

ORIGINAL ARTICLE

Narrative macrostructure and microstructure profiles of bilingual children with autism spectrum disorder: differentiation from bilingual children with developmental language disorder and typical development

Krithika Govindarajan*  and Johanne Paradis 

Department of Linguistics, University of Alberta, Edmonton, Canada

*Corresponding author. Email: krithika@ualberta.ca

(Received 27 July 2020; revised 24 August 2022; accepted 13 September 2022; first published online 24 November 2022)

Abstract

Children with autism spectrum disorder (ASD) show heterogeneous language profiles beyond early language delays. Understanding the second language profiles of bilingual children with ASD is important for clinical practice in diverse societies. Accordingly, we examined the narrative abilities of bilinguals with ASD, with developmental language disorder (DLD), and with typical development (TD) to determine which narrative components best differentiate bilinguals with ASD from the other groups. Participants were 29 bilingual children with ASD, DLD, and TD who were matched for age (mean = 6;8), nonverbal intelligence, and receptive vocabulary. Narratives were coded for macrostructure (story grammar (SG) scores, number of individual SG components) and microstructure (syntactic complexity, mean length of utterance, lexical diversity, and story length). The TD group had superior SG scores, included more SG components, and used longer utterances and more complex syntax than the ASD group, whereas no differences were found between the clinical groups. For SG components requiring perspective-taking abilities, the ASD group had worse performance than the TD and DLD groups. Our results suggest that bilingual children with ASD show weaknesses in both macrostructure and microstructure, which can overlap with children with DLD. The linguistic profiles of bilingual children with ASD and DLD are thus both overlapping and distinct.

Keywords: Bilingualism; autism; developmental language disorder; narratives

A comprehensive understanding of the linguistic profiles of bilingual children with autism spectrum disorder (ASD) is essential for clinical practice in culturally and linguistically diverse societies (Kay-Raining Bird, Genesee & Verhoeven, 2016; Paradis, Genesee & Crago, 2021; Paradis & Govindarajan, 2018). In comparison

© The Author(s), 2022. 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.

with monolingual children, bilingual children (i.e., children who are exposed to two languages, both simultaneously or sequentially) display more individual variability in their language abilities and developmental trajectories because their language learning context is more complex. For example, bilingual children's input space is divided, often unevenly, between their two languages, the onset of learning a second language (L2) can occur after the first language (L1) has been established (except for simultaneous bilinguals), and the quantity and quality of input and output can be different in each language and over time (Lauro, Core & Hoff, 2020; Paradis, 2019; Unsworth, 2016). This variation in the language abilities of bilingual children complicates the process of determining whether abilities in the majority language, typically the L2, that are below monolingual-based age expectations could be due to language input and length of exposure factors, or due to an inherent language disorder (Bedore & Peña, 2008; Paradis et al., 2021). This complication in assessment impacts not only the identification of developmental language disorder (DLD), but it also impacts the determination of whether bilingual children with ASD have normalized language or show evidence of language difficulties/disorder beyond early language delays. While children with ASD typically have delay in the onset of language and also exhibit lifelong deficits in social communication, their structural language development (vocabulary and morphosyntax) after the early years can vary. In other words, there are multiple language phenotypes in ASD (including children who remain minimally verbal, e.g., Ellis Weismer & Kover, 2015), with some children showing intact structural language abilities, while others present with structural language difficulties, and in some cases, these difficulties are similar to DLD (e.g., Durrleman & Delage, 2016; Meir & Novogrodsky, 2019; Roberts, Rice & Tager-Flusberg, 2004; Kjelgaard & Tager-Flusberg, 2001; Wittke et al., 2017). Therefore, heterogenous language trajectories in ASD, combined with the variation displayed in dual language learning, would complicate identifying the presence of language difficulties/disorder in bilingual children with ASD.

Most research to date on children with ASD who are exposed to two languages has focused on their capacity for bilingual development. This research has been conducted mainly with young, preschool age children and has compared bilinguals with ASD to monolinguals with ASD to understand whether the children exposed to two languages were lagging behind their monolingual counterparts for early language and other developmental milestones (e.g., Ohashi et al., 2012; Petersen, Marinova-Todd & Mirenda, 2012; Reetzke, Zou, Sheng, & Katsos, 2015; Valicenti-McDermott et al., 2013). These studies have overwhelmingly shown that dual language learning in the early years does not exacerbate the language delays and behavioral patterns in children with ASD. However, this body of research does not provide much insight into how dual language development unfolds in children with ASD past the early years and how their profiles of linguistic strengths and weakness compare to their bilingual peers with typical development (TD). Comparisons between bilinguals with ASD and with TD are crucial, as bilingual children with ASD are more likely to be similar to bilinguals with TD, rather than to monolinguals with ASD, in terms of their overall language learning environments (Paradis, 2016). Accordingly, the present study focused on the English L2 development of bilingual children with ASD in the school age years and included an age and language equivalent comparison group of bilinguals with TD.

In addition to TD age peers, the language abilities of monolingual children with ASD have also been considered in light of a DLD profile in order to understand the extent to which structural language abilities of children with ASD overlap with this other clinical group whose primary condition is language impairment (e.g., Durrleman & Delage, 2016; Meir & Novogrodsky, 2019; Tager-Flusberg & Joseph, 2003; Wittke et al., 2017). The logic behind comparing these groups lies in the potential similarities for structural language on the one hand, and the potential dissimilarities at the level of discourse pragmatics, or social communication, on the other hand. ASD is a neurodevelopmental disorder that is characterized by deficits in social communication, along with the presence of restricted and repetitive patterns of behavior (American Psychiatric Association (APA), 2013). In particular, children (and adults) with ASD show deficits at the discourse-pragmatic level of language comprehension and use, with notable problems in taking the perspective of a listener and attributing mental states to themselves or others, among other behaviors comprising the *theory of mind* construct (Gerenser & Lopez, 2017). While such discourse-pragmatic deficits are nearly universal in children with ASD, the presence and extent of structural language difficulties varies, as mentioned earlier. DLD is also a neurodevelopmental disorder, but clinically significant deficits are narrower than those for ASD. Children with DLD have below age expectations for language development, although they have had adequate exposure to the target language, normal hearing, no frank neurological damage, no evidence of intellectual disability, and no evidence of social communication deficits consistent with ASD (Leonard, 2014; Schwartz, 2017). Children with DLD show protracted development in all structural language domains, with morphosyntax being particularly affected beyond what their general delay might indicate (Leonard, 2014; Oetting & Hadley, 2017). Even though children with DLD, by definition, do not have intellectual disabilities, they exhibit mild deficits in perceptual and cognitive systems that could, in part, underlie their language learning difficulties (Leonard, 2014; Schwartz, 2017). Therefore, what some children with ASD could have in common with children with DLD is structural language abilities below age expectations; in contrast, where children with ASD could be expected to behave differently from children with DLD would be for social communication abilities, especially those related to theory of mind.

Narrative storytelling from wordless picture books is a measure of language abilities that is well suited for examining both discourse-pragmatic abilities as well as structural language abilities in a realistic communicative task. As explained in more detail below, this is because good storytelling involves socio-cognitive-linguistic interface skills which must take the listener's needs into account to produce a coherent story (macrostructure), as well as lexical and morphosyntactic skills (microstructure). For this reason, much research has focused on narratives as a measure of language abilities in children with ASD and with DLD (e.g., Banney, Harper-Hill & Arnott, 2015; Capps, Losh & Thurber, 2000; Losh & Capps, 2003; Mäkinen et al., 2014; Norbury & Bishop, 2003; Reilly, Losh, Bellugi & Wulfeck, 2004). Cross-disorder comparisons between ASD and DLD have been conducted using narratives (e.g., Engberg-Pedersen & Christensen, 2017; Norbury, Gemmel & Paul, 2014). But, as described below, there have been few studies of narrative skills in bilinguals with ASD, and to date, no cross-disorder study including bilinguals

with ASD and DLD. In response to this gap in knowledge, the present study compared the English-L2 narrative skills of school age bilingual children with ASD, TD, and DLD. More specifically, we sought to identify which narrative macrostructure and microstructure components differentiated the children with ASD from the other groups. In so doing, we aimed to contribute to our understanding of the similarities and differences between the L2 profiles of bilingual children with ASD and with DLD.

Narrative tasks: macro- and microstructure

There are different types of narratives used in language acquisition research, for example, fictional and personal narratives. In personal narratives, children relate something they have experienced, whereas, in fictional narratives, children tell a story about other children or characters following prompts. Fictional narratives can be elicited through retell tasks, in which a child repeats a story they have just heard or through story generation tasks, in which a child produces a story while looking at a wordless picture book. Henceforth, “narratives” refer to fictional narratives.

The narratives children produce are generally analyzed at two different levels, namely macrostructure and microstructure. The term macrostructure refers to the overall content and organization of the story. The story grammar model (Stein & Glenn, 1979) is the most widely used to study narrative macrostructure and was adopted for this study. Macrostructure analyses typically focus on children’s inclusion of story grammar components, the number of story episodes included, and the complexity of episode structures. In the story grammar model, narratives consist of six categories of information or story grammar components. According to this model, a story has (1) a *Setting* that introduces the time, place, and characters in the story, (2) an *Initiating Event* that sets up the problem or dilemma in the story, (3) an *Internal Response* or the character’s response to the Initiating Event, (4) an *Attempt* or an action of the character to solve the problem, (5) the *Outcome* or the result of the previous action, and (6) a *Response* or how a story character responds to the outcome (e.g., Iluz-Cohen & Walters). Stories contain at least one episode, that is, they contain three core story grammar units (namely, an initiating event, an attempt, and an outcome), but they may consist of more than one episode (Schneider, Dubé, & Hayward, 2005). It is important to note that producing story grammar components requires discourse-pragmatic skills, like taking a listener’s needs into account, in order to produce a coherent story. Skipping the initiating event or the consequence would render a story difficult to follow. Also noteworthy is that mentioning the internal responses or reactions of characters requires perspective-taking and theory of mind (e.g., Capps et al., 2000; Siller, Swanson, Serlin, & Teachworth, 2014; Tager-Flusberg & Sullivan, 1995), for example, understanding that different characters in the story can have different internal mental states.

Story grammar can be measured in different ways, but two ways that are relevant to our study are outlined here: (1) total story grammar scores per story, where the number of components included by the child are counted (e.g., Losh & Capps, 2003; Mäkinen et al., 2014; Norbury & Bishop, 2003; Norbury et al., 2014); (2) frequency of use of individual components (e.g., Diehl, Bennetto, & Young, 2006). The former is a measure of global story coherence, because the more components included, the

more coherent the story. The latter focuses on whether certain component scores are more or less likely to be produced by a child. This approach can reveal whether a child with ASD might be less likely to produce components that require perspective-taking.

In contrast to macrostructure, the term microstructure refers to a local level of analysis in which the linguistic structures used to produce stories are analyzed. It includes measures of productivity and measures of complexity (e.g., Baixauli, Colomer, Rosello, & Miranda, 2016; Justice et al., 2006). The term *productivity* refers to the amount of material produced in a narrative. This may be measured by looking at the total number of words produced (story length), the number of different words produced (lexical diversity), or by calculating the number of clausal-level elements (Justice et al., 2006; Mäkinen et al., 2014). Examining the mean length of utterances (MLU), use of complex syntax, or morphological errors are some ways of examining *complexity* or grammatical functioning (Altman, Armon-Lotem, Fichman, & Walters, 2016; Baixauli et al., 2016; Justice et al., 2006; Mäkinen et al., 2014). In sum, microstructure components can cover a wide range of linguistic features.

The following macrostructure components were examined in this study: total story grammar scores for individual stories as well as the frequency of individual story grammar components used, including characters' internal reactions, across the set of stories. For microstructure, we examined lexical diversity and story length in words (measures of productivity) as well as mean utterance length and use of complex syntax (measures of complexity).

Narrative macrostructure in monolingual children with ASD and DLD

Children with ASD

Studies have overwhelmingly found significant differences between children with ASD and TD controls for macrostructure, including total story grammar scores or the use of individual components (e.g., Banney et al., 2015; Diehl et al., 2006; Losh & Capps, 2003; Mäkinen et al., 2014; Norbury et al., 2014; Rumpf, Kamp-Becker, Becker & Kauschke, 2012; Smith Gabig, 2008; Suh et al., 2014; Tager-Flusberg, 1995; but see Norbury & Bishop, 2003; Young et al., 2005). A recent meta-analysis by Baixauli and colleagues found that macrostructure measures differentiated children with ASD from TD controls, with a large effect size (Baixauli et al., 2016). For example, Norbury et al. (2014) examined narratives in 6 ½–15 years old children with ASD. Children with ASD were matched to their typically developing peers on age, nonverbal ability, and language. In comparison to their typically developing peers, the children with ASD produced less coherent narratives and omitted important story components. Similarly, Mäkinen et al. (2014) examined narratives in 5–10-year-old Finnish-speaking children with ASD and TD. They found that the children with ASD produced stories with less informative content/fewer components than their TD peers. Similarly, Banney et al. (2015) examined narratives in 9–15-year-old children with ASD who were matched on age, nonverbal intelligence, and language skills to children with TD. The children with ASD produced less coherent stories with fewer story grammar components. As mentioned above, differences have also been found for individual story grammar components.

For example, Tager-Flusberg (1995) found that children with ASD were less likely to include outcomes (resolutions) in their stories than children with TD. Similarly, in Banney et al. (2015), children with ASD were less likely to include outcomes or internal responses in their stories compared to children with TD.

Children with DLD

There is an extensive body of research comparing the narrative skills of children with DLD to those of children with TD, but results are conflicting for narrative macrostructure. Some studies have found children with TD to obtain higher story grammar scores or include more narrative content, that is, more story grammar components, and hence, more coherent stories (e.g., Bishop & Donlan, 2005; Mäkinen et al., 2014; Norbury et al., 2014; Reilly et al., 2004; Torng & Sah, 2020); whereas, other studies have not found macrostructure to distinguish between TD and DLD groups (e.g., Dodwell & Bavin, 2008; Norbury & Bishop, 2003; Tsimplici, Peristeri & Andreou, 2016). The conflicting findings could be due, in part, to methodological differences. For example, in Norbury et al. (2014), participants were not matched on language abilities. Furthermore, studies have elicited stories in different ways. For example, while Mäkinen et al. (2014) used story generation, story recall tasks have also been used in other studies, such as in Dodwell and Bavin (2008). Differences in results can emerge depending on whether story retell or story generation was used; story generation being a more difficult task (Schneider, 1996; Schneider et al., 2005). Whatever the reason, the research with DLD stands in contrast to the research on children with ASD in terms of how consistent difficulties with macrostructure have been found.

Narrative microstructure in monolingual children with ASD and DLD

Children with ASD

Regarding productivity, while some studies have found children with ASD to produce shorter stories than TD controls (e.g., Norbury et al., 2014; Rumpf et al., 2012; Tager-Flusberg, 1995), others have not found significant differences for length (e.g., Banney et al., 2015; Mäkinen et al., 2014). Conflicting findings have also emerged for complexity. Children with ASD have been reported to produce shorter utterances (Mäkinen et al., 2014; Norbury et al., 2014; Smith Gabig, 2008; Tager-Flusberg, 1995) and use less complex syntax than their TD peers (e.g., Banney et al., 2015; Capps et al., 2000; Mäkinen et al., 2014), but similar patterns of performance for both groups have been reported for utterance length (e.g., Kauschke, van der Beek, & Kamp-Becker, 2016; Rumpf et al., 2014) as well as for syntactic complexity (e.g., Diehl et al., 2006; Rumpf et al., 2014). While it may appear difficult to generalize from these studies, Baixauli et al. (2016) reported in their meta-analysis that measures of productivity and complexity differentiate between children with ASD and children with TD, with a moderate effect size. The variation in the findings could be expected given the heterogeneity in structural language development in ASD. In sum, in contrast to the findings for narrative macrostructure, there are inconsistent findings on whether children with ASD are similar or dissimilar to their TD peers for narrative microstructure.

Children with DLD

Not surprisingly, given the structural language difficulties that define DLD, significant differences have often been found for narrative microstructure, for both measures of productivity (e.g., story length and lexical diversity) and measures of complexity such as MLU and the use of complex syntax (e.g., Colozzo et al., 2011; Fey et al., 2004; Mäkinen et al., 2014; Norbury & Bishop, 2003; Norbury et al., 2014; Reilly et al., 2004; Schneider, Hayward & Dubé, 2006). For example, children with DLD produce shorter stories and use a less diverse vocabulary when narrating than age-matched TD children (Fey et al., 2004; Colozzo et al., 2011; Reilly et al., 2004; but see Norbury & Bishop, 2003). Coming to measures of complexity, the results are consistent: children with DLD produce narratives with more grammatical errors and fewer complex sentences and have difficulties introducing referents (e.g., Colozzo et al., 2011; Fey et al., 2004; Mäkinen et al., 2014; Norbury & Bishop, 2003; Norbury et al., 2014). For example, in Mäkinen et al. (2014), narratives produced by age-matched Finnish children with and without DLD were analyzed. The children with DLD showed reduced syntactic complexity, produced shorter utterances as well as made more morphological errors when compared to the TD group. In sum, unlike the results for narrative macrostructure discussed earlier, studies are highly consistent in finding microstructure to be an area of weakness in DLD, especially for measures of complexity.

Direct comparisons of narrative abilities in monolingual children with ASD and with DLD

To date, only a handful of studies have compared the narratives produced by children with ASD to those produced by children with DLD (Colozzo, Morris & Mirenda, 2015; Goldman, 2008; Manolitsi & Botting, 2011; Norbury & Bishop, 2003, Norbury et al., 2014). For example, in Norbury and Bishop (2003), narratives produced by 6–10-year-old children with ASD were compared to those produced by age-matched children with DLD and with TD. No group differences were found between the clinical groups for either macrostructure or microstructure. Both clinical groups used less complex syntax and tense morphology and produced more ambiguous pronouns than the TD group. Norbury et al. (2014) also examined narratives produced by 6–15-year-old children with ASD, DLD, and TD. Although the ASD group had no structural language deficits on standardized assessments, both clinical groups patterned similarly and worse than the TD group for macrostructure and microstructure measures such as, inclusion of story components, the use of complex syntax, utterance length, story length, and lexical diversity. In contrast, other studies have found significant differences in macrostructure abilities between ASD and DLD. For example, in Colozzo et al. (2015), only children with ASD had significantly lower story grammar scores than children with TD, with the children with DLD occupying an intermediate position, and not differing significantly from either group. Similarly, in Manolitsi and Botting (2011), children with ASD group performed worse than children with DLD on a measure of story content and included characters' goals and actions less frequently than the DLD group. Further differences were found between children with ASD and children with

DLD in Goldman (2008) on two specific macrostructure measures: characters and outcomes. Taken together, these cross-disorder studies indicate that children with ASD and children with DLD can often show largely similar narrative profiles for microstructure but with differences for macrostructure.

Narrative macrostructure and microstructure abilities in bilingual children with ASD and DLD

Bilinguals with ASD

In Baldimtsi, Peristeri, Tsimpli, and Nicolopoulou (2016), narratives produced by 7–11-year-old bilingual children with ASD were compared to those produced by age-matched bilinguals with TD. Narratives were elicited in L2 Greek, and the participants had diverse L1 backgrounds. No significant differences were found between the ASD and TD bilingual groups for macrostructure. Group differences were also not found for complex syntax or lexical diversity. In contrast, Peristeri, Baldimtsi, Adreou and Tsimpli (2020) found significant differences between bilinguals with TD and bilinguals with ASD (Greek L2, diverse L1s), 7–12 years old, for both macro- and microstructure; the TD group included more story grammar components and used more complex syntax than the ASD group. Similarly, in Hoang, Gonzalez-Barrero and Nadig (2018), bilingual children with ASD, with diverse L1s, produced less coherent stories than TD bilinguals in a picture-sequencing task in their L2 (French), which was their dominant language. Microstructure measures were not examined in this study. Such limited and conflicting findings indicate that further research on the narrative skills of bilingual children with ASD is needed.

Bilinguals with DLD

Parallel to the monolingual literature, bilingual children with DLD are less skilled narrators than their bilingual age peers with TD, but with more consistent findings for microstructure (e.g., Altman et al., 2016; Boerma et al., 2016; Govindarajan & Paradis, 2019; Iluz-Cohen & Walters, 2012) than for macrostructure (Boerma et al., 2016 and Govindarajan & Paradis, 2019 vs. Altman et al., 2016 and Iluz-Cohen & Walters, 2012). Altman et al. (2016) and Iluz-Cohen and Walters (2012) examined narratives retold by bilingual English–Hebrew preschoolers with TD and with DLD in both languages. Both groups patterned similarly for macrostructure but differed on microstructure measures such as utterance length or the use of complex syntax. In contrast, other studies have found significant group differences for narrative macrostructure. In Boerma et al. (2016), 5–6-year-old bilingual children with and without DLD heard a model story in L2 Dutch and then produced a story with the support of pictures. The children with DLD produced fewer story grammar components than the children with TD. Similarly, in Govindarajan and Paradis (2019), narratives produced in L2 English by bilinguals with and without DLD aged 5–7 were examined. The bilinguals with DLD obtained significantly lower story grammar scores than the bilinguals with TD.

Importantly, unlike the research with monolinguals, no study to date has compared bilinguals with ASD to bilinguals with DLD on a narrative task.

The present study

For the present study, narrative language samples were gathered using a standardized narrative instrument from three groups of children (mean age = 6;8) who were acquiring English as an L2 with diverse L1 backgrounds: children with TD, ASD, and DLD. Groups were matched for age, nonverbal abilities, and L2 abilities. We asked the following research questions:

- (1) Does story coherence (story grammar components included) differentiate the narratives produced by the bilinguals with ASD from those of bilinguals with TD and with DLD?

This research question focuses on the story grammar components included within each story, that is, composite story grammar scores – a measure of narrative macrostructure. Therefore, this question focuses on children’s abilities to generate coherent stories from the separate picture sequences. Based on the existing research with monolinguals and bilinguals discussed earlier, the bilingual children with ASD were expected to include fewer story grammar components overall – have lower composite scores – than the bilinguals with TD and possibly show similar composite scores to the bilinguals with DLD.

- (2) What individual story grammar components differentiate the narratives produced by bilinguals with ASD from those of bilinguals with TD and with DLD?

This research question focuses on the use of story grammar components across all stories in order to determine if the bilinguals with ASD include certain story grammar components less frequently than the other groups. Narrative macrostructure is often measured by looking at composite story grammar scores, indexing story coherence within each story, as proposed for research question (1). However, looking at the use of individual story grammar components across several stories could potentially differentiate between the groups in another way. Given the social communication deficits common to children with ASD, the bilinguals with ASD in this study were expected to produce fewer story grammar components requiring perspective-taking abilities, for example, internal plans and reactions, than the other groups. Following prior monolingual research (e.g., Tager-Flusberg, 1995), bilinguals with ASD were also expected to produce fewer story outcomes than the other groups. Unlike other research with narratives and bilingual children with ASD and with DLD, this study gathered data from six different stories. Doing so enabled us to examine the frequency with which individual story grammar components were used by the children across stories; this approach to understanding macrostructure in narratives of bilinguals with ASD has not been employed in existing research.

- (3) What microstructure components differentiate the narratives produced by bilingual children with ASD from those of bilingual children with TD and with DLD?

The bilinguals with ASD were expected to pattern similarly to bilinguals with DLD with respect to narrative microstructure abilities, since this expectation is consistent with the majority of the existing research on cross-disorder comparisons between ASD and DLD discussed earlier. Both clinical groups were predicted to show inferior microstructure abilities to the bilinguals with TD, also consistent with prior research. More specifically, we expected both clinical groups to show reduced productivity (shorter stories in number of words and less diverse vocabulary) and complexity (shorter mean length of utterances and less use of complex syntax) in comparison with the bilingual TD group.

Method

Participants

Twenty-nine bilingual children participated in this study (9 with ASD, 10 with TD, and 10 with DLD) ranging in age from 5;4 to 9;1 (mean age = 6;8). The children were all L2 learners of with diverse L1 backgrounds. All children came from first-generation immigrant and refugee families where both parents were foreign-born and L2 speakers of English. Participants were in schools where English was the language of instruction and were living in an English majority-language city. The children with ASD were selected from a larger sample of 26 bilingual children with ASD who participated in a study that included language measures, cognitive measures, parent interviews, and parent questionnaires, with different goals than the present study (Paradis, Govindarajan & Hernandez, 2018). The larger sample of children showed a wide range of verbal abilities, including those who were minimally verbal. The children selected for the present study ($N = 9$) met the following inclusion criteria: (1) children were willing and able to produce oral narratives; and (2) children were willing and able to complete a test of receptive vocabulary and a test of nonverbal cognitive abilities. The children with TD and with DLD were chosen from larger samples of participants from previous studies (e.g., Paradis, Schneider & Sorenson Duncan, 2013) according to matching criteria with the ASD group (see below), but the sample in this study was not identical to that in any previous study. The Research Ethics Board at the University of Alberta, Canada, granted approval for this study.

Recruitment and background on children with DLD and ASD

Children with TD were recruited through schools as well as through agencies offering settlement assistance to newcomers. The children with DLD were diagnosed by certified referred speech-language pathologists who were working with them in a school setting. We specified to the speech-language pathologists that the children being referred to us needed to meet standard exclusionary criteria (e.g., no hearing loss, autism, or intellectual disabilities) (for more details about the DLD group, see Paradis et al., 2013). Children with ASD were also recruited through schools and from agencies offering assistance to newcomers. All the children referred to us had a clinical diagnosis of ASD established through an assessment protocol from a multidisciplinary team.

Our testing time with each child did not permit the inclusion of diagnostic measures specific to DLD or ASD, nor did we have access to health records for these

Table 1. Total and section scores from the Alberta Language Development Questionnaire for the TD, ASD, and DLD groups. Results of statistical analyses are in the final column

ALDeQ ^a	Group			Pairwise comparisons ^c
	ASD	DLD	TD	
Total: all sections	0.59(.12) ^b	0.44(.13)	0.81(.07)	ASD vs. TD $W = 78.5, p < .001$ ASD vs. DLD $W = 74, p = .02$ TD vs. DLD $W = 100, p < .001$
Section A: early milestones	.62(.29)	.39(.29)	.96(.08)	ASD vs. TD $W = 7, p = .002$ ASD vs. DLD $W = 79.5, p = .01$ TD vs. DLD $W = 97, p < .001$
Section B: current L1 abilities	0.35(.19)	0.25(0.17)	0.61(.21)	ASD vs. TD $W = 15, p = .02$ ASD vs. DLD $W = 66.5, p = .08$ TD vs. DLD $W = 91, p = .002$
Section C: behaviors and activity preferences	0.78(.16)	0.74(.21)	0.86(.11)	ASD vs. TD $W = 33, p = .35$ ASD vs. DLD $W = 53.5, p = .51$ TD vs. DLD $W = 64.5, p = .29$
Section D: family history	0.63(.39)	0.40(.46)	0.87(.32)	ASD vs. TD $W = 29, p = .14$ ASD vs. DLD $W = 58.5, p = .27$ TD vs. DLD $W = 77.5, p = .02$

Note.

^aSection and total scores can fall between 0.0 and 1.0.

^bmean (SD).

^cWilcoxon pairwise tests.

children. Therefore, we included a parent questionnaire, the Alberta Language Development Questionnaire (ALDeQ; Paradis, Emmerzael & Sorenson Duncan, 2010) as an additional source of information about children's early milestones (e.g., age at first word or age of first word combinations), their current L1 abilities, their behavior patterns and activity preferences, and family history (family members with language and learning difficulties (see Materials and Procedures for a description of the ALDeQ). The ALDeQ section and total scores for each group and analysis results are presented in Table 1. Kruskal–Wallis tests for the three groups followed by pairwise Wilcoxon tests were used to determine differences, and scores were also norm-referenced to the TD sample in Paradis et al. (2010). These analyses revealed the following: (1) TD had higher total scores than DLD and ASD, and DLD had lower total scores than ASD; TD total scores were within the normal range but the DLD and ASD scores were < -1.5 standard deviations below the mean. (2) TD had higher section A (early milestones) scores than the TD and ASD groups. (3) TD had higher section B (current L1 abilities) scores than DLD and ASD. (4) There was a significant difference between TD and DLD for scores in section D (family history). No group or pairwise differences emerged for section C (behavior and activity

preferences). In sum, the ALDeQ scores confirm the expected language development patterns based on the clinical diagnoses: children with DLD and ASD scored below the normal range of a larger sample of TD children and below the scores of the TD children in this study. The children with ASD and DLD were both delayed in their early milestones compared to the TD children and had weaker L1 abilities than their TD peers at the time of testing.

Group matching

As mentioned above, the children with TD and the children with DLD were selected from a larger sample on the basis of background variables that allowed them to be matched to the group with ASD. Because of our small sample size of children with ASD, we endeavored to create three closely matched groups on certain variables to ensure meaningful comparisons. First, the three groups had diverse L1 backgrounds, but participants were selected from previous studies in order to have similar distributions of L1 backgrounds to the ASD group (see Table A2 in the Appendix). Next, selection of the TD and DLD participants was made to create groupwise equivalencies for age, nonverbal cognitive abilities (Columbia Mental Maturity Scales, CMMS; Burgemeister, Blum & Lorge, 1972), general L2 abilities (Peabody Picture Vocabulary Test, PPVT-III; Dunn & Dunn, 1997), length of L2 exposure (input quantity; Alberta Language Environment Questionnaire, ALEQ, Paradis, 2011), and richness of L2 environment (input quality; Alberta Language Environment Questionnaire, ALEQ, Paradis, 2011). See *Materials and Procedures* for descriptions of the tests and questionnaire data used to obtain these variables. The descriptive variables and results of nonparametric Kruskal–Wallis tests followed by pairwise Wilcoxon comparisons (if significant) are presented in Table 2. Thus, participants in this study were matched groupwise on age, nonverbal intelligence, L2 receptive vocabulary, and richness of the L2 environment. However, in spite of our best efforts, the groups were not matched for length of L2 exposure even though they were matched for general L2 abilities. To account for the variance in our dependent variables that could be due to differences in L2 exposure rather than group, L2 exposure was entered as a covariate in the linear regression models. Following Kover and Atwood (2013), we complemented these matching analyses with analyses based on Cohen's *d* for the key variables of age, nonverbal abilities, and L2 abilities. There is no agreement for determining matching through effect sizes (Kover & Atwood, 2013), but small effect sizes indicate better matched groups. Note that in Table 2, the majority of the effect sizes are small, with the largest being .89 for the difference between nonverbal cognitive abilities of TD and DLD. In sum, taken together, both these analysis techniques suggest our groups are matched.

Materials and procedure

Children were tested by trained student research assistants at home or in schools, where they completed the narrative assessment, a nonverbal IQ test, and a test of receptive vocabulary. At home, parents were given questionnaires about their child's language learning history in L1 and L2 and their current language environment. A cultural broker or interpreter was present if the families desired so.

Table 2. Variables forming the basis of the matching criteria for the ASD, DLD, and TD groups

Background variable	Group			Group and pairwise comparisons
	ASD	DLD	TD	
Age (months)	85.2(16.1)	77.3(15.7)	79.1(6.8)	$\chi^2(2) = 2.89, p = .24$ TD-ASD: $d = 0.49$ TD-DLD: $d = 0.17$ ASD-DLD: $d = 0.51$
Nonverbal cognitive abilities (CMMS)	100.8(24.7)	97.7(11.4)	107.5(10.5)	$\chi^2(2) = 3.4, p = .18$ TD-ASD $d = -0.35$ TD-DLD $d = -0.89,$ ASD-DLD $d = -0.16,$
L2 abilities (PPVT)	84(12.1)	88.2(18.7)	90.2(20.6)	$\chi^2(2) = 1.32, p = .52$ TD-ASD $d = -0.37$ TD-DLD $d = -0.11$ ASD-DLD $d = .25$
L2 exposure in months (ALEQ)	55(15.8)	38 (17.9)	30.3(10.1)	$\chi^2(2) = 7.71, p = 0.02$ TD vs. ASD $W = 78.5, p = .007$ TD-ASD $d = 1.86$ TD-DLD $d = 1.00$ ASD-DLD $d = 0.53$
Richness of L2 environment (ALEQ)	0.76(.13)	0.64(.14)	0.65(.14)	$\chi^2(2) = 3.68, p = .16$ TD-ASD $d = 0.81$ TD-DLD $d = 0.88$ ASD-DLD $d = .07$

Note: CMMS, Columbia Mental Maturity Scales, standard mean = 100; PPVT = Peabody Picture Vocabulary Test, standard mean = 100; Richness of L2 environment, scores from 0 to 1.0.

The Edmonton Narrative Norms Instrument (ENNI; Schneider et al., 2005; http://www.rehabmed.ualberta.ca/spa/enni/about_the_enni.htm) was used to elicit narratives. The ENNI is a normed and standardized instrument that consists of two sets of three stories of increasing complexity, stories A1–A3 and B1–B3. Stories A1 and B1 contain a single episode, A2 and B2 contain two episodes, while A3 and B3 contain three episodes. Children are shown the picture books and asked to tell the stories while the experimenter sits in front of the child and cannot see the pictures. The stories produced by the children were then recorded, transcribed using the CHAT system (MacWhinney, 2000), and analyzed. The following macrostructure (story grammar) and microstructure (mean length of communicative unit [MLCU]),

syntactic complexity, number of different words/lexical diversity, and total number of words/story length) measures were examined in the children's stories.

Narrative macrostructure coding

Story grammar scores were calculated using rubrics specifically created for this study, rather than the ENNI scoring system. New scoring rubrics were created for two reasons. First, the ENNI manual contains scoring rubrics for only two out of six stories; so, four additional rubrics were created following the principles used for the two existing rubrics. Second, in the ENNI, reactions to story outcomes may include internal state terms such as *happy*, but also actions such as *say thank you*, behavioral manifestations of emotions such as *cry*, or even physical descriptions such as *wet*. As children with ASD have difficulties with perspective-taking, it is possible that they may produce fewer reactions, compared to actions that do not require perspective-taking abilities. Hence, in our scoring rubrics, we made a distinction between internal state terms produced as reactions, and actions or behavioral manifestations produced as the story grammar component of reactions (see also the Multilingual Assessment Instrument for Narratives [MAIN]; Gagarina et al., 2012). The scoring rubrics we created were used for scoring all six ENNI stories.

Each story was scored for the presence or absence of story grammar components by using the rubrics created for this study. The following story grammar components were scored: character introductions, settings (when and where the story events took place), initiating events (the event that sets off the story episode), internal responses (how characters respond to initiating events), internal plans (how characters plan on dealing with the initiating event), attempts (their attempts to so), outcomes (the results of their attempts), internal reactions (internal state terms produced as reactions to outcomes), and other reactions (other responses to outcomes such as actions or behavioral manifestations of emotions). The number of episodes ranged from one in stories A1 and B1, to three in the more complex stories A3 and B3. Similarly, the number of characters also differed, while stories A1 and B1 contained two characters, stories A2 and B2 introduced a third character, while stories A3 and B3 contained four characters. As stories A1–A3 are of increasing complexity, as are stories B1–B3, the maximum score possible was not identical for all stories (12 for stories A1 and B1, 24 for stories A2 and B2, and 36 for stories A3 and B3). Details on story grammar components with examples and instructions for scoring are given in the Appendix in Table A1.

Composite story grammar scores were calculated for each story (research question #1) by counting each story grammar component produced, yielding six composite story grammar scores, one for each story. Next, the number of each story grammar component produced across all six stories was counted (see research question #2). For example, we counted how many settings or initiating events the child included across all six stories. Note that for character introductions, we used a stringent scoring scheme in which only unambiguous introductions were counted. As such, introductions with pronouns were excluded as introducing characters with pronouns presupposes shared knowledge producing unambiguous characters would therefore require perspective-taking skills. Other story grammar components that would require perspective-taking skills are internal plans, internal responses, and

internal reactions. Because judgment is involved in scoring for story grammar components, 31% of the corpus was rescored by a separate research assistant. Comparisons of scoring for story grammar and story grammar components across stories yielded reliability of 82% and 85%, respectively. Discrepancies were settled through discussion and a final scoring was arrived at by consensus.

Narrative microstructure coding

(1) Utterance length: This refers to the mean utterance length in words across all stories and was calculated automatically by CLAN (MacWhinney, 2000) by looking at the MLCU. All utterances produced by the child, except for false starts, repetitions, and utterances not part of the storytelling were included. Higher scores reflect longer utterances/greater complexity. (2) Syntactic complexity: an index of syntactic complexity was calculated by dividing the number of independent and dependent clauses produced across all stories by the number of independent clauses produced. Higher scores mean the presence of more complex sentences. Fifty-five percent of the transcripts produced by the children with ASD were rescored for reliability by a separate research assistant. Comparisons of scoring for syntactic complexity yielded reliability of 87%. Comparisons for scoring for syntactic complexity in the other two bilingual groups yielded reliability of 98%. Any discrepancies were settled through discussion, and a final scoring was arrived at by consensus. (3) Lexical diversity: the number of unique word types used across all stories was calculated automatically by CLAN. (4) Story length: the number of word tokens used across all stories was calculated automatically by CLAN. This was used as a measure of productivity.

The Alberta Language Development Questionnaire (ALDeQ; Paradis et al., 2010; <https://www.ualberta.ca/linguistics/chescentre/questionnaires>). The ALDeQ is a parent questionnaire designed for use in linguistically and culturally diverse contexts. The ALDeQ includes sections that focus on (A) early milestones, (B) current abilities in the first language, (C) activity and behavior patterns shown by the child, and (D) family history of language and/or learning disabilities. The ALDeQ yields a total proportion score with a range from 0 to 1, as well as section scores. Lower scores on the ALDeQ are more typical of children with language disorder.

The Alberta Language Environment Questionnaire (ALEQ; Paradis, 2011; <https://www.ualberta.ca/linguistics/chescentre/questionnaires>). The ALEQ is a parent questionnaire with questions on language input factors, age, and family demographics. This questionnaire was administered to parents with the assistance of interpreters or cultural brokers. The ALEQ contains questions about the following topics: age of arrival in Canada, parents' self-rated proficiency in English, parent education, current language use by family members in the house (parents, other adults, siblings, and the target child), age at which the child started learning English in school, exposure to English measured in months (age of acquisition subtracted from the age at testing) as well as the richness of the English language environment. English language richness scores were calculated by examining the number of L2 enriching activities (such as book reading in English) the child was engaged in, as well as the frequency of these activities. A proportional score from 0 to 1 was calculated, with scores closer to 1 indicating richer English language environments.

The Columbia Mental Maturity Scales (CMMS; Burgemeister et al., 1972). The CMMS is a test of nonverbal intelligence in which children are shown patterns of increasing complexity and asked to identify the pattern that does not logically belong in a given sequence. Children who have a standard score greater than 80 score within the normal range on this test.

The Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997). The PPVT is a test of receptive vocabulary in which children are shown pictures and asked to identify the picture that corresponds to the word spoken by the experimenter. The PPVT has a standard score of 100, with the normal range being from 85 to 115.

Results

Analyses to address our three research questions were conducted using regression modeling. Linear regression was used to answer research questions 1 and 3 by using the `lm` function in R (R Core Team, 2017), with group (ASD, TD, or DLD) as the independent variable or fixed effect, and story grammar scores or microstructure scores as the dependent variable. Length of exposure to L2 English was also entered as a covariate fixed effect in the models to capture the variance in scores that might be due to differences in experience with L2 input, as our groups differed in their amount of L2 exposure. Therefore, there was a maximum of two fixed effects per model (group [categorical variable with three levels] and L2 exposure [continuous variable]). First, both fixed effects were entered, and the interpretation of the effect of group – our main variable of interest – is made in the results. If L2 exposure was also significant, this is also interpreted in the results. No significant interactions were found between the fixed effects. After selecting the optimal model, the fixed effects were examined for significance level: significant: $p < .05$, trend: $p < .08$ or nonsignificant $p > .08$). As our sample sizes were small, we have also reported trends ($p < 0.08$) for all research questions to indicate what might be of interest in future research with larger samples. Second, we used deviance comparison to arrive at the optimal parsimonious model for each dependent variable. That is, the AIC of the full model with both fixed effects was generated and compared with the reduced model with one fixed effect. The reduced model was chosen if the deviance (fit) was not improved by the addition of the second fixed effect (AIC function in R). The optimal models are reported in tables in the Appendix (see the foregoing sections of the number of each table). The ASD group was used as the reference level for group in the modeling. Thus, when group was significant, these models would show if ASD were different from TD or DLD. But, they did not allow us to compare TD and DLD. To do this, we ran models with parallel procedures to those described above, but with TD as the reference group in order to be able to know if DLD were different from TD. This was only done if group was significant. Note that doing so did not change the significance of group in the first set of models.

For our second research question, we used Poisson regression as we were examining count data. A series of Poisson regressions were fitted for each story grammar component using the `glmer` function in R (R Core Team, 2017). Group and L2 exposure were entered as fixed effects, and participant was entered as a random effect. Models were compared using the AIC function in R to ensure that optimal models

were selected for analysis. Again, the ASD group was used as the reference level for group in the modeling. However, to see if there were significant differences between the bilinguals with TD and the bilinguals with DLD, the TD group was used as the reference group in follow-up analyses.

Story grammar/macrostructure for each story

Significant group differences were found between the bilinguals with ASD and the bilinguals with TD for five of the six stories: A1 ($\beta = 4.32, t=2.60, p = .02$), A2 ($\beta = 4.98, t=3.74, p < .01$), A3 ($\beta = 5.24, t=2.60, p = .02$), B1 ($\beta = 2.16, t = 2.22, p = p < .04$), and B3 ($\beta = 4.23, t = 2.57, p = .02$). The bilinguals with DLD obtained higher scores than the bilinguals with ASD for only story B3 ($\beta = 4.16, t = 2.28, p = .03$). The bilingual DLD group obtained lower story grammar scores than the bilingual TD group for story A2 ($\beta = -3.97, t = -3.7, p = p < .01$), and a trend toward significance was found for story A3 ($\beta = -3.87, t = -1.93, p = 0.06$). Finally, while exposure was a significant predictor for story A2 ($\beta = 0.08, t = 2.43, p = .04$), no interaction was found between group and L2 exposure. These results have been summarized in Figure 1 and Table A3 (in the Appendix). Null results for story B2, for which neither group nor exposure predicted performance, have also been reported.

Story grammar/macrostructure components across stories

Group emerged as a significant predictor for character introductions, with both the TD ($\beta = 0.37, z = 2.71, p = .01$) and the DLD groups ($\beta = 0.28, z = 2.01, p = .04$) introducing more characters than the ASD group, who were more likely to introduce characters with pronouns. Significant group differences were found for initiating events with both the DLD ($\beta = -0.45, z = -2.52, p = .01$), and the ASD ($\beta = 0.62, z = 2.83, p = .005$) group being less likely to produce initiating events than the TD group. Similarly, the bilinguals with TD were more likely to include attempts ($\beta = 0.39, z = 2.42, p = .02$), and outcomes ($\beta = 0.47, z = 2.50, p = .01$) in their narratives than the bilingual ASD group. The DLD group produced fewer outcomes than the TD group ($\beta = -0.36, z = -2.33, p = .02$), and a trend toward significance was found for attempts ($\beta = 0.29, z = -1.89, p = .06$). The two clinical groups did not differ significantly from each other for any component except for the use of internal reactions: the bilingual DLD group produced more internal reactions than the bilingual ASD group ($\beta = 0.69, z = 2.28, p = .02$). Although the bilingual TD group did not differ from the bilingual ASD group for use of internal reactions, a trend toward significance was noted ($\beta = 0.53, z = 1.71, p = .09$). Finally, no interactions were found between group and L2 exposure, and L2 exposure emerged as a significant predictor only for initiating events ($\beta = 0.19, z = 2.06, p = .04$) and outcomes ($\beta = 0.16, z = 2.13, p = .03$). These results have been summarized in Figure 2 and Table A4 (in the Appendix).

No group differences were found for settings, internal responses, internal plans, and other reactions (non-internal). All groups of children were equally likely to produce these components. These null findings have also been reported in Table A4.

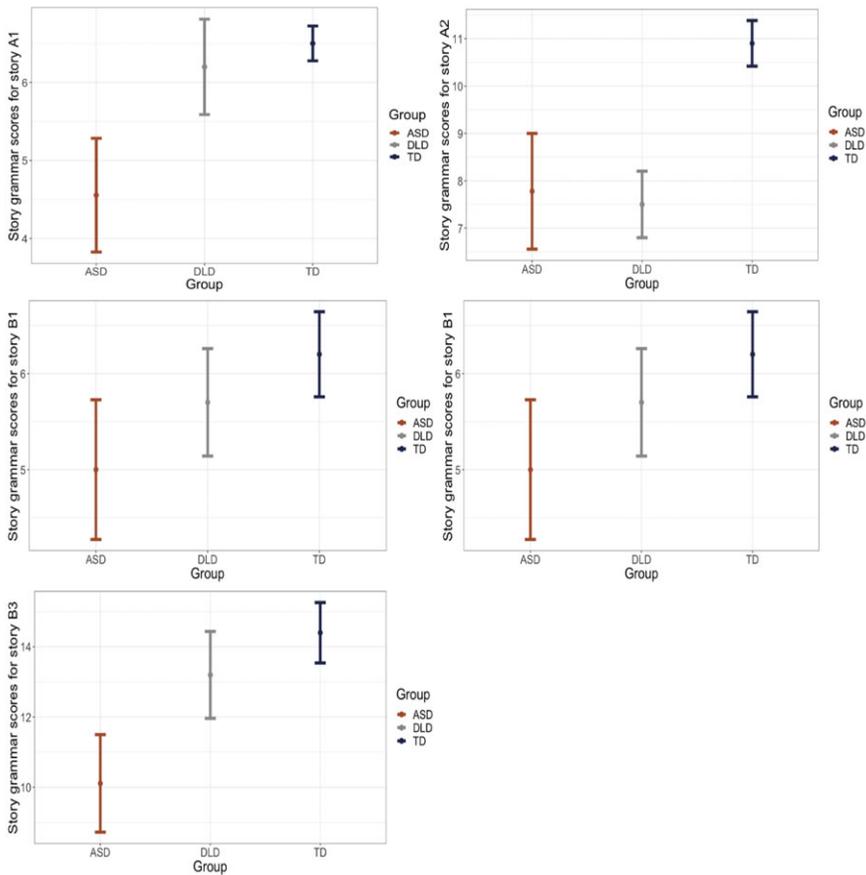


Figure 1. Score ranges for story grammar scores. Note that the scales differ across stories, as more complex stories have higher maximum scores.

Microstructure components across stories

Significant group differences emerged for utterance length and syntactic complexity. The optimal model for utterance length included group but not L2 exposure. Children with ASD produced shorter utterances than the children with TD ($\beta = 4.42, t = 2.48, p = .02$), but no differences were found with the DLD group. While the children with DLD did not differ significantly from the children with TD, there was still a trend toward significance ($\beta = -3.2, t = -1.18, p = .08$). Similarly, the ASD group used less complex syntax than the TD group ($\beta = 5.85, t = 2.91, p < .05$), but not the DLD group. Both clinical groups differed significantly from the TD controls for the use of complex syntax. L2 Exposure was not a significant predictor, nor were any interactions found between group and exposure. These results have been summarized in Figure 3 and Table A5 (in the Appendix).

In contrast to the above results, no group differences were found for lexical diversity or story length (productivity). These null results have also been included in Table A5.

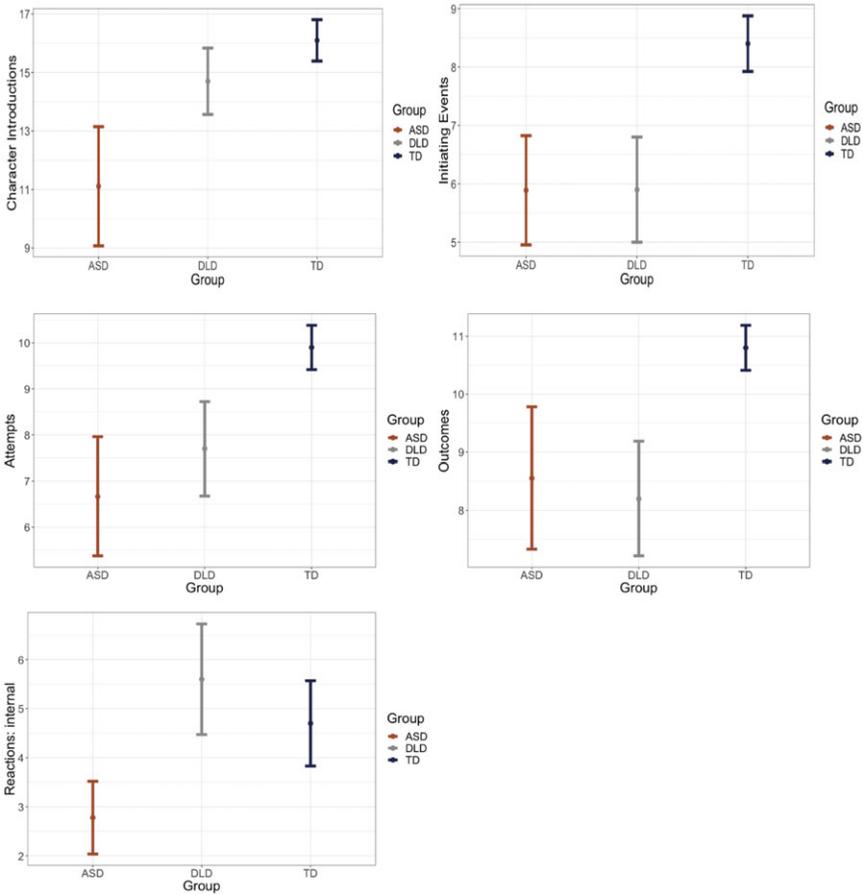


Figure 2. Score ranges for story grammar components.

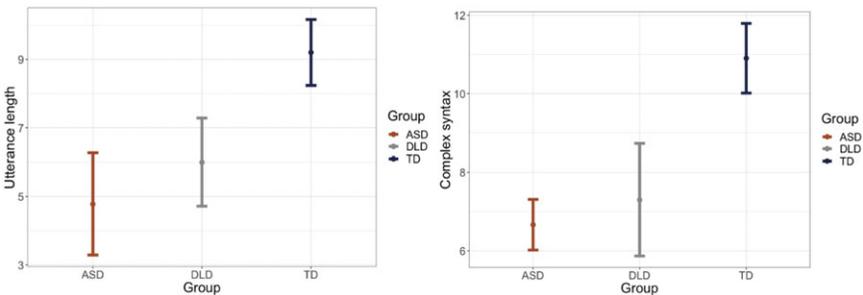


Figure 3. Score ranges for microstructure components.

Discussion

A greater understanding of the language and communication development of bilingual children with ASD is essential for clinical practice in diverse societies. The existing research on bilingual children with ASD has mainly focused on the early years and on comparisons between monolinguals and bilinguals with ASD to determine whether dual language learning would be too burdensome for children with ASD. Therefore, there is limited research on dual language development in school-aged children with ASD and how they compare to their bilingual TD peers. In addition to TD peers, bilingual children with DLD are an interesting comparison group for bilinguals with ASD because, while the children with ASD and DLD show some convergence in structural language difficulties, they often diverge for discourse-pragmatic language abilities. Accordingly, the objective of this study was to examine the narrative abilities of bilingual school age children with ASD, as referenced to those of TD bilinguals, as well as to bilinguals with DLD, in order to address these gaps in knowledge. Because narrative macrostructure taps into discourse-pragmatics skills while narrative microstructure taps into structural language skills, performance on a narrative task could differentiate between the linguistic profile of bilinguals with ASD from those of bilinguals with TD and with DLD. Before discussing the results, it is important to specify that, as a spectrum disorder, autism is characterized by great variation. Autistic individuals show great variation both in their core symptoms as well as in language and intelligence, with language abilities ranging from seemingly intact structural language to never acquiring functioning language (Tager-Flusberg, 2004; Tager-Flusberg et al., 2005). No study on autism can account for the spectrum. This study is no exception and therefore, we cannot draw conclusions about the entire spectrum of autistic individuals.

Macrostructure abilities in bilingual ASD: story grammar/coherence for each story

For our first research question, we examined story grammar scores for the individual stories produced by the bilingual children with TD, ASD, and DLD.

We predicted that the bilinguals with ASD would have lower story grammar scores than the bilinguals with TD, that is, that they would produce less coherent stories with fewer story grammar components and possibly be similar to the bilinguals with DLD. Concerning the ASD and TD comparisons, our predictions were largely supported by the data, as significant group differences were found between the bilinguals with ASD and the bilinguals with TD for five out of six stories. Why not in B2? This could just be an artifact of small sample sizes. While most studies on monolinguals with TD and with ASD have found significant differences for story grammar (e.g., Norbury et al., 2014), null results have also been reported (e.g., Young et al., 2005). Furthermore, a key difference between our study and the previous research on ASD and narratives was that we examined six different story narratives, while most studies have used a single story narrative task (e.g., Diehl et al., 2006; Norbury et al., 2014; Tager-Flusberg, 1995; but see Colozzo et al., 2015), which means our study included more samples of narrative abilities from each child than many previous studies. Because of this sampling difference, some null effects could be expected.

Regarding the second part of our prediction, our analyses found that the bilingual ASD and DLD groups showed no differences in their story grammar scores for the majority of the stories (five out of six). Our results therefore align with some previous research on cross-disorder comparisons with monolinguals (Norbury & Bishop, 2003; Norbury et al., 2014; but see Manolitsi & Botting, 2011). While the DLD group was only different from the ASD group (with DLD obtaining higher scores) for their scores on one story (B3), they were also only significantly lower than the TD group for one story A2, and with a trend toward significance for A3. This absence of a clear difference between DLD and TD echoes the inconsistent findings from the bilingual research where not all studies have found macrostructure differences between bilinguals with DLD and their TD age peers (Boerma et al, 2016 and Govindarajan & Paradis, 2019 vs. Altman et al., 2016 and Iluz-Cohen & Walters, 2012).

In sum, our results for story grammar scores dovetail with those from the monolingual literature (e.g., Baixauli et al., 2016) and indicate that: (1) producing a well-structured narrative is a challenge for children with ASD who produced less coherent stories than children with TD, and (2) that DLD and ASD pattern similarly for macrostructure when story grammar scores are examined. Being bilingual does not change this profile. However, looking at only story grammar scores may mask important differences between the narratives produced by groups, as certain story grammar components may be particularly challenging for children with ASD. Hence, we formulated our second research question focusing on individual story grammar components.

Macrostructure abilities in bilingual ASD: individual story grammar components

For our second research question, we examined whether the frequency of use of individual story grammar components differentiated the groups. This kind of a fine-grained analysis is less common in either the monolingual and the bilingual research on ASD and narratives. While some previous studies have reported differences between monolingual TD and ASD groups for individual components such as outcomes, internal responses, or introducing characters (Banney et al., 2015; Goldman, 2008; Tager-Flusberg, 1995), looking at individual story grammar components has not been explicitly framed as a research question in previous studies. To address our second question, we counted the number of story grammar components produced by children across all six stories. For the core narrative components, we predicted that the bilingual ASD group would produce fewer outcomes than the bilingual TD group. Next, we predicted significant differences between the bilingual ASD group and the other two bilingual groups for narrative components relying on perspective-taking abilities: unambiguous character introductions, internal plans, internal responses, and reactions.

The bilinguals with ASD included fewer outcomes than the bilinguals with TD, in line with our predictions and prior research with monolinguals (Banney et al., 2015; Goldman, 2008; Tager-Flusberg, 1995). When it came to all the core narrative components, initiating events, attempts, and outcomes, both clinical groups patterned similarly, and they both had lower scores than the TD bilingual group (for attempts: trend only for DLD). Regarding perspective-taking skills, children with ASD introduced fewer characters in their stories than both the children with TD and the

children with DLD, in line with our prediction. (Recall that we used a stringent scoring scheme for character introductions where introductions with a pronoun, which were frequent in the bilingual ASD group, were not counted.) Partly consistent with our prediction, significant group differences were found for the number of internal reactions, with the bilingual ASD group producing significantly fewer internal reactions than the bilingual DLD group. The bilingual ASD group did not differ significantly from the TD group, however; although, a trend toward significance was found, suggesting that a significant result may have emerged with larger participant groups. Finally, we did not find any differences for two other components that required perspective-taking skills, internal plans, and internal responses (*cf.* Banney et al, 2015). However, regardless of group, children produced very few internal plans or responses, which might be indicative of developmental trends in the production of narratives (Berman & Slobin, 1994) as these components are included more often in the narratives of older children and adolescents.

In sum, our results suggest that a fine-grained analysis of individual story grammar components may reveal group differences that are masked by composite story grammar scores. When looking at only composite story grammar scores, both bilinguals with ASD and bilinguals with DLD presented similar profiles. However, breaking down story grammar scores revealed some differences between these two groups. Although both clinical groups in our study were similar in their core narrative components, components requiring perspective-taking abilities were more challenging for ASD than for DLD. These components may be particularly useful for distinguishing between children with ASD and children with DLD.

Microstructure abilities in bilinguals with ASD: productivity and complexity

For our third research question, we examined the children's performance with the following microstructure components: lexical diversity, story length (productivity), and syntactic complexity and utterance length (complexity). We predicted that the bilingual ASD group would differ from the TD group on both the measures of productivity and the measures of complexity. In addition, we expected the bilinguals with ASD to pattern similarly to the bilinguals with DLD (Norbury & Bishop, 2003; Norbury et al., 2014). These predictions were partially supported. We found differences between ASD and TD for the measures of complexity (utterance length and complex syntax), but not for the measures of productivity (lexical diversity and story length; see below). The bilinguals with ASD in our study produced shorter utterances and used less complex syntax than the bilinguals with TD, consistent with findings the monolingual literature (Norbury & Bishop, 2003; Norbury et al., 2014). While measures of productivity have been found to differentiate between children with ASD and children with TD, with a moderate effect size (Baixauli et al., 2016), no group differences were found in this study. Bilinguals in all three groups produced stories of similar length and used a similarly diverse vocabulary. Differences between children with ASD and their TD peers for productivity measures of length and lexical diversity may be less apparent when they are producing stories with structured supports (Losh & Capps, 2003); for example, the participants in our study were constrained in their storytelling by the picture sequences. Productivity differences might be more apparent on a less structured

narrative task or spontaneous conversation. Furthermore, all three groups in this study were matched on receptive vocabulary at the outset, and this might have contributed to the absence of differences for expressive lexical diversity.

As predicted, no differences were found between the bilingual ASD and the bilingual DLD group for either the measures of productivity or complexity. Like the bilinguals with ASD, the bilingual DLD group also used significantly less complex syntax than the TD group, and a trend for the same pattern emerged for utterance length. As mentioned in the introduction, while ASD is characterized by deficits in discourse pragmatics, the presence and extent of structural language difficulties varies in ASD. Our findings add to the increasing evidence from the monolingual research that, in addition to the well-documented difficulties with pragmatics, some children with ASD also display deficits with structural language skills. It must be noted that shorter utterances and reduced syntactic complexity can be, in part, attributable to limited narrative abilities. Further research is required to determine whether children with ASD show reduced complexity because of narrative abilities, or because of core structural language deficits.

Narrative difficulties in bilingual ASD: L2 exposure

Recall that our groups were matched for L2 vocabulary abilities but were not matched for L2 exposure, with the ASD group having significantly more L2 exposure than the TD group. Therefore, we entered length of L2 English exposure in the models to capture any variance it might have on children's abilities with macro- and microstructure components. Exposure to L2 only emerged as a significant predictor for the global story grammar scores for story A2, and for the components *initiating events* and *outcomes* across stories. Furthermore, no interactions were found between exposure and group. As such, the bilinguals with ASD did not seem to benefit from their additional L2 exposure in their performance on the narrative task.

Conclusions and limitations

This study contributes to the emergent body of research on bilinguals with ASD and narratives and is the first to conduct a cross-disorder comparison with bilingual populations with ASD and DLD. This study is consistent with findings of monolingual children with ASD and shows that school age children with ASD, whether they are monolingual or bilingual, show deficits in both narrative macrostructure and microstructure. Overall, in comparison to TD controls, these bilingual children with ASD produced stories with reduced story content and used less complex syntax and shorter utterances. In terms of global story grammar scores and microstructure measures, the bilinguals with ASD mainly patterned similarly to the bilinguals with DLD and differed from the bilinguals with TD. Finally, our results with these bilinguals align with much evidence from the monolingual research indicating that difficulties with structural language can extend to an overlap between some children with ASD and children with DLD. While both clinical groups overlapped on structural language skills and global story grammar scores (except for story B3), differences were found for individual narrative macro-structure components that require perspective-taking abilities, such as character introductions, and internal reactions

to story outcomes: these were a relative weakness for ASD but a relative strength for DLD. Thus, this study found that the linguistic profile of these bilingual children with ASD and DLD are both overlapping and distinct, and that components requiring perspective-taking skills are particularly useful in distinguishing between these two clinical groups.

Finally, we would like to acknowledge certain limitations to our study. This study had a small sample size, which limits the generalization of our findings. Thus, there is a need for additional cross-disorder comparisons with larger bilingual groups to ascertain if the results of this study are borne out in others. Next, as mentioned earlier, autism is a *spectrum* disorder and no study on autism can claim to be generalizable to the entire population of autistic individuals. We were also unable to include language measures for group matching beyond L2 receptive vocabulary as measured by the PPVT. Thus, there is a need for further cross-disorder bilingual comparisons involving groups matched not only on age, receptive vocabulary or intelligence, but also on expressive language skills. Finally, the small sample size limited how many fixed effects could be entered in our model, thus limiting our ability to comprehensively analyze the role of additional individual difference factors in bilingual development with ASD – a worthwhile goal for future research.

References

- 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.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.
- Baixauli, I., Colomer, C., Rosello, B., & Miranda, A. (2016). Narratives of children with high-functioning autism spectrum disorder: A meta-analysis. *Research in Developmental Disabilities*, *59*, 234–254. <https://doi.org/10.1016/j.ridd.2016.09.007>
- Baldimtsi, E., Peristeri, E., Tsimpli, I.M., & Nicolopoulou, A. (2016). Bilingual children with high functioning autism spectrum disorder: Evidence from oral narratives and non-verbal executive function tasks. In J Scott and D Waughtal (Eds.), *Proceedings of the 40th annual Boston University conference on language development* (pp. 18–31). Retrieved from <http://www.lingref.com/buclid/40/BUCLD40-02.pdf>
- Banney, M. R., Harper-Hill, K., & Arnott, W.L. (2015). The Autism diagnostic observation schedule and narrative assessment: Evidence for specific narrative impairments in autism spectrum disorders. *International Journal of Speech-Language Pathology*, *17*(2), 159–171. <https://doi.org/10.3109/17549507.2014.977348>
- Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, *11*, 1–29.
- Berman, R.A., & Slobin, D. I. (1994). *Relating events in narrative: A cross-linguistic developmental study*. Hillsdale, NJ: L. Erlbaum.
- Bishop, D., & Donlan, C. (2005). The role of syntax in encoding and recall of pictorial narratives: Evidence from specific language impairment. *British Journal of Developmental Psychology*, *23*, 25–46. <https://doi.org/10.1348/026151004x20685>
- 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. <https://doi.org/10.1111/1460-6984.12234a>
- Burgemeister, B., Hollander Blum, L., & Lorge, I. (1972). *Columbia Mental Maturity Scale*. San Antonio, TX: Psychological Corporation.

- Capps, L., Losh, M., & Thurber, C. (2000). "The frog ate the bug and made his mouth sad": Narrative competence in children with autism. *Journal of Abnormal Child Psychology*, 28(2), 193–204.
- Colozzo, P., Gillam, R.B., Wood, M., Schnell, R., & Johnston, J.R. (2011). Content and form in the narratives of children with Specific Language Impairment. *Journal of Speech, Language and Hearing Research*, 54(6), 1609–1627. [https://doi.org/10.1044/1092-4388\(2011/10-0247\)](https://doi.org/10.1044/1092-4388(2011/10-0247))
- Diehl, J., Bennetto, C., & Young, E. (2006). Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *Journal of Abnormal Child Psychology & Health*, 34, 87–102.
- Dodwell, K., & Bavin, E. L. (2008). Children with specific language impairment: An investigation of their narratives and memory. *International Journal of Language & Communication Disorders*, 43(2), 201–218. <https://doi.org/10.1080/13682820701366147>
- Dunn, L., & Dunn, L. (1997). *Peabody picture vocabulary test—Third edition*. Circle Pines, MN: AGS.
- Durrleman, S., & Delage, H. (2016). Autism spectrum disorder and specific language impairment: Overlaps in syntactic profiles. *Language Acquisition*, 23(4), 361–386. <https://doi.org/10.1080/10489223.2016.1179741>
- Ellis Weismer, S., & Kover, S. T. (2015). Preschool language variation, growth, and predictors in children on the autism spectrum. *Journal of Child Psychology and Psychiatry*, 56(12), 1327–1337. <https://doi.org/10.1111/jcpp.12406>
- Engberg-Pedersen, E., & Christensen, R. (2017). Mental states and activities in Danish narratives: Children with autism and children with language impairment. *Journal of Child Language*, 44(5), 1192–1217. <https://doi.org/10.1017/S0305000916000507>
- Fey, M., Catts, H., Proctor-Williams, K., Tomblin, J., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. *Journal of Speech, Language & Hearing Research*, 47(6), 1301–1318.
- Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūniene, I., Bohnacker, U., & Walters, J. (2012). MAIN: Multilingual assessment instrument for narratives. *ZAS Papers in Linguistics*, 56, 1–140.
- Gerenser, J., & Lopez, K. (2017). Autism spectrum disorders. In Schwartz, R. (Ed.), *Handbook of Child Language Disorders* (pp. 82–108). New York, NY: Routledge.
- 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. <https://doi.org/10.1016/j.jcomdis.2018.10.001>
- Hoang, H., Gonzalez-Barrero, A.M., & Nadig, A. (2018). Narrative skills of bilingual children with autism spectrum disorder. *Discours [En ligne]*, 23, mis en ligne le 21 décembre 2018, consulté le 05 décembre 2019. <http://journals.openedition.org/discours/9856>; <https://doi.org/10.4000/discours.9856>
- 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, 58–74.
- Goldman, S. (2008). Brief report: Narratives of personal events in children with autism and developmental language disorders: Unshared memories. *Journal of Autism and Developmental Disorders*, 38, 1982–1988.
- Colozzo, P., Morris, H., & Miranda, P. (2015). Narrative production in children with autism spectrum disorder and specific language impairment. *Canadian Journal of Speech-Language Pathology & Audiology*, 39(4), 316–332.
- Justice, L. M., Bowles, R. P., Kaderavek, J. N., Ukrainetz, T. A., Eisenberg, S. L., & Gillam, R. B. (2006). The index of narrative microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech Language Pathology*, 15(2), 177–191. [https://doi.org/10.1044/1058-0360\(2006/017\)](https://doi.org/10.1044/1058-0360(2006/017))
- Kauschke, C., van der Beek, B., & Kamp-Becker, I. (2016). Narratives of girls and boys with autism spectrum disorders: Gender differences in narrative competence and internal state language. *Journal of Autism and Developmental Disorders*, 46, 840–852. <https://doi.org/10.1007/s10803-015-2620-5>
- Kay-Raining Bird, E., Genesee, F., & Verhoeven, L. (2016). Bilingualism in children with developmental disorders: A narrative review. *Journal of Communication Disorders*, 63, 1–14.
- Kjelgaard, M.M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, 16(2-3), 287–308. <https://doi.org/10.1080/01690960042000058>

- Kover, S. T., & Atwood, A. K.** (2013). Establishing equivalence: Methodological progress in group matching design and analysis. *American Journal on Intellectual and Developmental Disabilities*, **118**(1), 3–15. <https://doi.org/10.1352/1944-7558-118.1.3>
- Lauro, J., Core, C., & Hoff, E.** (2020). Explaining individual differences in trajectories of simultaneous bilingual development: Contributions of child and environmental factors. *Child Development*, **91**(6), 2063–2082.
- Leonard, L.** (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.
- Losh, M., & Capps, L.** (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of Autism and Developmental Disorders*, **33**(3), 239–251.
- MacWhinney, B.** (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.
- Mäkinen, L., Loukusa, S., Leinonen, E., Moilanen, I., Ebeling, H., & Kunnari, S.** (2014). Characteristics of narrative language in autism spectrum disorder: Evidence from the Finnish. *Research in Autism Spectrum Disorders*, **8**(8), 987–996.
- Manolitsi, M., & Botting, N.** (2011). Language abilities in children with autism and language impairment: Using narrative as a additional source of clinical information. *Child Language Teaching and Therapy*, **27**(1), 39–55.
- Meir, N., & Novogrodsky, R.** (2019). Syntactic abilities and verbal memory in monolingual and bilingual children with High Functioning Autism (HFA). *First Language*. <https://doi.org/10.1177/0142723719849981>
- Norbury, C., Gemmell, T., & Paul, R.** (2014). Pragmatics abilities in narrative production: A cross-disorder comparison. *Journal of Child Language*, **41**(3), 485–510. <https://doi.org/10.1017/S030500091300007X>
- Norbury, C. F., & Bishop, D. V.** (2003). Narrative skills of children with communication impairments. *International Journal of Language and Communication Disorders*, **38**, 287–313.
- Oetting, J. B., & Hadley, P. A.** (2017). Morphosyntax in child language disorders. In R. G. Schwartz (Ed.), *The handbook of child language disorders* (pp. 365–391). New York, NY: Psychological Press
- Ohashi, J. K., Miranda, P., Marinova-Todd, S.H., Hambly, C., Fombonne, E., Szatmari, P., & The Pathways in ASD Study Team.** (2012). Comparing early language development in monolingual- and bilingual- exposed young children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, **6**(2), 890–897.
- Paradis, J.** (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech and Hearing Services in the Schools*, **36**, 172–187.
- Paradis, J.** (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, **1**, 213–237.
- Paradis, J.** (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language and Hearing Research*, **58**, 171–182.
- Paradis, J.** (2019). English second language acquisition from early childhood to adulthood: The role of age, first language, cognitive and input factors. In M. Brown & B. Dailey (Eds.), *Proceedings of the 43rd Boston University conference on language development* (pp. 11–26). Somerville, MA: Cascadilla Press.
- Paradis, J., Emmerzael, K., & Sorenson Duncan, T.** (2010) Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, **43**, 474–497.
- Paradis, J., Genesee, F., & Crago, M.** (2021). *Dual language development and disorders: A handbook on bilingualism and second language learning* (3rd Ed.). Baltimore, MD: Brookes.
- Paradis, J., & Govindarajan, K.** (2018). Bilingualism and children with language and communication disorders. In D. Miller, F. Bayram, J. Rothman and L. Serratrice (Eds.), *Bilingual language and Cognition: The state of the science across its subfields. [Sibil 54]* (pp. 347–370). Amsterdam: John Benjamin.
- Paradis, J., Govindarajan, K., & Hernandez, K.** (2018). Bilingual development in children with autism spectrum disorder from newcomer families. *Bilingual & ASD Report*, 1–12. <https://doi.org/10.7939/R31V5BT9X>
- 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.
- Peristeri, E., Baldimtsi, E., Adreou, M., & Tsimpli, I. M.** (2020). The impact of bilingualism on the narrative ability and the executive functions of children with autism spectrum disorders. *Journal of Communication Disorders*, **85**, 105999.

- Petersen, J. M., Marinova-Todd, S. H., & Mirenda, P. (2012). Brief report: An exploratory study of lexical skills in bilingual children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *42*(7), 1499–1503.
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Reetzke, R., Zou, X. K., Sheng, L., & Katsos, N. (2015). Communicative development in bilingually exposed Chinese children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, *58*, 813–835.
- 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*, 229–247.
- Roberts, J., Rice, M. L., & Tager-Flusberg, H. (2004). Tense marking in children with autism. *Applied Psycholinguistics*, *25*, 429–448. <https://doi.org/10.1017/S0142716404001201>
- Rumpf, A.-L., Kamp-Becker, I., Becker, K., & Kauschke, C. (2012). Narrative competence and internal state language of children with Asperger Syndrome and ADHD. *Research in Developmental Disabilities*, *33*, 1395–1407. <https://doi.org/10.1016/j.ridd.2012.03.007>
- Schneider, P. (1996). Effects of pictures versus orally presented stories on story retellings by children with language impairment. *American Journal of Speech-Language Pathology*, *5*, 86–96.
- Schneider, P., Hayward, D., & Dubé, R. V. (2006) Storytelling from pictures using the Edmonton narrative norms instrument. *Canadian Journal of Speech-Language Pathology and Audiology*, *30*, 224–238.
- Schneider, P. V., Dubé, R., & Hayward, D. (2005). The Edmonton narrative norms instrument. Retrieved June 3 2017, from www.rehabresearch.ualberta.ca/enni
- Schwartz, R. (2017). Specific language impairment. In R. Schwartz (Ed.), *Handbook of child language disorders* (pp. 3–43). New York: Psychology Press.
- Siller, M., Swanson, M. R., Serlin, G., & Teachworth, A. G. (2014). Internal state language in the storybook narratives of children with and without autism spectrum disorder: Investigating relations to theory of mind abilities. *Research in Autism Spectrum Disorders*, *8*(5), 589–596. <https://doi.org/10.1016/j.rasd.2014.02.002>
- Smith Gabig, C. (2008). Verbal working memory and story retelling in school-age children with autism. *Language, Speech, and Hearing Services in Schools*, *39*, 498–511.
- Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing* (pp. 53–102). Norwood, NJ: Ablex.
- Suh, J., Eigsti, I. M., Naigles, L., Barton, M., Kelley, E., & Fein, D. (2014). Narrative performance of optimal outcome children and adolescents with a history of an autism spectrum disorder (ASD). *Journal of Autism and Developmental Disorders*, *44*, 1681–1694. <https://doi.org/10.1007/s10803-014-2042-9>
- Tager-Flusberg, H. (1995). ‘Once upon a rabbit’: Stories narrated by autistic children. *British Journal of Developmental Psychology*, *13*, 45–59.
- Tager-Flusberg, H. (2004). Strategies for conducting research on language in autism. *Journal of Autism and Developmental Disorders*, *34*(1), 75–80. <https://doi.org/10.1023/b:jadd.0000018077.64617.5a>
- Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London, Series B* *358*, 303–314.
- Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders: Diagnosis, development, neurobiology, and behavior* (pp. 335–364). New York, NY: John Wiley & Sons Inc.
- Tager-Flusberg, H., & Sullivan, K. (1995). Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics*, *16*, 241–256.
- Torng, P. C., & Sah, W. H. (2020). Narrative abilities of Mandarin-speaking children with and without specific language impairment: Macrostructure and microstructure. *Clinical Linguistics and Phonology*, *34*(5), 453–478. <https://doi.org/10.1080/02699206.2019.1655097>
- Tsimpli, I. M., Peristeri, I. E., & Andreou, M. (2016). Narrative production in monolingual and bilingual children with specific language impairment. *Applied Psycholinguistics*, *37*(1), 195–216.
- Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 103–121). Washington, DC: The American Psychological Association.

- Valicenti-McDermott, M., Tarshis, N., Schouls, M., Galdston, M., Hottinger, K., Seijo, R., Shulman, L., & Shinnar, S. (2013). Language differences between monolingual English and bilingual English-Spanish young children with autism spectrum disorders. *Journal of Child Neurology*, *28*(7), 945–948. <https://doi.org/10.1177/0883073812453204>
- Wittke, K., Mastergeorge, A., Ozonoff, S., Rogers, S., & Naigles, L. (2017) Grammatical impairment and autism spectrum disorder: Exploring language phenotypes beyond standardized testing. *Frontiers in Language Sciences*. <https://doi.org/10.3389/fpsyg.2017.00532>
- Young, E. C., Diehl, J. J., Morris, D., Hyman, S. L., & Bennetto, L. (2005). The use of two language tests to identify pragmatic language problems in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools*, *36*(1), 62–72.

Appendix A

Table A1. Story Grammar Scoring Rubric

Story grammar component	Description	Example
Characters	<p>One point for introducing a character clearly. Introductions with pronouns were not counted.</p> <p>The number of characters in the story determined how many points children could score. Stories A1 and B1 contained two characters, stories A2 and B2 contained three characters, and stories A3 and B3 contained four characters.</p>	<p><i>once there was a giraffe and a elephant playing with one or three balls</i> (Child 14, ASD, L1 Spanish, 6;5, 60 months of L2 exposure)</p>
Setting	<p>One point for providing information about the setting.</p>	<p><i>After four months, it was a July and they went to the sandbox</i> (Child 5, ASD, L1 Mandarin, 8;0, 60 m months of L2 English Exposure)</p>
Initiating events	<p>One point for mentioning the initiating event that sets the story episode in motion.</p> <p>The number of initiating events possible ranged from one to three depending on story complexity and the number of episodes.</p>	<p><i>and then he dropped it in the water by accident</i> (Child 6, ASD, L1 Mandarin, 9;6, 71 months of L2 English exposure)</p>
Internal responses	<p>One point for mentioning how characters reacted to the initiating event.</p> <p>Depending on story complexity, a child could score from one to three points.</p>	<p><i>And dog got mad at the rabbit</i> (Child 70, TD, L1 Farsi, 6;0, 37 months of L2 English exposure)</p>
Internal plans	<p>One point for mentioning how characters planned to deal with the initiating event.</p> <p>Depending on story complexity, a child could score from one to three points.</p>	<p><i>Failure elephant decided to run</i> (Child 5, ASD, L1 Mandarin, 8;0, 60 m months of L2 English Exposure)</p>

(Continued)

Table A1. (Continued)

Story grammar component	Description	Example
Attempts	One point for indicating how characters attempted to attain their goal. Depending on story complexity, a child could score from one to three points.	<i>and then he <u>tries</u> to get it out.</i> (Child 6, ASD, L1 Mandarin, 9;6, 71 months of L2 English exposure)
Outcomes	One point for indicating the outcome or the consequence of the attempt. Depending on story complexity, a child could score from one to three points.	<i>and then he got it out and give it back to the giraffe.</i> (Child 6, ASD, L1 Mandarin, 9;6, 71 months of L2 English exposure)
Reactions to outcomes: internal	How characters reacted to the outcomes. Only internal state terms were counted. The number of story episodes and the number of characters determined the number of points children could score (two to nine).	<i>the giraffe is so <u>happy</u> that he got his toy back.</i> (Child 6, ASD, L1 Mandarin, 9;6, 71 months of L2 English exposure)
Reactions to outcomes: others	How characters reacted to story outcomes. Actions (physical and verbal) as well as manifestations of emotions such as <i>cry</i> were included. The number of story episodes and the number of characters determined the number of points children could score (two to nine).	<i>he <u>said</u>, thank you</i> (Child 14, ASD, L1 Spanish, 6;5, 60 months of L2 exposure)

Table A2. L1 backgrounds of participants in each group

Group	L1	Number of speakers
ASD	Mandarin	4
	Spanish	3
	Cantonese	1
	Arabic	1
DLD	Cantonese	3
	Vietnamese	1
	Spanish	4
	Urdu	1
	Arabic	1
TD	Arabic	4
	Farsi	1
	Cantonese	1
	Mandarin	1
	Spanish	1
	Urdu	1
	Punjabi	1

Table A3. Model results for story grammar scores

	Predictor	Estimate	SE	t	p
A1	Intercept	4.56	0.57	7.98	<.001***
	Group TD	1.94	0.79	2.47	0.02*
	Group DLD	1.64	0.79	2.09	0.05.
A2	Intercept	3.64	1.88	1.94	0.06
	Group TD	4.98	1.33	3.74	0.001***
	Group DLD	1.00	1.21	0.83	0.42
	L2 Exposure	0.08	0.03	2.42	0.02*
A3	Intercept	11.56	1.46	7.90	<.001***
	Group TD	5.24	2.02	2.60	0.02*
	Group DLD	1.84	2.02	0.92	0.37
B1	Intercept	2.87	1.38	2.09	0.05
	Group TD	2.16	0.97	2.22	0.04*
	Group DLD	1.36	0.88	1.54	0.14
	L2 Exposure	0.04	0.02	1.71	0.010
B2	Intercept	5.28	1.79	2.94	0.007**
	Group TD	2.01	1.26	1.59	0.13
	Group DLD	0.55	1.52	0.48	0.63
	L2 Exposure	0.03	0.03	1.13	0.27
B2	Intercept	7.11	0.76	9.40	<.001***
	Group TD	1.19	1.04	1.14	0.27
	Group DLD	-0.01	1.0	-0.01	0.99
B3	Intercept	10.11	1.21	8.35	<.001***
	Group TD	4.29	1.67	2.57	0.02*
	Group DLD	3.09	1.67	1.86	0.08.

Note: For task, the bilingual ASD group was the reference level. Exposure = length of exposure to English measured in months; * =significance code when $p < .05$; ** =significance code when $p < .01$; *** =significance code when $p < .001$; . =trend toward significance.

Table A4. Model results for story grammar components

	Predictor	Estimate	SE	t	p
Characters	Intercept	2.40	0.11	22.16	<.001***
	Group TD	0.37	0.14	2.71	0.007**
	Group DLD	0.28	0.14	2.01	0.04*
Initiating events	Intercept	1.61	0.16	9.95	<.001***
	Group TD	0.62	0.22	2.83	0.004**
	Group DLD	0.18	0.21	0.86	0.39
	Exposure	0.19	0.09	2.01	0.04*
Attempts	Intercept	1.89	0.13	14.69	<.001***
	Group TD	0.39	0.16	2.42	0.02*
	Group DLD	0.14	0.17	0.84	0.40
Outcomes	Intercept	2.00	0.13	14.89	<.001***
	Group TD	0.47	0.19	2.50	0.01*
	Group DLD	0.11	0.17	0.65	0.52
	Exposure	0.16	0.08	2.13	0.03*
Reactions: internal	Intercept	0.93	0.25	3.78	<.001***
	Group TD	0.53	0.31	1.71	0.08.
	Group DLD	0.69	0.31	2.28	0.02*
Settings	Intercept	4.67	0.50	9.30	<.001***
	Group TD	0.83	0.69	1.20	0.24
	Group DLD	-0.27	0.69	-0.39	0.70
Internal responses	Intercept	1.22	0.44	2.79	0.01*
	Group TD	-0.12	0.60	-0.20	0.84
	Group DLD	0.28	0.60	0.46	0.64
Internal plans	Intercept	0.89	0.39	2.24	0.03*
	Group TD	0.91	0.55	1.67	0.11
	Group DLD	-0.29	0.55	-0.52	0.60
Reactions: others	Intercept	4.11	0.84	4.89	<.001***
	Group TD	0.59	1.56	0.51	0.62
	Group DLD	0.19	1.56	0.16	0.83

Note: For task, the bilingual ASD group was the reference level. Exposure = length of exposure to English measured in months; * = significance code when $p < .05$; ** = significance code when $p < .01$; *** = significance code when $p < 0.001$.

Table A5. Model results for microstructure components

	Predictor	Estimate	SE	t	p
Utterance length	Intercept	4.78	1.29	3.69	0.001**
	Group TD	4.42	1.78	2.48	0.02*
	Group DLD	1.22	1.78	0.69	0.49
Syntactic complexity	Intercept	6.67	1.10	6.04	<.001***
	Group TD	4.23	1.52	2.78	0.01*
	Group DLD	0.63	1.52	0.42	0.68
Lexical diversity	Intercept	9.00	0.91	9.88	<.001***
	Group TD	0.70	1.26	0.56	0.58
	Group DLD	-1.30	1.26	-1.04	0.31
Story length	Intercept	8.87	0.97	9.91	<.001***
	Group TD	1.93	1.34	1.44	0.16
	Group DLD	0.43	1.34	0.32	0.75

Note: For task, the bilingual ASD group was the reference level. Exposure = length of exposure to English measured in months; * =significance code when $p < .05$; ** =significance code when $p < .01$; *** =significance code when $p < .001$.

Cite this article: Govindarajan, K. and Paradis, J. (2022). Narrative macrostructure and microstructure profiles of bilingual children with autism spectrum disorder: differentiation from bilingual children with developmental language disorder and typical development. *Applied Psycholinguistics* 43, 1359–1390. <https://doi.org/10.1017/S0142716422000406>