Journal of Child Language (2025), **52,** 1437–1447
doi:10.1017/S0305000924000394

CAMBRIDGE UNIVERSITY PRESS

BRIEF RESEARCH REPORT

Cognitive skills differentially influence narrative macrostructure in bilinguals' L1 and L2

Freideriki Tselekidou[1,2], Elizabeth Stadtmiller[3], Assunta Süss[1], Katrin Lindner[3] and Natalia Gagarina[1]

[1]Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin, Germany; [2]Humboldt-Universität zu Berlin, Berlin, Germany and [3]Ludwig-Maximilians Universität München, Germany
Corresponding author: Natalia Gagarina; gagarina@leibniz-zas.de

(Received 18 November 2022; revised 15 June 2024; accepted 08 July 2024; first published online 27 December 2024)

Abstract
This study explored cognitive effects on narrative macrostructure in both languages of 38 Russian-German bilinguals aged 4;6 to 5;1, while controlling for demographic factors (sex, socioeconomic status) and language proficiency. Macrostructure was operationalised as story structure (SS) and story complexity (SC) using the Multilingual Assessment Instrument for Narratives. Nonverbal cognitive subtasks assessing shifting (Figure Ground), visual memory (Form Completion), and inhibition (Attention Divided) were administered. None of the cognitive skills influenced SS; however, they differentially impacted SC: Figure Ground affected Russian SC, while Form Completion affected German SC. Findings advance our understanding of how cognition affects oral narratives in bilingual preschoolers.

Keywords: bilingual narratives; cognition; Russian; German

Introduction

Narratives are indispensable for everyday communication and later academic success (Wallach, 2008). Narrative production, be it personal or picture-based, is a complex activity requiring not only language-specific skills, such as vocabulary and morphosyntax (Berman & Slobin, 1994), but also cognitive skills, such as the ability to recognise the (pictorial) components and connect them into a meaningful whole, i.e., macrostructure (Gagarina & Bohnacker, 2022).

To monitor pictorial stimuli and to connect the events in a causal-temporal order, a storyteller relies on cognitive skills, such as shifting, memory, and inhibition (Mozeiko et al., 2011). Shifting helps discriminate objects or protagonists from complex backgrounds and generate a complete episode. Memory is needed, since constructing a narrative requires the ability to retain the visually presented stimuli and connect the story elements into a cohesive whole in a temporal sequence (logical order). Finally, inhibition is crucial for constructing a cohesive story, since it requires the ability to hold attention to two or more characters or actions simultaneously, organise them, and inhibit irrelevant information while producing the story.

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

Research on the effects of cognitive skills on narrative production is still inconclusive and deals predominantly with monolingual children and children with developmental language disorders (DLDs). Khan (2013) found that shifting predicted narrative production in three- to six-year-old monolinguals, while Veraksa et al. (2020) reported a significant relationship between macrostructure and visual working memory of pictures and spatial arrangements in typically developing (TD) children aged five to six. Additionally, correlations were found between macrostructure production and attention in both TD and DLD monolinguals aged six to nine (Duinmeijert et al., 2012).

Recently, Oshchepkova et al. (2022) found that narrative macrostructure in six- to seven-year-old Russian-speaking monolinguals and bilinguals is affected by cognitive flexibility (a measure of shifting), working memory, and inhibition with both groups showing similar macrostructure and cognitive flexibility. Similarly, Clark-Whitney and Melzi (2023) described significant contributions of inhibitory control to personal narratives' macrostructure in English-Spanish bilingual preschoolers.

These studies employed diverse approaches to measuring macrostructure which impedes comparability. Moreover, bilingual research is limited to atypical language development, or to comparisons between monolingual and bilingual English-speaking populations. Consequently, further research is needed to understand cognitive effects on narrative macrostructure in both languages of bilinguals and to find out if cognitive mechanisms underlying narratives vary due to linguistic differences. The present study endeavours to address this gap.

Narrative macrostructure

Previous research employed different pictorial stimuli and grammar models to assess macrostructure in picture-based narratives. The seminal study of Berman and Slobin (1994) study using *Frog, Where are you?* (Mayer, 1969) and the research that followed (Fiestas & Peña, 2004; Paradis et al., 2010; Pearson, 2002; Uccelli & Páez, 2007) used the story grammar model (Stein & Glenn, 1979), suggesting that stories consist of single components that are organised around the setting and episodes.

In the last years the multidimensional model of narrative organisation has served as a base for the Multilingual Assessment Instrument for Narratives (LITMUS – MAIN (henceforth, MAIN) Gagarina et al., 2012, Gagarina et al., 2019 as part of Language Impairment Testing in Multilingual Settings battery). This model provides a unified approach to assess narrative organisation (i.e., macrostructure) with picture-based stories. The core constituent of macrostructure in MAIN is an episode including a Goal (G) of a protagonist, an Attempt (A) to reach the goal, an Outcome (O) of the protagonist's activities and, additionally, internal states (ISs) as initiating and resulting events of these activities. A and O constitute the factual components (that can be directly seen in the pictures), whereas G and IS are the inferred ones. Each MAIN story consists of three comparable episodes with the same number of components. Story Structure (SS) is operationalised as the sum of all components of all episodes in a narrative (quantity dimension) including the introductory setting. Story complexity (SC) assesses narrator's ability to combine the core episodic components (quality dimension), with the co-occurrence of G, A, and O forming the highest level of complexity.

For SS, studies using MAIN reported differences across bilinguals' two languages, showing that it is affected by language proficiency: higher language proficiency resulted in higher SS (Lindgren & Bohnacker, 2022; Kapalková et al., 2016; Lindgren et al., 2023, for a

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

comprehensive overview; Tribushinina et al., 2022). In contrast, SC was found to be similar in bilinguals' two languages which might indicate that it is less dependent on language proficiency (Lindgren & Bohnacker, 2022).

Narratives, demographic, and linguistic factors

Bilinguals' narrative skills are affected by various demographic and linguistic factors. For instance, sex can affect narrative production, with girls displaying better understanding of other people's thoughts and feelings than boys, especially at preschool age (Sheldon & Engstrom, 2005). Additionally, socioeconomic status (SES) is a determinant factor in children's language development since it can lead to less language disparities in both the first (L1; Hoff, 2013) and second languages (L2; Armon-Lotem et al., 2011). Finally, vocabulary affects SS in both languages of bilinguals (Uccelli & Páez, 2007). However, other studies (Lindgren & Bohnacker, 2022) found that vocabulary impacts SS in the less developed home language only, which authors attribute to the high proficiency in the societal language.

The present study

The present study explores how cognitive skills affect narrative macrostructure, operationalised as SS and SC, in both languages of Russian-German-bilingual preschoolers aged 4;6 to 5;1. Specifically, we ask how shifting (Figure Ground), memory (Form Completion), and inhibition (Attention Divided) affect macrostructure production when sex, SES, and language proficiency (measured by lexical and morphosyntactic skills) are controlled for. We additionally ask whether cognitive skills affect SS and SC in Russian and German in a similar way.

Based on previous findings (Khan, 2013; Veraksa et al., 2020), we expect the cognitive subtasks to positively affect macrostructure in both languages, after controlling demographic factors and language proficiency. Given that SC is less dependent on language skills (Lindgren & Bohnacker, 2022), we expect stronger effects of cognitive skills on SC than SS in both languages.

Method

Participants

Thirty-eight bilingual Russian-German children (16 boys, 22 girls) with a mean age 4;8 (range = 4;6–5;1, *SD* = 0;2) participated in the study and were tested in both Russian and German. Their mean AoO for L2 German was 25 months (range: 0–48, *SD*=12.87). All children had at least one Russian-speaking parent, always a first-generation immigrant (five fathers and one mother were non-Russian-speaking; home language data were missing from one set of parents and one father in a family with a Russian-speaking mother). The children attended full-time monolingual German or bilingual Russian-German nursery schools in Berlin or Bavaria.

Written parental consent was provided prior to the experiment. Parents filled out a questionnaire (based upon the Russian *Language Proficiency Test for Multilingual Children*, Gagarina et al., 2010) targeting the children's AoO, the family's use of German and Russian, and SES.

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

Table 1. Participants' age, IQ, and AoO for German

Variable	Mean (SD)	Range
Age	4;8 (0;2)	4;6–5;1
IQ	102.63 (8.21)	83–120
AoO German (in months)	24.95 (13.01)	0–48

An overview of the demographic information is presented in Table 1.

Following the Berliner Senate's manual (Senatsverwaltung für Wissenschaft, Gesundheit, Pflege und Gleichstellung, 2018), SES was classified on a 4-point scale based on parental education: (4) university degree, (3) a diploma program equivalent to college entrance requirements (Germany: Abitur, Russia: Attestat o Srednem Polnom Obščem Obrazovanii, or comparable in other country), (2) basic general education (Germany: Realschulabschluss, Russia: Attestat o Srednem Obščem Obrazovanii or comparable in other country), (1) German Hauptschulabschluss (max. nine years of school in total). In the analysis, we used the mean of both parents' scores. Parents were educated in Germany, Russia, Kazakhstan or Estonia (see Table 2).

All parents had completed at least 9 years of schooling. Data on education was missing from one father and one mother; in these cases, the other parent's score was used. Additionally, two mothers and one father did not fill out the country where they received their highest education (cf. Table 2).

Participants' hearing, neurological, socioemotional, and language development as well as IQ were within the norm. The participants were tested in their homes or preschools in a quiet room. The study was approved by the German Linguistics Society ethics committee, and it was carried out according to the Declaration of Helsinki.

Materials and procedure

Language proficiency

Language proficiency was assessed with productive and receptive lexicon and sentence comprehension tasks. Parallel tests were used in both languages: the *Sprachstandstest Russisch für mehrsprachige Kinder* (SRUK; Gagarina et al., 2010) in Russian and the *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen* (PDSS; Kauschke & Siegmüller, 2009) in German. In both tasks, the total score is measured by the number of correctly labelled items on production and reception. For Russian, the total score is 72 (36 verbs and nouns each, 26 reception and 10 production), whereas for German, it is 80 (20 verbs and 20 nouns, each in production and reception). For sentence

Table 2. Overview of parental education

	1					**2**				**3**				**4**					**NA**
Country	**G**	**R**	**Ot**	**NA**	**Sum**	**G**	**R**	**Ot**	**Sum**	**G**	**R**	**O**	**Sum**	**G**	**R**	**Ot**	**NA**	**Sum**	
Mothers	3	2	1	1	**7**	3	0	1	**4**	4	6	1	**11**	6	6	2	1	**15**	**1**
Fathers	4	2	2	0	**8**	9	3	4	**16**	3	2	3	**8**	3	1	0	1	**5**	**1**

Note: 1–4 represent the scale measuring parental education, G is for parents completing their education in Germany, R for Russia, Ot for other countries (e.g., Kazakhstan or Estonia), and NA for missing information.

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

comprehension, the SRUK grammar comprehension subtest (Gagarina et al., 2010) was used for Russian and TROG-D: *Test zur Überprüfung des Grammatikverständnisses* (Fox, 2011) for German. In TROG-D, the participant listens to a sentence and chooses the correct picture out of four possibilities, including three distractors (lexical or morphosyntactic). There are 21 blocks of four sentences, and a block is scored as incorrect if one or more sentences are identified incorrectly. The test ends after five consecutive incorrect blocks. SRUK uses 11 blocks, each containing two similar structures. A composite score for language proficiency in each language was calculated by summing up the scores from the three tasks. Thus, the maximum composite scores were 83 for Russian and 101 for German.

Narrative tests

Narrative production (telling) was elicited using the Baby Goats and Baby Birds stories from the MAIN (Gagarina et al., 2012, 2019). Native speakers of Russian and German administered the test in a monolingual mode with a mean interval of 18 days between languages. Story content (Baby Birds/ Baby Goats) and order of testing (L1/L2) were counterbalanced across participants. Twenty-two children were tested in German first, and 16 first in Russian. Following the protocol (see Gagarina et al., 2012), the experimenter puts three envelopes in front of the child and asks him/her to choose one and look at the sequence of the pictures in the story. Then, the story is folded up so that the child sees only the first two pictures (then four, then all six pictures), and s/he is asked to tell the story in such a way that only the child can have an overview over the sequence of pictures (non-joint attention mode).

The narratives were transcribed verbatim in CHAT (CHILDES; MacWhinney, 2000). The transcribers and coders were trained native speakers of the respective language, and their transcripts were double-checked by a third person for accuracy. The procedure included transcribing one-third of audio files per language and performing a Cohen kappa reliability score with computerised language analysis; the agreement on the transcribed words was over 80%.

The analysis of macrostructure production was operationalised by SS (according to the MAIN protocol) and SC in the following way. For SS, setting information (one for time and one for place, maximum two points), one point each for the production of an episode component: Goal (G), Attempt (A), Outcome (O), and an internal state as an initiating event and as a reaction (maximum five points) in the three episodes of the story were calculated. Thus, the maximum score was 17 points. Following previous studies (Yang et al., 2023), SC had a maximum score of 3 per episode (maximum 9 for all three). It was scored as follows: for factual components: A, O or AO score 1; for inferred components: G, GA, or GO score 2; for full GAO score 3.

Cognitive tests

For the current study, the subtests "Figure Ground," "Form Completion," and "Attention Divided" from the *Leiter International Performance Scale* test (Leiter-3; Roid et al., 2013) were administered. The "Attention Divided" subtest was performed together with the narrative task. The two other subtests were performed six months earlier, as part of the IQ testing for inclusion in the study. Following the test manual, the test was performed nonverbally. Each subtest allows for a teaching trial, which can be repeated up to three

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

times. There are no time limits. According to the manual, the subtests described below show a high internal consistency, with alphas between .80 and .94 for ages three to five.

Figure Ground measures the ability to identify embedded figures within a complex background, and it tests shifting ability. In this subtask, the child is presented with an image (stimulus) and two to five cards. The cards show a figure or item that is located ("hidden") on the stimulus image. The child must identify and point to the figures in the stimulus. There are a total of 12 stimulus images, and a point is awarded for each item the child correctly locates (maximum score 33). The stopping rule is six cumulative incorrect responses.

Form Completion assesses visual working memory; it requires that the child goes back and forth to recognise a whole picture from randomly displayed pieces. The child is presented with an image and, in the first two trials, must push blocks together to resemble the stimulus image. Then, the child sees a card with an incomplete item and must point to the complete version of the item within the stimulus image. There are 15 stimulus images; each correctly placed block or card gives a point (maximum score 36). The test is stopped if a child gives seven cumulative incorrect responses.

The subtest Attention Divided tests inhibition skills. It consists of ten cards (four of which show a red triangle) and 12 yellow disks. The child has two tasks. S/he must throw the 12 yellow discs into a yellow cup as quickly as possible. At the same time, the examiner flips over cards (one per second). The child must – while placing the yellow discs – slap the card that has been flipped over if it is a red triangle. If the card does not show a red triangle, the child should not interact with it. The test ends when the examiner has flipped over all 10 cards. There is no stop rule. The subtest Attention Divided results in two scores (max. 16 for each), one for correctly performed actions and one for incorrectly performed actions (a slap on a card without a triangle, a missed triangle, or any yellow disks remaining outside of the cup when the task is finished). Since the test manual describes the incorrect score as identifying a greater set of possible executive function problems (Roid et al., 2013), this score was used in the analysis.

Data analysis

All statistics were performed in R (R Core Team, 2018). In the first step, we compared the performance in language proficiency as well as narrative macrostructure across the two languages in a paired sample two-tailed t-test. For language proficiency, we used a composite score by summing up the correct responses from all three language tests.

Then, we calculated separate hierarchical linear regression models with SS and SC as dependent variables, to find out which predictors affected the two macrostructure dimensions. In total, there were four models, two for each language separately. All four models considered the same predictors with the same order: demographic factors of sex, SES, language proficiency, and cognitive skills (Figure Ground, Form Completion, Attention Divided).

Since our study is exploratory, we have combined hierarchical steps with nonhierarchical, with theoretical motivation to first control for demographic variables and language proficiency, and then to consider the effect of the three cognitive skills. However, the order in which the individual predictors (independent variables) were added within the cognitive skills group: Figure Ground and Form Completion was stepwise rather than hierarchical. In fact, to more accurately understand the effects of

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

both tasks, we calculated models using both orders of entry (first Figure Ground and first Form Completion).

The order of the steps in all four models was thus:

1. Sex
2. SES (average parental education)
3. Language proficiency German/Russian
4. Figure Ground/Form Completion
5. Figure Ground/Form Completion
6. Attention Divided

Results

Only significant findings ($p \leq .05$) are reported.[1] We describe the different orders of entry only where the results differ. For all the models used, please see the Appendix.

Differences across the two languages

The children's proficiency in Russian was significantly higher than in German, $t(37) = -8.16$, $\alpha = .05$ $p < .001$. Moreover, German proficiency has a much wider range (15%-83% correct) than Russian (43%-88% correct). Although the maximum score was similar, the minimum one was much lower in German. This indicates high variability in German proficiency among participants.

Similarly, the narrative macrostructure across languages was compared. For SS, the scores in Russian ($M = 5.42$, $SD = 1.90$) were significantly higher than for German ($M = 3.82$, $SD = 2.33$), $t(37) = -4.03$, $\alpha = .05$, $p < .001$. For SC, no significant difference between Russian ($M = 3.63$, $SD = 1.37$) and German ($M = 2.90$, $SD = 1.83$), $t(37) = -1.83$, $\alpha = .05$, $p = .07$) was found.

Story structure

The model used to predict SS in Russian never significantly explained the variance in the data. The model for German SS including demographic factors and German language proficiency ($F(4,34) = 4.13$, $p = 0.01$, $R^2 = .27$) was significantly better than the models with only demographic variables, which did not significantly explain the variance ($\Delta F(1,34) = 9.35$, $p = .004$, $\Delta R^2 = .20$). Within this model, the individual variable German language proficiency was significant ($t(34) = 3.06$, $\beta = .08$, $p = .004$). No other predictor additions (in any order) were significant.

Story complexity

The models for Russian SC were not significant when they included demographic variables or Russian language proficiency. The model including Figure Ground ($F(4,33) = 2.85$,

[1]The demographic factors sex and SES were always controlled for first, but were never significant predictors in any of the models. Similarly, Attention Divided was never significant.

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

$p = .04$, $R^2 = .26$) was significantly more predictive, $\Delta F(1,33) = 5.40$, $p = .03$, $\Delta R^2 = .12$. Figure Ground was the best individual predictor ($t(33) = 2.32$, $p = .03$), followed by Russian language proficiency, which had a significant negative effect ($t(33) = -2.18$, $\beta = -.05$, $p = .04$). This is most probably an example of suppressor regression, implying that the two variables are correlated. This could not be confirmed numerically using any classic test for correlation, however (Pearson's, $p = .06$; Spearman's, $p = .05$; or Kendall's, $p = .05$). No other steps significantly improved the model. When Form Completion was added before Figure Ground, the models never reached significance.

The model for German SC remained nonsignificant when it included demographic variables and German language proficiency. When Figure Ground was added first, the model was not significant. Though the addition of Form Completion resulted in significant improvement, the model including it remained nonsignificant.

When Form Completion was added first, it significantly improved the model, ($\Delta F(1,33) = 6.10$, $p = .02$, $\Delta R^2 = .14$). The model including sex, SES, German language proficiency, and Form Completion was significant ($F(4,33) = 2.82$, $p = .04$, $R^2 = .26$). Form Completion was a significant individual predictor ($t(33) = 2.47$, $\beta = .05$ $p = .02$). Adding Figure Ground as the next step did not improve the model, and in fact made the model to lose significance.

Discussion

The present study examined the effect of cognitive skills on macrostructure production (SS and SC) in both languages of 38 Russian-German-speaking preschool children, after controlling demographic factors (sex and SES) and language proficiency (measured by lexical and morphosyntactic skills in both languages). Subsequently, the performance in macrostructure dimensions across the two languages was also compared.

Macrostructure production was analyzed for SS and SC using MAIN (Gagarina et al., 2012, 2019). The cognitive tasks were assessed nonverbally via Leiter-3 (Roid et al., 2013), motivated by their relevance for telling a picture-based story: Figure Ground (shifting), Form Completion (visual working memory), and Attention Divided (inhibition).

Results showed that SS was higher in the L1 Russian than in the L2 German, aligning with findings by Tribushinina et al. (2022). However, this contrasts with studies such as Lindgren and Bohnacker (2022), which found SS to be higher in the societal language Swedish than in the home language German. For SC, as predicted, there were no differences across languages, thus confirming that it is less language-dependent (Lindgren & Bohnacker, 2022).

The controlled demographic factors (sex, SES) did not add significant variance in either SS or SC in both Russian and German, which is in contrast with previous research on bilinguals (Armon-Lotem et al., 2011; Sheldon & Engstrom, 2005). Language proficiency was a significant predictor of SS, not of SC; however, its effect was only measurable in the weaker language, L2 German SS, not L1 Russian. The differing effects of language proficiency on SS across Russian and German could be explained by variability in vocabulary. Specifically, children's language proficiency in Russian was high, meaning they had enough vocabulary to produce a story. In contrast, the German vocabulary scores were lower with larger variability, meaning some children had limited German vocabulary. This could prohibit them from producing a coherent SS. The same pattern was found in Lindgren and Bohnacker (2022), where language proficiency significantly affected SS in the weaker language but not in the dominant Swedish.

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

After controlling demographic factors and language proficiency, cognitive skills did not affect SS either in Russian or German. However, they affected SC in both languages in different ways. While Figure Ground was a significant predictor of Russian SC (together with a negative effect of language proficiency), for German SC, only a model including Form Completion reached significance. Finally, Attention Divided did not affect macrostructure in any of the languages.

Our results on cognitive effects on narratives follow previous literature in which macrostructure was influenced by shifting (Khan, 2013) and visual working memory (Veraksa et al., 2020). Nevertheless, they are in contrast with previous studies on the positive effect of inhibition on narrative macrostructure (Clark-Whitney & Melzi, 2023). This might be because the Attention Divided test was conducted six months earlier than the other cognitive tests, potentially resulting in much lower performance and therefore lack of significant impact on narrative macrostructure.

Since SC did not differ between the two languages, we ask the question: *Why does Figure Ground affect L1 Russian SC and Form Completion L2 German SC?* The differential effects of cognitive skills on the two languages could be linked to differences in language proficiency. In Russian, children had higher language proficiency, which negatively affected SC when Figure Ground was added to the model. This suggests children focusing more on specific linguistic details to produce correct grammar in the narrative and thus paying less attention to creating more complex episodes in Russian. Consequently, children made frequent shifts to depict suitable objects and combine them to create a cohesive story. Contrariwise, children had lower proficiency in German which led to focusing more on holding in memory visually presented stimuli and connecting the story elements into a coherent whole.

Nevertheless, future studies will be necessary to truly differentiate the various cognitive skills affecting storytelling over time and across different language pairs of bilinguals. Furthermore, they should refrain from administering cognitive tests at different time points, particularly with preschool-aged children who are still developing their cognitive skills.

In sum, to our knowledge, studies investigating the effect of cognitive skills on narratives have not been conducted in both languages of bilinguals. Thus, we consider our study exploratory. The findings advance our understanding of how cognitive factors impact bilinguals' SC by supporting that shifting and visual working memory have a differential influence on SC in L1 and L2 of preschool bilinguals, when sex, SES, and language proficiency are controlled for.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0305000924000394.

Acknowledgements. Data from this study were collected in the research project "Verbal and Non-Verbal Indicators for Identifying Specific Language Impairment in Successive Bilingual Preschoolers (DFG Az. LI 410/5-1 & Az. GA 1424/3-1)" funded by the German Research Foundation. We thank the German Research Foundation as well as all those who participated in the project for data collection, transcription and coding. Finally, we thank the two anonymous reviewers and the Action Editor of this Journal for their insightful and thorough comments.

Competing interest. The authors declare none.

References

Armon-Lotem, S., **Walters, J.**, & **Gagarina, N.** (2011). The impact of internal and external factors on linguistic performance in the home language and in L2 among Russian-Hebrew and Russian-German

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

preschool children. *Linguistic Approaches to Bilingualism*, **1**(3), 291–317. https://doi.org/10.1075/lab.1.3.04

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

Clark-Whitney, E., & **Melzi, G.** (2023). Executive function and narrative language abilities in emergent bilingual preschoolers: An exploratory study. *Language, Speech, and Hearing Services in Schools*, **54**(2), 584–599. https://doi.org/10.1044/2022_LSHSS-21-00190

Core Team, R. (2018, December 20). R. A language and environment for staistical computing. R version 3.5.2. Retrieved from www.R-project.org

Duinmeijert, 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.

Fiestas, C. E., & **Peña, E. D.** (2004). Narrative discourse in bilingual children: Language and task effects. *Language, Speech, and Hearing Services in Schools*, **35**, 155–168.

Fox, A. (2011). *Test zur Überprüfung des Grammatikverständnisses: TROG-D* (5th ed). Idstein, Germany: Schulz-Kirchner Verlag.

Gagarina, N., & **Bohnacker, U.** (2022). Storytelling in bilingual children: How organization of narratives is (not) affected by linguistic skills and environmental factors. *Linguistic Approaches to Bilingualism*, **12**(4), 407– 412. https://doi.org/10.1075/lab.00035.int

Gagarina, N., **Klassert, A.**, & **Topaj, N.** (2010). Russian language proficiency test for multilingual children. In *ZAS papers in linguistics*, **54**. Berlin: ZAS

Gagarina, N., **Klop, D.**, **Kunnari, S.**, **Tantele, K.**, **Välimaa, T.**, **Balčiūnienė, I.**, **Bohnacker, U.**, & **Walters, J.** (2012). Multilingual assessment instrument for narratives (MAIN). *ZAS Papers in Linguistics*, **56**. Berlin: ZAS.

Gagarina, N., **Klop, D.**, **Kunnari, S.**, **Tantele, K.**, **Välimaa, T.**, **Bohnacker, U.**, & **Walters, J.** (2019). MAIN: Multilingual assessment instrument for narratives – revised. *ZAS Papers in Linguistics*, **63**, 20. https://doi.org/10.21248/zaspil.63.2019.516

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, **49**(1), 4–14.

Kapalková, S., **Polisenska, K.**, **Markova, L.**, & **Fenton, J.** (2016). Narrative abilities in early successive bilingual Slovak–English children: A cross-language comparison. *Applied Psycholinguistics*, **37**, 145–164.

Kauschke, C., & **Siegmüller, J.** (2009). *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen*. Munich: Elsevier.

Khan, K. S. (2013*). The relation between cognitive skills and language skills in typically developing 3 ½ to 6 year olds. Ph.D. Thesis.* Philadelphia, PA: The Pennsylvania State University.

Lindgren, J., & **Bohnacker, U.** (2022). 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*, **12**(4), 479– 508.

Lindgren, J., **Tselekidou, F.**, & **Gagarina, N.** (2023). Multilingual assessment instrument for narratives: Recent developments and new language adaptations. *ZAS Papers in Linguistics*, **65**, 132. https://doi.org/10.21248/zaspil.65.2023.606.

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

Mayer, M. (1969). *Frog, where are you?* New York: Dial Books.

Mozeiko, J., **Le, K.**, **Coelho, C.**, **Krueger, F.**, **Grafman, J.** (2011). The relationship of story grammar and executive function following TBI. *Aphasiology*, **25**, 826–835.

Oshchepkova, E. S., **Shatskaya, A. N.**, **Dedyukina, M. I.**, **Yakupova, V. A.**, & **Kovyazina, M. S.** (2022). The relationship between cognitive flexibility, bilingualism and language production: Evidence from narrative abilities in senior preschoolers from the Republic of Sakha (Yakutia). *RUDN Journal of Language Studies, Semiotics and Semantics*, **13**(1), 125–143. 10.22363/2313-2299-2022-13-1-125-143

Paradis, J., **Genesee, F.**, & **Crago, M. B.** (2010). *Dual language development and disorders: A handbook on bilingualism and second language learning* (2nd ed.). Baltimore, MD: Paul H. Brookes.

Pearson, B. Z. (2002). Language and mind in the stories of bilingual children. In L. Verhoeven & S. Strömqvist (Eds.), *Narrative development in a multilingual context* (pp. 135–174). Amsterdam: John Benjamins.

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press

Roid, G., **Miller, L.**, **Pomplun, M.**, & **Koch, C.** (2013). *Leiter-3: International performance scale*. 3rd ed. Wood Dale, IL: Stoelting Company.

Senatsverwaltung für Wissenschaft, Gesundheit, Pflege und Gleichstellung (Ed.). (2018). *Grundauswertung der Einschulungsdaten in Berlin 2017*. https://www.berlin-suchtpraevention.de/wpcontent/uploads/2019/04/2018_Einschulungsuntersuchung_Berlin.pdf

Sheldon, A., & **Engstrom, H.** (2005). Two systems of mutual engagement: The co-construction of gendered narrative styles by American preschoolers. In J. Thornborrow & J. Coates (Eds.), *The sociolinguistics of narrative* (pp. 171–192). Amsterdam: John Benjamins.

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

Tribushinina, E., **Irmawati, M.**, & **Mark, P.** (2022). Macrostructure in the narratives of Indonesian-Dutch bilinguals: Relation to age and exposure. *Linguistic Approaches to Bilingualism* **12**(4), 540–570.

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.

Veraksa, A., **Bukhalenkova, D.**, **Kartushina, N.**, & **Oshchepkova, E.** (2020). The Relationship between executive functions and language production in 5–6-year-old children: Insights from working memory and storytelling. *Behavioral Science*, **10**(52), 1–1220.

Wallach, G. P. (2008). *Language intervention for school-age students: Setting goals for academic success*. Mosby Elsevier.

Yang, W., **Chan, A.**, & **Gagarina, N.** (2023). Left-behind experience and language proficiency predict narrative abilities in the home language of Kam-speaking minority children in China. *Frontiers in Psychology*, **13**, Article 1059895. https://doi.org/10.3389/fpsyg.2022.1059895

Cite this article: Tselekidou, F., Stadtmiller, E., Süss, A., Lindner, K., & Gagarina, N. (2025). Cognitive skills differentially influence narrative macrostructure in bilinguals' L1 and L2. *Journal of Child Language* **52,** 1437–1447, https://doi.org/10.1017/S0305000924000394

https://doi.org/10.1017/S0305000924000394 Published online by Cambridge University Press