen, D. B., & Spencer, T. D. (2016). Using narrative intervention to accelerate canonical story grammar and complex language growth in culturally diverse preschoolers. *Topics in Language Disorders*, 36(1), 6–19.
- Petersen, D. B., Thompson, 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.
- Pico, D. L., Hessling Prah, A., Biel, C. H., Peterson, A. K., Biel, E. J., Woods, C., & Contesse, V. A. (2021). Interventions designed to improve narrative language in school-age children: A systematic review with meta-analyses. *Language, Speech, and Hearing Services in Schools*, 52(4), 1109–1126.
- R Core Team. (2017). *R: A language and environment for statistical computing*. Retrieved from <https://www.R-project.org/>
- Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). “Frog, where are you?” Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language*, 88(2), 229–247.
- Restrepo, M. A., Morgan, G. P., & Thompson, M. S. (2013). The efficacy of a vocabulary intervention for dual-language learners with language impairment. *Journal of Speech, Language, and Hearing Research*, 56(2), 748–765.
- Rezzonico, S., Chen, X., Cleave, P. L., Greenberg, J., Hipfner-Boucher, K., Johnson, C. J., Milburn, T., Pelletier, J., Weitzman, E., & Girolametto, L. (2015). Oral narratives in monolingual and bilingual preschoolers with SLI. *International Journal of Language & Communication Disorders*, 50(6), 830–841.
- Roesch, A. D., & Chondrogianni, V. (2016). “Which mouse kissed the frog?” Effects of age of onset, length of exposure, and knowledge of case marking on the comprehension of wh-questions in German-speaking simultaneous and early sequential bilingual children. *Journal of Child Language*, 43(3), 635–661.
- Scott, C. M., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research*, 43(2), 324–339.
- Soto-Corominas, A. (2021). Morphology and L1 acquisition. In A. Fábregas, V. Acedo-Matellán, G. Armstrong, M.C. Cuervo, & I. Pujol Payet (Eds.), *The Routledge Handbook of Spanish Morphology* (pp. 513–525). London: Routledge.
- Spencer, T., Peterson, D., & Restrepo, M.A. (2017). Puente de Cuentos. A bilingual response intervention program for preschool children. Language Dynamics Group.
- Spencer, T. D., Petersen, D. B., Restrepo, M. A., Thompson, M., & Gutierrez Arvizu, M. N. (2019). The effect of Spanish and English narrative intervention on the language skills of young dual language learners. *Topics in Early Childhood Special Education*, 38(4), 204–219.
- Squires, K. E., Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., Bohman, T. M., & Gillam, R. B. (2014). Story retelling by bilingual children with language impairments and typically developing controls. *International Journal of Language & Communication Disorders*, 49(1), 60–74.
- Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school. In R.O. Freedle (Ed.), *Advances in discourse processes: vol. 2. New directions in discourse processing* (pp. 53–120). Norwood, New Jersey: Ablex.
- Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

- Thordardottir, E.** (2010). Towards evidence-based practice in language intervention for bilingual children. *Journal of Communication Disorders*, **43**(6), 523–537.
- Thordardottir, E., Cloutier, G., Ménard, S., Pelland-Blais, E., & Rvachew, S.** (2015). Monolingual or bilingual intervention for primary language impairment? A randomized control trial. *Journal of Speech, Language, and Hearing Research*, **58**(2), 287–300.
- Tsimpili, I. M., Peristeri, E., & Andreou, M.** (2016). Narrative production in monolingual and bilingual children with specific language impairment. *Applied Psycholinguistics*, **37**(1), 195–216.
- Unsworth, S.** (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, **16**(1), 86–110.
- Walters, J.** (2005). *Bilingualism: The sociopragmatic-psycholinguistic interface*. London: Routledge.
- Wickham H.** (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.
- Wiley, R. W., & Rapp, B.** (2019). Statistical analysis in Small-N designs: using linear mixed-effects modeling for evaluating intervention effectiveness. *Aphasiology*, **33**(1), 1–30.
- Zehler A., Fleischman H., Hopstock P., Stephenson T., Pendzick M., & Sapru S.** (2003). *Policy report: Summary of findings related to LEP and SPED-LEP students*. Arlington, VA: Development Associates.

Appendix

Table A1. Results of likelihood ratio tests predicting total macrostructure score

	<i>AIC</i>	<i>BIC</i>	<i>logLik</i>	<i>deviance</i>	χ^2	<i>Df</i>	<i>p</i>
Group	658.59	670.24	−325.30	650.59	3.54	1	0.06
Language	609.98	621.63	−300.99	601.98	52.16	1	<.001
PM	607.29	627.68	−296.65	593.29	8.69	3	0.03
AoB	605.76	629.06	−294.88	589.76	3.53	1	0.06
Group*Language	607.11	633.33	−294.56	589.11	4.18	2	0.12
Group*PM	610.90	642.94	−294.45	588.90	4.39	4	0.36
Language*PM	611.52	640.65	−295.76	591.52	1.77	3	0.62
Language*AoB	603.67	629.88	−292.83	585.67	7.62	2	0.02

Table A2. Results of likelihood ratio tests predicting probability of producing Story Grammar elements

Fixed factors	<i>AIC</i>	<i>BIC</i>	<i>logLik</i>	<i>deviance</i>	χ^2	<i>df</i>	<i>p</i>
Group	1042.50	1187.80	−491.00	982.00	5.47	1	0.02
Language	989.65	1140.30	−466.83	927.65	54.34	1	<.001
PM	979.70	1144.90	−455.87	911.70	15.96	3	0.001
Element	955.97	1150.30	−437.98	875.97	35.73	6	<.001
Group*Element	965.03	1188.50	−436.52	873.03	2.94	6	0.82
Language*Element	963.30	1186.80	−435.65	871.30	4.67	6	0.59
Group*Language*Element	977.68	1264.30	−429.84	859.68	16.29	19	0.64
Group*PM*Element	996.22	1409.20	−413.11	826.22	49.75	45	0.30

Note: Element = Story Grammar Element.

Table A3. Results of likelihood ratio tests predicting three microstructure measures (TW, NDW, C-units)

Fixed factors	AIC	BIC	logLik	deviance	χ^2	Df	p
TW							
Group	1104.50	1116.20	−548.28	1096.50	6.26	1	0.01
Language	1106.30	1120.60	−548.03	1096.00	0.50	1	0.48
PM	1108.50	1128.90	−547.26	1094.50	2.03	3	0.57
AoB	1102.20	1116.80	−546.11	1092.20	4.33	1	0.04
Group*Language	1105.60	1126.00	−545.82	1090.60	0.58	2	0.75
Group*PM	1112.00	1144.00	−544.97	1090.00	2.27	6	0.89
Language*PM	1113.50	1148.40	−544.74	1089.50	2.74	7	0.91
Language*AoB	1105.60	1126.00	−545.78	1091.60	0.65	2	0.72
NDW							
Group	918.06	955.92	−446.03	892.06	5.85	1	0.02
Language	905.36	946.14	−438.68	877.36	14.70	1	<.001
PM	900.73	950.24	−433.36	866.73	10.64	3	0.01
AoB	896.33	948.76	−430.17	860.33	6.39	1	0.01
Group*Language	894.80	950.15	−428.40	856.80	3.53	1	0.06
Group*PM	901.38	962.55	−429.69	859.38	0.95	3	0.81
Language*PM	901.62	962.79	−429.81	859.62	0.71	3	0.87
Language*AoB	894.57	949.91	−428.29	856.57	3.76	1	0.05
C-units							
Group	673.48	685.13	−332.74	665.48	3.21	1	0.07
Language	675.19	686.84	−333.6	667.19	1.50	1	0.22
PM	679.93	697.4	−333.96	667.93	0.77	3	0.86
AoB	676.02	687.67	−334.01	668.02	0.67	1	0.41
Group*Language	675.96	693.43	−331.98	663.96	4.74	3	0.19
Group*PM	683.21	712.34	−331.61	663.21	5.48	7	0.60
Language*PM	683.68	712.80	−331.84	663.68	5.02	7	0.66
Language*AoB	678.27	695.75	−333.14	666.27	2.42	3	0.49

Note: TW- Total words; NDW- number of different words.

Table A4. Results of likelihood ratio tests predicting complexity and accuracy

	<i>AIC</i>	<i>BIC</i>	<i>logLik</i>	<i>deviance</i>	χ^2	<i>df</i>	<i>p</i>
Complexity							
Group	−225.16	−213.51	116.58	−233.16	2.07	1	0.15
Language	−225.62	−213.97	116.81	−233.62	2.52	1	0.11
PM	−219.33	−201.85	115.66	−231.33	0.23	3	0.97
AoB	223.11	−211.46	115.56	−231.11	0.01	1	0.92
Group*Language	−223.89	−206.41	117.94	−235.89	4.79	3	0.19
Group*PM	−219.64	−190.51	119.82	−239.64	8.54	7	0.29
Language*PM	218.93	−189.81	119.47	−238.93	7.83	7	0.35
Language*AoB	−221.89	−204.41	116.94	−233.89	2.79	3	0.43
Accuracy							
Group	172.84	184.49	−82.42	164.84	0.19	1	0.67
Language	161.01	172.66	−76.51	153.01	12.01	1	0.001
PM	160.16	180.55	−73.08	146.16	6.85	3	0.08
AoB	161.81	176.38	−75.91	151.81	1.19	1	0.27
Group*Language	150.13	167.60	−69.06	138.13	14.88	2	0.001
Group*PM	151.97	186.92	−63.99	127.97	10.16	6	0.12
Language*PM	148.06	183.01	−62.03	124.06	14.07	6	0.03
Language*AoB	145.24	186.02	−58.62	117.24	6.82	2	0.03

Cite this article: Dadon Jorno, A., Fichman, S., Walters, J., & Altman, C. (2025). Bilingual narrative intervention among Russian-Hebrew bilingual preschool children with developmental language disorder. *Applied Psycholinguistics*. <https://doi.org/10.1017/S0142716425000062>